THE UNIVERSITY OF SYDNEY

Report 4

Improving understanding of the effectiveness of the EQUIPS programs: The relationship between program dosage and reoffending

Authors: Marlee Bower, Lexine Stapinski, and Emma Barrett

Contact

The Matilda Centre for Research in Mental Health and Substance Use

Level 6, Jane Foss Russell Building,
G02, Camperdown NSW 2006

Marlee.Bower@sydney.edu.au

www.sydney.edu.au/matilda-centre/

CRICOS 00026A

Acknowledgements

We would like to thank the members of our Aboriginal Reference Group for this project, Dr Michael Doyle, Ms Louise Lynch and A/Prof Peter Malouf for their valuable conversations and advice prior to starting this project which influenced the thinking behind this first report. Thanks also to Corrective Services NSW for commissioning this research and Dr Mark Howard for his invaluable feedback, edits and input.

Thanks to Amelia Manks for her assistance with compiling the literature review.

Cite as: Bower, M., Stapinski, L., & Barrett, E. (2020) Improving understanding of the effectiveness of the EQUIPS programs: The relationship between program dosage and reoffending. The Matilda Centre for Research in Mental Health and Substance Use.

THE UNIVERSITY OF SYDNEY

Table of Contents

Scope and Summary

This report is the fourth in a series of reports produced by the Matilda Centre for Research in Mental Health and Substance Use (University of Sydney) to examine implementation processes associated with the Explore, Question, Understand, Investigate, Practice, Succeed (EQUIPS) programs run by Corrective Services NSW (CSNSW). This report examines the relationship between dosage of EQUIPS program sessions and reoffending outcomes. The findings of this report are subject to finalisation an remain unpublished to date. The report is intended for internal review at this stage.

In order to maximise methodological control and rigor, analyses were restricted to offenders who were referred to and participated in EQUIPS programs during their first custodial episode (N=4345). In this sample, 22.8% (n=990) completed one EQUIPS program and 5.5% (n=239) completed two programs. Reoffending was defined as any proven offence of a 'serious' nature occurring in the first 12 months post release.

Key findings of this report are listed below.

- Completing one EQUIPS program was associated with a 24% decrease in the likelihood of serious reoffending during 12 months post-release, compared to those who did not complete an EQUIPS program. These observed benefits were associated with program *completion* rather than participation.

- There were no significant differences in reoffending outcomes when comparing those who had completed two EQUIPS programs to those who had completed one program

- Compared to medium and med-high risk offenders, for high risk offenders there was a greater increase in risk of reoffending associated with increases in modular treatment dose (i.e. two vs one modules completed)

- There was no evidence that certain combinations of two programs were more effective at reducing reoffending when compared to those who completed only one EQUIPS program

Background

Since 2015, the EQUIPS suite of programs, described in detail in Report One, Two and Three, has been one of the main Corrective Services NSW (CSNSW) strategies to reduce reoffending amongst offenders who receive custodial or community sentences. The EQUIPS suite of programs was developed by CSNSW as a form of correctional intervention for offender rehabilitation in both custodial and community settings (Juarez and Howard, 2018). These programs were developed in reference to the Risk-Needs-Responsivity (RNR) model, primarily to aid offenders in practicing strategies for reducing antisocial behaviour and promote prosocial behaviour.[1] Within the RNR model, offender treatment programs try to address offenders' criminogenic 'needs' which are aspects of risk factors which are amenable to change and are identified as having contributed to an offender offending in the first place (Mann et al., 2010). According to this model, offending behaviour will likely continue to occur if criminogenic risks/needs are not addressed and changed. Following this RNR framework, CSNSW determines which offenders require treatment and the type of treatment they require according to a needs assessment based on each offender's:

- current and future risk of reoffending (risk principle), where higher risk offenders are prioritised for interventions
- criminal and antisocial behavioural needs (needs principle), and
- responsiveness and capacity to engage in treatment (responsivity principle) (Andrews et al., 2011)

The EQUIPS suite is grounded in a CBT framework and is comprised of four programs:

- ***EQUIPS Foundation:*** A general program that is not specific to offence type, presented as a standalone intervention for general offending behaviour, or as a precursor for participation in other programs. It aims to introduce the offender to rehabilitative interventions, reduce generalised risk of reoffending and increase participation in prosocial opportunities.[2]
- ***EQUIPS Addiction:*** A program offering support for participants to minimise addictive behaviours. This program is split between group processes and self-management tasks with a focus on aligning skill development to their personal experiences.[3]
- ***EQUIPS Domestic Abuse:*** This program encourages offenders to accept responsibility for their intimate partner violence and abuse offence-related behaviours. There is a focus on increasing their level of accountability to minimise future behaviours. This program is based on a psycho-behavioural framework with a strong therapy-based delivery.[4]
- ***EQUIPS Aggression:*** A program focused on increasing participants' behavioural control and their ability to manage negative life experiences. This approach focuses on the direct and peripheral causes of aggressive behaviour in an attempt to minimise future aggression.

Each program has five modules comprising four sessions of two hours each (or 40 hours combined).[5] While each program can be delivered as a standalone intervention, offenders can be referred to multiple EQUIPS programs if they are deemed to have different needs to be addressed, and even repetitions of the same program, if it is determined that they require extended 'treatment dosage' or maintenance of therapeutic benefit, in accordance with their identified criminogenic needs and case management pathway.

---

[1] See New South Wales Corrective Services Compendium pp. 16.
[2] Ibid.
[3] Ibid.
[4] Ibid.
[5] See CSNSW policy for EQUIPS delivery (for internal use only).

Definitions of dosage in offender treatment programs

As the RNR model suggests, offenders who pose the highest risk of recidivism have the greatest potential to benefit from treatment and should be prioritised for offender treatment programs (Makarios et al., 2014). Existing literature has tended to define offender program dosage in two ways; either focusing on the *number of days* offenders have spent in treatment, or more commonly, the *number of treatment hours* an offender has completed (treatment hours are usually grouped in increments of 100 hours). Using the prescribed methodology of treatment hours, Sperber et al. (2013) examined the interaction between level of dosage and criminogenic risk in an U.S. Community-Based Correctional Facility. Sperber et al. (2013) showed that increasing program dosage according to level of risk resulted in reductions in recidivism. For the low- and medium-risk offenders, there was a 9% reduction in recidivism when their dosage was increased from low and medium dosage, to medium and high dosage, respectively. For high-risk offenders, recidivism dropped from 81% to 57% when their treatment dosage was higher than 200 hours compared to below 200 treatment hours. While recidivism rates were reduced slightly for low- and medium-risk offenders, the strongest effect was seen in higher risk offenders, providing support for matching dosage to offender risk.

Sperber et al. (2013) findings were replicated in a later study, with a similar pattern of results found when the defined dosage categories were narrowed to increments of 50 treatment hours, instead of the prescribed 100-hour increments. In this study, Makarios, Sperber and Latessa (2014) also increased the highest dosage to 300+ hours of treatment, rather than the 200+ hours reported in Sperber et al. (2013). Despite these changes in dosage increments, the authors similarly found a relationship between dosage and recidivism, with reductions in recidivism seen when dosage increased for each risk level.

While these results are promising, they indicated an upper limit of dosage effectiveness. Very small decreases (2%) in recidivism were seen in medium-high risk offenders beyond 300 hours of treatment, suggesting a saturation point may exist for those most at risk. Makarios et al. (2014) indicated a saturation point also exists for low-medium and medium-risk offenders. While increasing treatment dose for these offenders from under 100 hours to 150 hours reduced recidivism by 23%, increasing their dosage beyond 150 hours resulted in a 16% increase in recidivism, suggesting there is also not a linear relationship between dose and reoffending for these lower risk levels.

Studies defining dosage as number of days or months in a treatment program, rather than treatment hours, have demonstrated a similar trajectory of recidivism. Studies have indicated that a high dosage does not need to be classified as 200 or 300 hours, as shorter treatment dosage have been associated with similar reductions in rates of recidivism. Haerle's (2016) study examining dosage in a youth Violent Offender Treatment Program (VOTP) for youth with high-risk of reoffending. Dosage was operationalised in this study by the number of days in VOTP and divided it into three measurements: i) Any dosage; ii) weak dosage (less than 4.5 months in treatment); and iii) strong dosage which (more than 4.5 months in treatment). Offenders were expected to complete three, four-hour weekly group therapy sessions. While the program was designed to last 6 months, the average period offenders spent in VOTP was 4.5 months. The authors considered whether the 4.5 month 'tipping point' for offender participation could be due to issues with program content or non-program factors. For example, the 4.5-month point coincided with the beginning of the more emotionally intense portion of the program, content which could deter some youth from continuing to attend. Alternatively, external disruptions and unforeseen circumstances like release to parole supervision, or transfer to an adult facility due to aging out of the juvenile system, could prompt some participants to exit programs prematurely.

Haerle (2016) first ran a pre-condition analysis and showed that youth treated with any dosage of VOTP were 14% less likely to recidivate during a 3-year follow up. Within the 3-year follow up, 62.1% of youth receiving a VOTP dosage were re-arrested, compared with 68.9% non-VOTP participants. A strong dosage of VOTP reduced the likelihood of re-arrest by 40%. Similar to studies defining dosage through the number of hours completed, Haerle (2016) showed a weak dosage had no effect on recidivism, and that youth receiving no dosage reoffended at a lower rate than those who received a weak dosage. Haerle (2016) provides strong support that serious and violent offenders benefit from higher dosage in treatment.

Factors impacting the relationship between program dosage and reoffending

Feasibility of lengthy programs

While previous studies have predominantly used the number of hours or days in treatment as a classification of dosage, the concept of increasing dosage to 200+ hours of treatment for high-risk offenders comes with limitations. While increasing dosage based on level of risk has demonstrated efficacy in reducing recidivism, Makarios et al. (2014) recognised that the likelihood of an offender engaging in lengthy treatment programs in its' entirety is low. In their study, this notion is reflected by the small number of high-risk offenders undertaking 300+ hours of treatment (n=24). Makarios et al. (2014) acknowledged that the unpredictable nature of the justice system is a significant barrier for the completion of lengthy programs, where transfers, paroles and releases occur sporadically, disrupting program completion.

Program quality

Previous dosage studies have highlighted the quality of the treatment program as a mediator of the risk-dosage mechanism. Sperber and Lowenkamp (2017) discussed the importance of following CBT principles to adequately adhere to the general responsivity principle. They argued that programs offering non-CBT programs at adequate dosage produced reductions in recidivism, while programs offering CBT with insufficient dosage produced increased recidivism. Interestingly, they suggested that programs offering CBT with a sufficient dosage produced the largest reductions in recidivism. In their study, dosage was defined as total number of hours spent in treatment. However, regarding their arguments concerning the *quality* of treatment, increases in weekly CBT role plays was associated with decreases in recidivism. They reported a 25% and 18% reduction in recidivism in medium and high-risk offenders, respectively, when role-plays were used more than 3 times a week (Sperber & Lowenkamp, 2017). This result demonstrates that the quality of treatment is just as important as correct dosage, and that programs should be heavily grounded in CBT.

Simourd and Olver (2019) extend on the notion that treatment quality mediates dosage and risk. They argue that to see any outcome on recidivism, programs should aim to influence the attributes of offenders and their circumstances (Simourd & Oliver, 2019). Essentially, treatment needs to be an appropriate service for its targeted offender. Dosage, whether it is defined as number of treatment hours or days in treatment, only matters if offenders are enrolled in an evidence-based program or suite of programs (Simourd and Olver, 2019). In their study, which included programs focused on 'skill acquisition', reductions in recidivism doubled when programs prioritised criminogenic needs and general responsivity, not just offender risk. This finding further supports the notion that dosage may lose relevance if the 'wrong' treatment is prescribed.

Although the prescribed dosage guidelines of 100, 200 and 300 plus hours of treatment were intended to be a recommendation for managing recidivism, Simourd & Oliver, (2019) expressed concern towards how these guidelines have morphed into 'best practice'. It is clear dosage should not be considered a 'magic number', but should rather be a mix of offender risk, duration, content, and intensity of delivery. A study by Swartz, Lurigio, and Slomka (1996) indicated that when treatment programs incorporated a combination of these mediating factors, similar reductions in recidivism are achieved in shorter treatment periods, providing further evidence that program content, rather than just dosage alone, is important in assessing program efficacy. Although their study did not include criminogenic risk, they examined the efficacy of a prison-based substance abuse program (IMPACT). Overall, recidivism rates were 51%, but when length of stay was taken into consideration, offenders who attended the program for 91 to 150 days had the lowest rates of recidivism (41%). Recidivism rates increased when attendance was higher than 150 days. From this, the authors deduced the optimum dosage for optimum reductions in recidivism rates was 90-150 days, which falls well below the previously recommended threshold of 200-hour dosage.

Participation in treatment versus treatment completion

The available evidence suggests that the conceptualisation of dosage extends beyond time in treatment and is more aligned with attendance in targeted programs, specific to an offender's criminogenic needs. Programs that target these criminogenic factors have shown to have a larger impact on reducing recidivism, however the question arises as to whether these results are attributed to participation or actual completion of programs.

Rahman and Poynton (2018) evaluated the effect of commencing the EQUIPS Domestic Violence Program (DVEQUIPS) on both general offending and domestic violence related offending in high-risk offenders. The DVEQUIPS program involves 20 two-hour sessions (40 hours in total) delivered in weekly modules. In the absence of an RCT design and in order to examine offenders who were similar in terms of eligibility and risk level, Rahman and Poynton (2018) compared recidivism rates of 1,273 offenders who either participated in the program ('starters') to those who were referred to the program but did not start ('non-starters'). Starters are defined as completing four or more sessions. The evaluation indicated that fewer offenders who started DVEQUIPS reoffended *with any offence* 12 months post-release compared to non-starters, with the observed offending rate among starters being 6% less than among non-starters (Rahman & Poynton, 2018). Surprisingly, rates of domestic violence related reoffending between starters and non-starters were not significantly different.

Conversely, Zhang et al. (2019) examined the role program completion plays in conceptualising dosage. Evaluating the same DVEQUIPS program, Zhang et al. (2019) examined outcomes for offenders who were referred to, participated in, and completed the program. Similarly, Zhang et al. (2019) reported that participation in the domestic abuse program resulted in a lower rate of 12.2% for any reoffending in the follow up period, compared to offenders who were referred to the program. No significant reductions in domestic violence related offences were observed. Interestingly, offenders who completed the program demonstrated significantly lower levels of general and violent reoffending (30% and 25%, respectively) compared to those who participated (Zhang et al., 2019). Compared to untreated offenders, those who completed the domestic violence program were 82% less likely to reoffend, but this did not meet the threshold for statistical significance.

Both studies indicated that participation in the DVEQUIPS was associated with reductions in the likelihood of reoffending as well as lower odds of any violent reoffending within 12 months (Rahman and Poynton, 2018; Zhang et al., 2019). Notably, *completion* of the DVEQUIPS program was associated with the highest reduction in reoffending in both general offences as well as violent and domestic abuse offences. These results offer support for the use of program completion, rather than just treatment participation, in the conceptualisation of dosage.

EQUIPS and dosage

The RNR model (Andrews and Bonta, 2010) provides an important framework for delivering behaviour change interventions, like EQUIPS, to offenders by articulating *what* should be delivered, *who* it should be delivered to, and *how* it should be delivered. From an implementation perspective, the risk principle also raises important implications for the *quantity* of intervention that should be delivered, or 'treatment dosage'. The risk principle largely addresses considerations of treatment dosage in relative terms, whereby higher risk offenders are prioritised for delivery of units (e.g. sessions or hours) of intervention, and receive a greater intensity of units of intervention, relative to lower risk offenders. Effective implementation of interventions also requires consideration of dosage in absolute terms, or the amount of units of intervention needed to promote behaviour change and achieve outcomes such as reduction in risk of reoffending.

Previous research evidence supports moving away from a prescriptive 200+ hours as the optimal treatment does for high-risk offenders and moving instead towards a conceptualisation of dosage that incorporates program content, quality, and completion. In line with this evidence, the EQUIPS suite of programs were developed to target behaviour change within a modular design, where offenders are flexibly referred according to individual needs to participate in, and complete, up to four programs of differing focus and content. Offenders can be referred to one or a combination of 2 or more programs, comprising a foundational program (Foundation) for general offending, as well as specialised programs from Aggression, Addiction and Domestic Abuse. The time spent in treatment for each of the EQUIPS programs consists of 20 sessions amounting to 40 hours. Typically, offenders are referred to either 1 or 2 programs, with a total dosage amounting to either 40 or 80 hours. EQUIPS program completion is defined as completing 80% or more of any EQUIPS program content

An examination of the relationship between module dosage of EQUIPS and reoffending outcomes, and how this may be moderated by offender risk, will help to inform an understanding of CSNSW's current model for delivering interventions to offenders. It is important that such an examination consider specific individual and operational (system-level) factors that may play a role in this relationship (such as those identified in Reports One, Two and Three).

Aims

This report seeks to answer the following questions

1. Is increasing dosage, defined as completing a greater number of EQUIPS modules, associated with reductions in reoffending?
2. Does the relationship between EQUIPS dosage and reoffending vary as a function of severity of the offender's reoffending risk?
3. Are particular EQUIPS module combinations more effective than others?

Method

Participants

A study of administrative data obtained from the Offenders Integrated Management System (OIMS)and offending data held in the Reoffending Database (ROD) at the NSW Bureau of Crime Statistics and Research (BOCSAR) was conducted to explore the impact of modular dosage of EQUIPS programs on likelihood of reoffending post-release from supervision. Data was obtained for all adult offenders managed by CSNSW who had been referred to an EQUIPS program in custody or in the community between 2 January 2015 (the implementation of EQUIPS) and 31 December 2018. This resulted in a total of 61,459 referrals to EQUIPS programs attributed to 18,963 unique offenders. The target sample was defined by their referral to any of the EQUIPS programs in custody. Relevant search functions applied to the CSNSW Offender Information Management System (OIMS) were used to identify the sample of offenders. Ethical approval to conduct this research was obtained from CSNSW, The University of Sydney HREC (2019/730) and the Aboriginal Health and Medical Research Council (AH&MRC) HREC (1560/19).

There is complexity involved in measuring the efficacy of programs for those who are referred within custody and then go on to participate in community settings (often whilst on parole). This is because the likelihood of reoffending is greater in the period immediate on release, meaning that likelihood of participating in/completing programs amongst this group is often confounded with a reduced likelihood of recidivism. Similarly, there are complexities for the sample referred via the community that make it difficult to define an appropriate comparison group and address the research questions rigorously. In view of these and other considerations, it was decided that this study would investigate only those referred to, and participating in, programs within custodial settings.

Due to these methodological considerations, the sample for this study was limited to offenders who were referred to and participated in EQUIPS programs whilst in their first custodial episode. Offenders were excluded from analysis if they; i) were referred to and completed programs in the community; ii) were referred to programs in custody and then completed them post-release or on parole in the community; iii) if an offender was referred through a custodial staff member but did not have a custodial period associated with their referral; iv) were not in custody when the referral occurred. Program referrals and participations that occurred during participant's subsequent CSNSW episodes (e.g., if someone reoffended post-release and returned to custody) were also excluded. In order to give offenders adequate time to participate in and complete programs (approximately 10 weeks) that they were referred to in custody, offenders who received their first referral less than 10 weeks (or ≤69 days) before the end of their custodial episode were removed from analysis. This left a total offender sample referred to EQUIPS of 4345.

Design

For the purposes of this study a quasi-experimental design was employed. While it would have been preferable to run a RCT with treatment and control groups, this was not possible due to methodological and ethical practicalities related to the nature of the population/study.

For measurement of program completion in this study, only offenders who had completed at least 80% of their total EQUIPS sessions were deemed to have completed the program, that is attending at least 16 of 20 EQUIPS sessions during their index custodial episode. This fits with the EQUIPS manual requirements that 80% of EQUIPS program sessions are needed for a

complete dose. Sessions were considered more representative participants' engagement with course content than hours elapsed in treatment, which may incorrectly detect repeated sessions as course progression. The comparison group was defined as those who were referred to the EQUIPS program but did not participate. In comparing outcomes between the treatment and comparison groups there is a risk of self-selection bias due to the fact that certain types of participants are more likely to be motivated to complete EQUIPS programs they are referred to and potentially also more motivated to not reoffend. In order to counter this of this risk of bias, inverse probability weighting techniques, described further in the Analysis section below, were used to balance the two groups in terms of personal covariates.

The design allowed for comparison of different treatment outcomes, including

i) those who completed one compared to zero EQUIPS programs, and
ii) those who completed two compared to one EQUIPS programs.

The analyses exploring whether completing 2 compared to 1 EQUIPS programs help to overcome some of the self-selection bias that is a problem in the 1 vs 0 analysis. This is because unlike in the case of completing 1 vs 0 programs, comparing those who completed 2 vs 1 programs mean we are potentially comparing groups that are more equivalent in terms of being motivated to do programs.

Measures

The administrative data were obtained from the CSNSW Offenders Integrated Management System (OIMS). This includes the Level of Service Inventory – Revised (LSI-R; Andrews & Bonta, , 2001), an actuarial risk assessment tool designed to classify an offender's risk of reoffending and identify criminogenic needs. The LSI-R has 54 items grouped into 10 subscales: Criminal History, Education/Employment, Finances, Family/Marital, Accommodations, Leisure/Recreation, Companions, Alcohol/Drug, Emotional/Personal, and Attitude/Orientation. LSI-R total scores are generally used to predict recidivism (risk) whereas subscale scores are used to identify criminogenic needs.

Data was linked to offending data held in the Reoffending Dataset (ROD) at the Bureau of Crime Statistics and Research (BOCSAR). This included all NSW criminal court appearances and outcomes for a given offender in the first 12 months of free-time after release from custody (with or without parole). The ROD data used in this study contains all NSW court appearances finalised by September 2020. In line with previous research evaluating the EQUIPS Domestic Abuse program, ***reoffending was defined as any 'proven offence' of a 'serious' nature with an offence date occurring in the first 12 months of 'free-time' according to court finalisation data.*** Serious reoffending was chosen over a broader general reoffending outcome in an effort to exclude more 'trivial' offences that may be more reflective of circumstances rather than serious offending (e.g., breaches of parole). Additionally, the current NSW Premiers Priority to reduce reoffending amongst the state's most serious offenders and that the EQUIPS programs are designed to target more serious behaviours, such as aggression, violence, and drug offences. In this study, serious reoffending included, to the exclusion of all other offences, homicide and related offences, acts intended to cause injury, sexual assault and related offences, abduction and kidnapping, robbery, extortion and related offences, unlawful entry with intent, theft and related offences, fraud, deception and related offences, the import, export, dealing, trafficking, manufacture and/or cultivation of illicit drugs (NSW Government, 2020).

All other relevant variables are listed and described in Table 1 in the Appendix.

Data analysis and model specification

Descriptive statistics and graphical representations of data are provided to illustrate EQUIPS program completion patterns and their relationship with reoffending post-release. Research questions were addressed through four stages of analysis and all analyses were conducted using STATA version 14.

Analysis One

The first stage aimed to estimate effects on reoffending amongst offenders who had been referred to EQUIPS who had completed one EQUIPS program ('completers') compared to those who did not participate in or complete any programs ('non-participators'). Because we used a modular definition of program dosage which conceptualises dosage as the number of programs completed, offenders who started but did not complete programs ('non-completers') were excluded in analysis. Subsequently, a sensitivity analysis was carried out to assess whether excluding 'non-completers' led to a biased estimation of treatment effect. In this analysis reoffending outcomes were compared between 1) a combined sample of program completers and non-completers, and 2) non-participators.

Inverse Probability Weighting (IPW) to balance differences related to self-selection into groups

Determining the effects of therapeutic treatments on offending outcomes in observational data (where participants are not randomised but self-select into treatment and non-treatment groups) needs to consider between-group differences in participant characteristics that may bias the results. As EQUIPS programs are voluntary, offenders who choose to complete programs could differ from those who do not complete in ways that confound the treatment effects. For example, people who are younger may be less likely to attend an EQUIPS program when referred, but may also be more likely to reoffend post-release, which would make them more likely to reoffend regardless of program completion, confounding the effect of program completion on reoffending outcomes. When random assignment to treatment or control conditions is not possible, ethical or feasible, other statistical tools such as Inverse Probability Weighting (IPW) can overcome the inherent selection bias in whether an offender engages in treatment or not and isolates the effects of treatment from the characteristics of participants by balancing covariates between treated and comparison groups (Garrido et al., 2014). Weighting techniques used to 'balance' the treatment and comparison groups are preferred over other similar methods, like propensity score matching, as they allow for preserving a large majority of the study sample needed to maximise precision in estimating treatment effects (Desai and Franklin, 2019).

In IPW, data are weighted according to the inverse of the probability of the offender being assigned to the observed treatment group given their personal characteristics. In other words, an IPW approach 'balances' the treatment group (those who completed EQUIPS) and comparison groups (those who were referred but that never commenced EQUIPS) in terms of characteristics that may predispose them to either engage or not engage in EQUIPS. We selected variables to balance the groups based on our findings about factors associated with participation and completion of EQUIPS programs in our previous research (Bower et al., 2020a; 2020b; 2021) and based on existing theory about recidivism and program engagement (Hanson, 2002, Crites and Taxman, 2013, Olver et al., 2011). We then tested these variables to make sure they were useful to include statistically, based on model statistics on whether they assisted to balance the treatment and comparison groups.

The final list of covariates chosen were:

- Length of index custodial episode
- Socioeconomic status (SEIFA IRSD score)
- Aboriginality
- Accumulated conviction count
- Cumulative time in prison
- Age at referral
- LSI-R Risk (as most recently assessed at referral date)
- Number of referrals to EQUIPS during index custodial episode

The *pscore* Stata command was used to create a single propensity score using the selected covariates. This syntax checks whether the resulting propensity score has an adequately similar distribution (or 'balance') between treated and comparison groups (Garrido et al., 2014). The *pscore* command also checks whether propensity scores produced is equivalent in the treatment and comparison groups within each of the five quintiles. If they are not equivalent, then some of the quintiles are split into smaller blocks. In cases where balance cannot be achieved in smaller blocks, the covariates included in the propensity model were re-evaluated until satisfactory balance between groups was achieved.

The *qui dr* command was used to transform the propensity score into a IPW weight (Garrido et al., 2014). The *teffects ipwra* package was used to calculate a treatment effect and the overlap assumption in the propensity scores between treatment and comparison groups, i.e., whether there is overlap in the likelihood or probability for participants in the treatment and comparison groups of getting each treatment level (Stata, 2021). The *ipwra* command allows for a doubly robust analysis, meaning additional covariates above propensity-score weighting can be added to predict the reoffending outcome.[6] We used the Average Treatment Effect (ATE), from which estimates can be interpreted as the effect of the treatment when the whole study population is treated with the treatment under investigation (completing EQUIPS) versus the reference treatment (never commencing EQUIPS) (Desai and Franklin, 2019).

In some instances, the *teffects ipwra* command was not suitable for examining the outcomes of interest, for example when, i) when the predictor or outcome was binary , ii) a predictor variable had more than three response options or iii) an interaction effect between treatment type and another covariate needed to be tested. In these instances, a logistic regression was carried out, weighting the analysis by the IPW weight developing using the *qui dr* command and controlling for the same 'treatment model' covariates used in the *teffects ipwra* analysis above. Data driven *bfit* command in Stata *(Cattaneo et al., 2013)* was used to derive the best combination of covariates that predicted reoffending and was included as a covariate within 'doubly robust' models. In the *teffects* model these included Aboriginality, age, SEIFA IRSD score, accumulated conviction count and cumulative time in prison. In the logistic regression models, which could handle more

[6] There has been debate about whether Inverse Probability Models should use 'doubly robust' methodologies. On the one hand, there are arguments that 'treatment status' should be the only independent variable used to derive treatment effect estimates because covariates are expected to be balanced in the weighted population, as per Desai & Franklin (2019). On the other hand, 'doubly robust' models allow for unique covariates associated with program completion in the same analysis which ensures that only one of the models – the treatment model or the outcome model – needs to be specified correctly to obtain correct estimates of treatment effect. Data driven *bfit* command in Stata was used to derive the best combination of covariates that predicted reoffending and was included as a covariate within 'doubly robust' models.

complicated modelling, several extra covariates were included in the model based on data-driven 'best fit' analyses (using the *bfit* command): including an interaction between age and age (age^2), an interaction between age and length of custodial episode, an interaction between age and SEIFA IRSD score and an interaction between age and accumulated conviction count.

Analysis Two

The second stage aimed to explore whether incremental increases in completion of EQUIPS modules have additive effects on reoffending outcomes, relative to the effect sizes observed in the first stage. The primary analysis involved comparing offenders who were referred to and completed two EQUIPS programs, to offenders who only completed one program, controlling for the number of referrals they received. Insufficient sample size meant that we were unable to compare reoffending outcomes of those referred to and completed three EQUIPS programs vs. offenders who completed two programs.

An advantage of this analysis is that it may be less susceptible to systemic selection biases in terms of offender characteristics that may be associated with willingness or capacity to take up referrals to the EQUIPs programs. Nonetheless, it was expected that offenders in the treatment vs comparison group would differ on some characteristics, and hence inverse probability weighting was similarly applied to this analysis, in order to balance the groups in terms of offender characteristics. These analyses enabled us to address core questions about the relationship between incremental increases in EQUIPS module dose, and the reoffending outcomes that these programs are intended to reduce.

Analysis Three

The third stage aimed to examine whether treatment effect associated with incremental increases in EQUIPS module dosage vary as a function of offender's LSI-R risk level. A logistic regression was carried out, incorporating the covariates and inverse probability weights developed in Analysis Two. The primary research question about moderation of treatment effects according to criminogenic risk was examined by main effects of treatment completion and an interaction term for categorical LSI-R Risk assessment (with 'medium' risk as the reference group, to 'med-high' and 'high' risk) and treatment group (Completing one program vs two programs).

Analysis Four

The final analysis explored the moderation effect of the types of program combinations involved, relative to those who only completed one EQUIPS program. These analyses examined the incremental benefit associated with different combinations of two programs. Each set of two-program combinations were tested, including:

- Foundation and Addiction
- Foundation and Domestic Abuse
- Foundation and Aggression
- Addiction and Aggression
- Addiction and Domestic Abuse
- Aggression and Domestic Abuse

Program combinations were dummy-coded to examine relative effects on subsequent reoffending outcomes. A logistic regression was carried out, with reoffending as an outcome, incorporating the covariates and inverse probability weights utilised in Analysis Two with the dummy coded treatment combination predictor as the treatment variable.

Results

Out of the 4345 offenders in this custody referral dataset, 3078 (70.8%) never went on to participate in the program (defined as attending at least one session). Almost a quarter (22.8% or n=990) completed one EQUIPS program; 5.5% (n=239) completed two EQUIPS programs and 0.9% (n=38) completed three EQUIPS programs. Within the first custodial episode for which offenders received an EQUIPS referral, participants received between 1-7 referrals. Table 2 reports the frequency and proportion of i) EQUIPS referrals received and ii) EQUIPS program completions for offenders during the first custodial episode that they received a referral.

Table 1 Frequency and proportion of EQUIPS referrals and program completions

	Referrals		Completions	
Number	Number of participants (n)	Proportion of participants (%)	Number of participants (n)	Proportion of participants (%)
0	0	0	3078	70.8
1	1675	38.6	990	22.8
2	1507	34.7	239	5.5
3	932	21.5	38	0.9
4	195	4.5		
5	26	0.6		
6	7	0.2		
7	3	0.1		
Total	4345	100	4345	100

Table 3 illustrates the specific programs completed by the participants who completed one program. The most commonly completed program was Addiction, with close to half of participants completing this program.

Table 2 Types of EQUIPS programs completed amongst offenders who completed one program

Program type	Number of participants (%)
Addiction	484 (48.9%)
Foundation	261 (26.4%)
Aggression	170 (17.1%)
Domestic Abuse	75 (7.6%)
Total	990 (100%)

Table 4 depicts specific program combinations completed among offenders who completed two programs. The most common combination was Foundation and Addiction (41.8%).

Table 3 presents program type data amongst offenders who completed two EQUIPS programs

Program combination	Number of participants (%)
Foundation and Addiction	100 (41.8)
Addiction and Aggression	81(33.9)
Foundation and Aggression	24 (10.0)
Addiction and Domestic Abuse	22 (9.2)
Foundation and Domestic Abuse	6 (2.5)
Aggression and Domestic Abuse	3 (1.3)
Total	236 (98.7*)

*Note. *Percentages do not add up to 100% because three participants completed the same program twice and were excluded from this table.*

1) Is increasing dosage, defined as completing a greater number of EQUIPS modules, associated with reductions in reoffending?

Analysis one

Of those who completed one program through the custodial pathway (n=990), 24.5% reoffended within 12 months post-release. Of those who were referred but did not complete programs in the custodial pathway (n=3078), 30.4% reoffended within 12 months. An inverse probability weighted regression analysis was conducted, where covariates listed in Appendix Table 1 below were balanced between treatment groups. A Chi Squared chi2(9) = 10.32, *p*=.33, overidentification test for covariate balance showed that we cannot reject the null hypothesis that the IPW model balanced all covariates. That is, there was no evidence that the groups remained unbalanced on the covariates after the weighting procedure. The standardized weighted and unweighted mean differences can be observed in Appendix Table 5. The test of distributional overlap between the variables showed ample overlap between the two groups, as can be observed in Appendix Figure 1. This is a good thing, as it means the groups had become quite similar through the weighting procedure. A binary logistic model was carried out, testing program completion and known covariates on reoffending, which was statistically significant (LR Chi2 (11)=329.68, *p*≤.0001). Controlling for all other covariates, ***having completed one EQUIPS program was associated with a 24% decrease in the likelihood of reoffending in the 12 months post-release, compared to those who did not complete an EQUIPS program*** (robust SE: .07, *p*=.003, 95%CI .63, .91).

A sensitivity analysis tested whether beneficial effects of treatment would be observed within an intention-to-treat framework, where the treatment group was defined as participants who commenced treatment, irrespective of whether they completed. Using this approach, the effect of treatment vs no treatment was not significant (2% reduction in offending; *p*=.10, 95%CI -.05, .00), accounting for all covariates. This result indicates that ***the observed benefits on recidivism are associated with completing the treatment program, rather than participation per se.***

Analysis two

Overall, of those who completed two EQUIPS programs in custody (n=239), 28.5% (n=68) went on to seriously offend in the 12 months post-release. Of those who completed one EQUIPS program in custody (n=990), 24.6% (n=243) went on to seriously offend in the 12 months post-release. To assess the incremental effects of increased modular dose, a binary logistic regression was carried out comparing offenders who completed two EQUIPS programs compared to offenders who completed one EQUIPS program whilst in custody (model statistics: Chi2(9) = 57.26, *p*≤.0001, n=1143). Participants were inversely weighted on their propensity to complete programs and other relevant covariates (see Table 6 and Figure 2, Appendix). Covariates that were associated with likelihood of serious offending were controlled for within the model (i.e., those listed in Table1). ***Completing two EQUIPS programs compared to completing one EQUIPS program was associated with an 1% decrease in the odds of serious reoffending in the 12 months post-release and this difference was not statistically significant*** (*p*=.959).

2) Does the relationship between EQUIPS dosage and reoffending vary as a function of severity of the offender's reoffending risk?

Analysis Three

The third stage examined whether treatment effects associated with increased EQUIPS module dosage vary as a function of offender's LSI-R risk level. A binary logistic regression was conducted which repeated Analysis Two, adding in the main effect of LSI-R and an interaction term between LSI-R and program completion. There was only one significant interaction effect found between treatment and LSI-R ratings (*p*=.034, 95%CI, 1.12, 19.89). Compared to medium and med-high risk offenders, ***for high risk offenders there was a greater increase in risk of reoffending associated with increases in modular treatment dose*** (i.e. two vs one modules completed; see Figure 3 below). There was no significant interaction effect amongst those assessed as 'medium-high' relative to those assessed as 'medium' risk (*p*=.878, 95%CI, .45, 2.53).

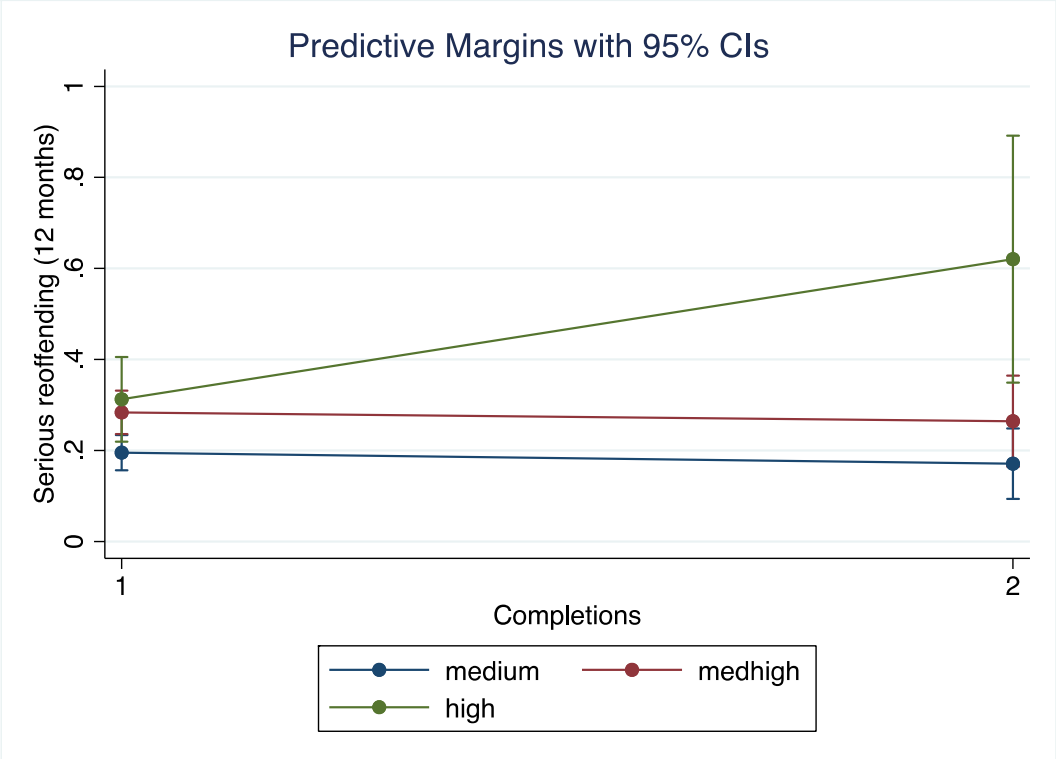

Figure 3. Predictive margins of the interaction effect between LSI-R risk level, and modular treatment dose (i.e.number of program completions) on the likelihood of serious reoffending. Amongst high risk offenders there was a greater increase in risk of reoffending associated with increases in modular treatment dose.

3) Are particular EQUIPS module combinations more effective than others?

Analysis Four

There was no evidence that certain combinations of two programs were more effective at reducing serious reoffending when compared to those who completed only one EQUIPS program (*p*=.14 - .97). The analysis controlled for covariates associated with serious reoffending (see Table 1). Figure 4 below illustrates the marginal reoffending means of different group combinations.

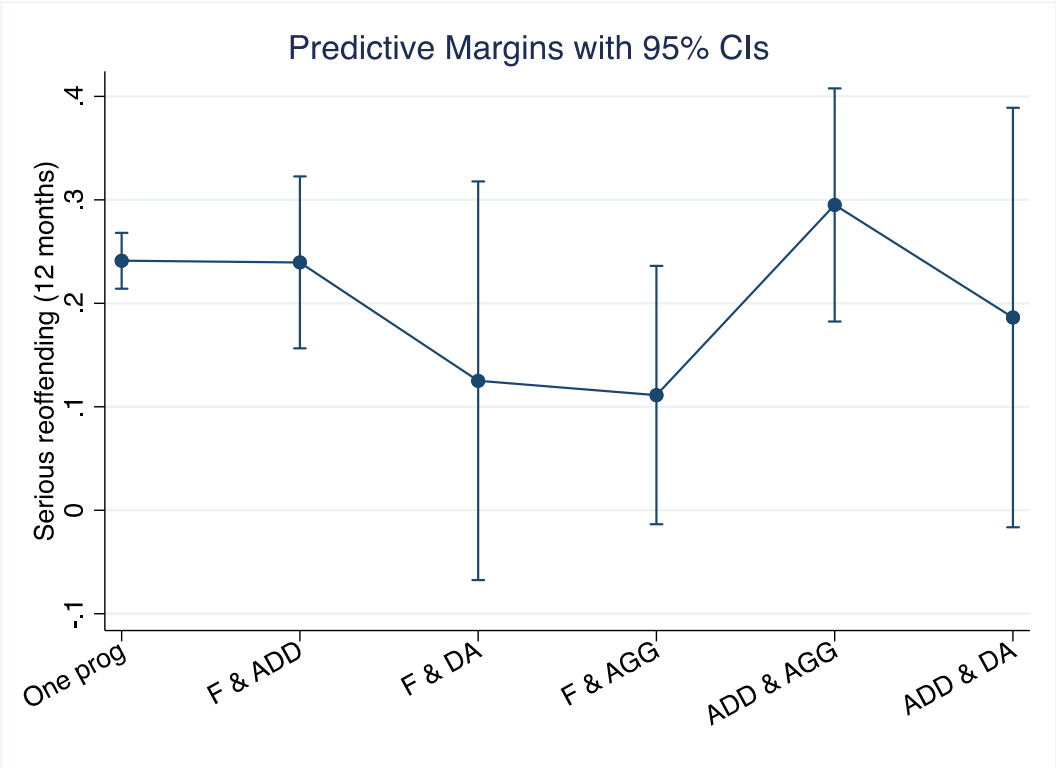

Figure 4. Marginal means for reoffending associated with different EQUIPS program combinations
Note. One prog = Completed only one program; F = Foundation; DA = Domestic Abuse; AGG = Aggression; ADD = Addiction

RESULTS SUMMARY

- In this sample (N=4345), 22.8% completed one EQUIPS program and 5.5% completed two programs. The most commonly completed program was Addiction (48.9%).
- ***Having completed one EQUIPS program was associated with a 24% decrease in the likelihood of serious reoffending in the 12 months post-release***, compared to those who did not complete an EQUIPS program. The observed benefits on recidivism are associated with completing the treatment program, rather than participation.
- ***There were no significant differences in reoffending outcomes when comparing completion of two EQUIPS programs vs. completion of one EQUIPS program***
- Compared to medium and med-high risk offenders, ***for high risk offenders there was a greater increase in risk of reoffending associated with increases in modular treatment dose*** (i.e. two vs one modules completed)
- ***There was no evidence that certain combinations of two programs were more effective at reducing serious reoffending*** compared to completion of one program

Discussion

Among custodial offenders, completing in an EQUIPS program was found to be associated with a reduction in the likelihood of serious offending 12 months post-release. Specifically, those who had completed one program lead to a 24% reduction in reoffending, compared to those who did not complete a program. Sensitivity analysis revealed that the beneficial effect of the EQUIPS program on serious reoffending was associated with completing the program rather than participating, per se. This finding is comparable to that found by Zhang et al (2019), finding that completing EQUIPS Domestic Abuse in community settings was associated with a 25% reduction in the hazard of violent reoffending and a 30% reduction in the hazard of any reoffending compared to matched offenders and accounting for variance in the survival period (Zhang et al, 2019). It is more modest effect than has been found in lengthier offender programs, such as completing ~18 weeks of VOTP (versus 10 weeks of EQUIPS), which reduced the likelihood of re-arrest by 40% at a three-year follow-up (Haerle, 2016). It is important to acknowledge that these findings only apply to custodial offenders on the first custodial episode in which they received a referral to EQUIPS and completed this EQUIPS program whilst in custody. This excludes offenders who complete EQUIPS in the community post-release.

Nonetheless, there was little evidence that increasing dosage in the EQUIPS programs was associated with additive reductions in reoffending. Custodial offenders who completed two EQUIPS programs in custody were no less likely to seriously reoffend than those who completed one program. This is despite this group completing double the 'dosage' of EQUIPS sessions, which evidence suggests should have an additive impact on reducing reoffending (REF). The risk principle of RNR mandates that the higher risk an offender is assessed to be, the more they are likely to benefit from extra program dosage. However, our evidence did not support this relationship with regards to EQUIPS data. In fact, there was a significant interaction effect suggesting that offenders assessed as higher-risk showed an increase in likelihood of reoffending when completing two programs, compared to one program, when compared to medium risk. There are numerous reasons why this may be the case. One is that being referred to, and completing, two EQUIPS programs may be indicative of more complex treatment needs, potentially more than can be addressed by EQUIPS programs, and subsequently higher likelihood

of reoffending. But for this report, the comparison group also received two referrals so were also likely to have higher needs. Other studies have identified iatrogenic effects associated with low-risk offenders partaking in offender programs, but to our knowledge, this is the first study finding a similar effect amongst high-risk offenders. Our statistical certainty around this effect is limited by the small number of offenders who completed two offender programs.

There was also no evidence that completing particular combinations of EQUIPS programs had differential effect on serious reoffending outcomes, compared to completing one program. It is common practice within CSNSW to pair EQUIPS Foundation, the program targeting 'general' offending behaviour, with a more targeted program, such as Aggression, Addiction or Domestic Abuse. However, there was no evidence that such a pairing had any extra benefit on reoffending than pairing two targeted programs. However, statistical power in this analysis was limited by the very small sample sizes in some program pair groups (particularly in relation to Domestic Abuse), potentially limiting the reliability of this result.

Taken together, these results suggest that program 'stacking', the completion of multiple programs as part of a treatment plan does not appear to be the same as increasing 'dosage'. This may explain differences between our findings and others (Haerle, 2016) which provided strong support that serious and violent offenders benefit from higher dosage in treatment. Engaging in a longer program that is specifically targeted to an offender's rehabilitation needs, like the VOTP assessed by Haerle, may be very different to completing multiple shorter programs addressing different needs in the same timeframe. Another potential cause of difference could be because of the difference in program content between 'stacking' and non-stacking models. Different EQUIPS have similar formats and include similar therapeutic techniques (e.g., offence mapping), which may create some redundancy and repetition, potentially reducing efficacy for participants. A singular program is arguably less likely to have this problem. Another potential difference between these cohorts could be related to potential 'treatment group' effects: longer programs allow for development of a supportive therapeutic community amongst program participants, compared to the change in group structure experienced with multiple programs. Qualitative research exploring the experience of participating in two EQUIPS programs is necessary to clarify and understand these effects.

Strengths and Weaknesses

This study had several methodological strengths. It utilised a large dataset of comprehensive administrative data. In order to deal with the complexities in the dataset and many potential confounding factors, including different program referral/participation locations (community vs. custody) and time-frames (within custody vs. on parole), analyses focused on treatment completed in the custodial setting and on serious offending outcomes. This maximised control over factors that would potentially bias the analyses and increased statistical confidence in results. It is also the first exploration of the relationship between all EQUIPS programs and reoffending outcomes.

In our effort to focus analyses and maximise methodological control and rigor, analyses were restricted to a sample of 990 offenders who had completed one program and 239 who had completed two. A larger sample may have revealed other findings (e.g., the effects on reoffending of completing a greater number of modules) and explored associations among offenders who were both referred to and completed programs within community-corrections settings. Other limitations included those related to use of ROD dataset, which only includes offending in the NSW context, meaning our analyses do not capture reoffending that may have occurred in other Australian states and territories. Also, whilst incredibly comprehensive, our use of administrative

data that is not designed specifically for the research use means that we cannot rule out the possibility of administrative errors, such as false or accidental EQUIPS referrals.

Finally, our analysis methodology was a clear strength of this study, allowing us to effectively create a pseudo randomised study by balancing the treatment and comparison groups, in terms of characteristics that may predispose them to either complete or not engage in EQUIPS. However, as with most analysis techniques, the use of IPW did have some limitations. Unlike actual randomised study cohorts, IPW can also create conceptual equivalence or 'exchangeability' between treated and untreated cohorts with respect to measured covariates, meaning there is a risk samples may retain differences driven by unmeasured covariates (Desai and Franklin, 2019), which can bias treatment estimates. This bias. This bias may grow as the relationship between measured and unmeasured covariates increases (Garrido et al., 2014). Given the voluntary nature of EQUIPS programs, offenders who commenced EQUIPS could differ from those who did not commence in ways that cannot be observed/measured in this data. For example, the suitability of offenders to the program, or personal motivation to change behaviour were not measured but may impact both the likelihood of program completion and reoffending.

Policy Implications

Based on the results of this study, there is evidence for the efficacy of EQUIPS programs on serious reoffending and therefore, on the benefit of the continued delivery of this program for eligible offenders. Additionally, results indicated that offenders should be encouraged to *complete* EQUIPS programs as benefits were evident for those who had completed, rather than having participated in EQUIPS programs. There is, however, limited evidence in this study to support the delivery of two EQUIPS programs on reoffending programs, regardless of what specific mix of EQUIPS programs are delivered.

At the time of writing this report, the current NSW Premier's Priority was "to increase program hours to the highest risk offenders to ensure that a greater proportion of these offenders receive the level of treatment or services that evidence suggests is needed to be effective." They note that, they will "increase the average hours of treatment to at least 160 hours for 3000 higher risk offenders exiting prison each year" (NSW Government, 2020). However, our results and past research (Simourd and Olver, 2019) suggest that any extra dosage and support needs to be tailored and targeted towards an offender's individual rehabilitation needs, in line with the RNR concept of 'responsivity'. As noted by Simourd and Olver (2019) reductions in recidivism are maximised when programs prioritise criminogenic needs and general responsivity, not just offender risk, tailored to offender's circumstances, attributes and skills. The relationship between treatment dosage and reoffending can lose relevance if an inappropriate or unspecialised treatment program is prescribed. There is no one-size-fits-all approach to offender rehabilitation and increasing hours of unspecialised programs may not necessarily add extra benefit to offenders' likelihood of recividism.

Conclusions

Completing an EQUIPS program is significantly associated with reductions in serious reoffending. This finding supports existing evidence for benefits to delivering an EQUIPS program in custodial settings to eligible offenders. This study found no evidence for stacking two programs or for particular program combinations adding any extra benefit to reoffending outcomes over and above completing one program, but it is important to note that examination of a larger sample of participants and community samples could reveal more findings in this respect.

References

ANDREWS, D. A. & BONTA, J. 2001. *Level of Service Inventory–Revised (LSI-R): User's manual. ,* North Tonawanda, NY, Multi-Health Systems.

ANDREWS, D. A. & BONTA, J. 2010. Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law,* 16**,** 39-55.

ANDREWS, D. A., BONTA, J. & WORMITH, J. S. 2011. The Risk-Need-Responsivity (RNR) Model:Does Adding the Good Lives Model Contribute to Effective Crime Prevention? *Criminal Justice and Behavior,* 38**,** 735-755.

BOWER, M., STAPINSKI, L., NEWTON, N. & BARRETT, E. 2020a. Identifying the demographic, sentencing and criminogenic features of offenders entering the EQUIPS programs in custody and in the community. Report commissioned by NSW Department of Communities and Justice: Corrective Services NSW.

BOWER, M., STAPINSKI, L., NEWTON, N. & BARRETT, E. 2020b. The individual and operational (system level) factors associated with program participation and program completion among referred offenders to the EQUIPS programs in custody and in the community. Report commissioned by NSW Department of Communities and Justice: Corrective Services NSW.

BOWER, M., DOYLE, M., STAPINSKI, L., NEWTON, N. & BARRETT, E. 2021. Factors associated with EQUIPS program participation and completion for Aboriginal offenders in custody and the community. Report commissioned by NSW Department of Communities and Justice: Corrective Services NSW.

CATTANEO, M., DRUKKER, D. & HOLLAND, A. 2013. Estimation of multivalued treatment effects under conditional independence. *Stata Journal,* 13**,** 407-450.

CRITES, E. L. & TAXMAN, F. S. 2013. The responsivity principle: determining the appropriate program and dosage to match risk and needs. *In:* TAXMAN, F. S. & PATTAVINA, A. (eds.) *Simulation Strategies to Reduce Recidivism.* New York: Springer.

DESAI, R. J. & FRANKLIN, J. M. 2019. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. *BMJ***,** l5657.

GARRIDO, M. M., KELLEY, A. S., PARIS, J., ROZA, K., MEIER, D. E., MORRISON, R. S. & ALDRIDGE, M. D. 2014. Methods for constructing and assessing propensity scores. *Health services research,* 49**,** 1701-1720.

GOVERNMENT, N. 2020. *Premier's Priorities - Data Information Sheet* [Online]. Premier & Cabinet. Available: **https://www.nsw.gov.au/sites/default/files/2020-11/Reducing-recidivism-in-the-prison-population-data-information-sheet.pdf** [Accessed 6/04/2021 2021].

HAERLE, D. R. 2016. Dosage Matters:Impact of a Violent Offender Treatment Program on Juvenile Recidivism. *Youth Violence and Juvenile Justice,* 14**,** 3-25.

HANSON, R. K. 2002. Recidivism and Age: Follow-Up Data From 4,673 Sexual Offenders. *Journal of Interpersonal Violence,* 17**,** 1046-1062.

JUAREZ, T. & HOWARD, M. 2018. Assessing offender change over treatment: The influence of treatment context on self-reported antisocial attitudes. *In:* RESEARCH BULLETIN (ed.). Corrective Services New South Wales: New South Wales Government.

KRONER, D. G. & TAKAHASHI, M. 2012. Every session counts: The differential impact of previous programmes and current programme dosage on offender recidivism. *Legal and Criminological Psychology,* 17**,** 136-150.

MAKARIOS, M., SPERBER, K. G. & LATESSA, E. J. 2014. Treatment Dosage and the Risk Principle: A Refinement and Extension. *Journal of Offender Rehabilitation,* 53**,** 334-350.

MANN, R. E., HANSON, R. K. & THORNTON, D. 2010. Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. . *Sexual Abuse: A Journal of Research and Treatment,* 22**,** 191-217.

OLVER, M. E., STOCKDALE, K. C. & WORMITH, J. S. 2011. A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. *J Consult Clin Psychol,* 79**,** 6-21.

RAHMAN, S. & POYNTON, S. 2018. Evaluation of the EQUIPS Domestic Abuse Program *Crime and Justice Bulletin No. 211.* Sydney: NSW Bureau of Crime Statistics and Research.

SIMOURD, D. J. & OLVER, M. 2019. Prescribed correctional treatment dosage: Cautions, commentary, and future directions. *Journal of Offender Rehabilitation,* 58**,** 75-91.

SPERBER, K. G., LATESSA, E. J. & MAKARIOS, M. D. 2013. Examining the Interaction between Level of Risk and Dosage of Treatment. *Criminal Justice and Behavior,* 40**,** 338-348.

SPERBER, K. G. & LOWENKAMP, C. T. 2017. Dosage is more than just counting program hours: The importance of role-playing in treatment outcomes. *Journal of Offender Rehabilitation,* 56**,** 433-451.

STATA. 2021. *teffects intro advanced: Advanced introduction to treatment effects for observational data* [Online]. Available: https://www.stata.com/manuals/teteffectsintroadvanced.pdf#teteffectsintroadvanced [Accessed 2021].

Swartz, J. A., Lurigio, A. J., & Slomka, S. A. (1996). The impact of IMPACT: An assessment of the effectiveness of a jail-based treatment program. *Crime & Delinquency*, *42*(4), 553-573.

WAN, W.-Y., POYNTON, S., VAN DOORN, G. & WEATHERBURN, D. 2014. Parole supervision and reoffending. *In:* CRIMINOLOGY, A. I. O. (ed.) *Trends & issues in crime and criminal justice.* Canberra: Australian Institute of Criminology.

ZHANG, Y., WEI, Z., HOWARD, M. V. A. & GALOUZIS, J. 2019. Evaluation of EQUIPS treatment pathways for domestic violence offenders in New South Wales. *In:* STATISTICS, C. R. E. A. (ed.) *Research Publication.* Sydney: NSW Government Corrective Services NSW.

Appendices

Table 4. Key variables used in the current report

Variable	Meaning
Defining groups	
Program non-participation	Offenders who were referred to EQUIPS on at least one occasion but did not participate in any EQUIPS programs [source: OIMS].
Program non-completion	Offenders who were recorded by CSNSW programs staff as having attended at least one session of an EQUIPS program that they had been referred to [source: OIMS].
Program completion	Only offenders who had completed at least 80% of their total EQUIPS sessions were deemed to have completed the program, that is attending at least 16 of 20 EQUIPS sessions during their index custodial episode. This fits with the EQUIPS manual requirements that 80% of EQUIPS program sessions are needed for a complete dose. Sessions were considered more representative participants' engagement with course content than hours elapsed in treatment, which may incorrectly detect repeated sessions as course progression [source: OIMS].
Outcome variable	
Serious Reoffending	The NSW Bureau of Crime Statistics and Research (BOCSAR) Reoffending Database (ROD) was used to link all 'serious' finalised NSW criminal court appearances and outcomes for a given offender in the first 12 months of free-time after release from custody (with or without parole). The ROD data used in this study contains all NSW court appearances finalised by September 2020. Serious Reoffending was defined as per the NSW Government definition including, to the exclusion of all other offences, homicide and related offences, acts intended to cause injury, sexual assault and related offences, abduction and kidnapping, robbery, extortion and related offences, unlawful entry with intent, theft and related offences, fraud, deception and related offences, the import, export, dealing, trafficking, manufacture and/or cultivation of illicit drugs(NSW Government, 2020). These were identified using the following ANZSOC codes: 01, 02, 03, 51, 06, 07, 08, 09, 101, 102 and 103. [source: ROD].
Covariates	
1. <u>Individual Factors</u>	
Demographics	
Age at referral	Offender's age at the time that they were referred to a particular EQUIPS program [source: OIMS]
Gender	Gender of participant at time of referral, either male, female or unknown/undetermined. [Source: OIMS]
Culturally and Linguistically Diverse (CALD)	Corrective Services collects data around Cultural diversity (Country of birth) and Linguistic Diversity (English as a second language). In the dataset CALD variable, there were four categories YY, YN, NY, NN. Y's were interpreted as 'yes' and N as 'no' to either Culturally Diverse or

	Linguistically Diverse. In the current study CALD status is defined as those who are *both* Culturally and Linguistically Diverse, or those identified in the dataset as 'YY'. [Source: OIMS]
In a relationship	Relationship status is conceptualised as those who reported being in a married or de facto relationship on entry to custody. Those who were in the following categories were understood as not being in a married or de facto relationship: never married, divorced, widowed or separated. [Source: OIMS]
Socioeconomic status – geographic location of origin	
SEIFA Relative Advantage and Disadvantage score	*Relative socio-economic advantage and disadvantage (IRSAD)* - A low IRSAD score indicates an area that has relatively greater disadvantage in terms of income, occupation, education and even internet connection.
ABS Remoteness Index	Australian Statistical Geography Standard-Remoteness Area is a geographical classification which defines locations in terms of remoteness, i.e., the physical distance of a location from the nearest urban centre and therefore, relative access to major services. Remoteness is categorised in terms of: Major cities of Australia, Inner regional Australia, Outer regional Australia, Remote Australia and Very remote Australia. [Source: OIMS]
Criminogenic Risk-Needs (LSI-R Risk category and subcategories)	
LSI-R Risk Category[1]	CSNSW uses the Level of Service Inventory-Revised (LSI-R) actuarial risk assessment tool to measure offender's criminogenic risk/needs. Total risk scores are calculated by adding risk factor scores. Offenders who score from 0-13 on the LSI-R are considered low risk offenders, 14-23 low-medium risk offenders, 24-33 medium risk offenders, 34-40 recognized as medium-high risk, and 41-54 are considered high risk offenders. Referrals to EQUIPS are recommended amongst those who are deemed to be of medium-high to high risk of reoffending using the LSI-R. This variable includes an offender's recent risk category recorded against their most recent LSI-R assessment. [Source: OIMS]
Historical and Current Most Serious Offences needs	
Higher Conviction Count (over lifetime)	Number of convictions over lifetime, both in custody and community-based sentences. [Source: OIMS]
More time in prison over lifetime	Cumulative time that each offender had spent incarcerated over their lifetimes [Source: OIMS]
Most serious offence associated with index sentence (custody and community)	Australian and New Zealand Standard Offence Classification (ANZSOC) Most Serious Offence (MSO) summarised into categories based on those used in Wan et al. (2014). Categories included serious violent offence, non-serious violent offence, property offence, breach of court order and driving offence. The remaining types of offences were aggregated into an 'other' group.
2. <u>Operational (system-level) Factors</u>	
Employed at time of referral	Indicating offenders were employed in their correctional centre at the time of referral to EQUIPS. Binary responses (yes/no). [Source: OIMS]
Receiving education at time of referral	Indicating offenders were undergoing education in their correctional centre at the time of referral to EQUIPS. Binary responses (yes/no). [Source: OIMS]
Parole attached to sentence	Offenders who had a parole component attached to their community or custodial sentence. Binary responses (yes/no). [Source: OIMS]

Not needing to move centres to complete	An indicator that an offender started attending the program at the same centre at which their referral took place. Binary responses (yes/no).
Months between EQUIPS program commencement and index referral	The number of months between the first EQUIPS referral on 2 Jan 2015 and an inmate's EQUIPS referral. An indicator of the amount of time that had elapsed since the EQUIPS suite of programs commenced. This was chosen as a covariate as it was deemed important to control for possible improvements in program referral and delivery, the longer that a program had been running.
Months between sentence start and first program attendance	The number of months between the start of an offender's sentence (either custodial or community, depending on referral pathway) and the first day they attended an EQUIPS program.
Months between sentence end and first program attendance	The number of months between the first day an offender attended an EQUIPS program and the end of their sentence (either custodial or community, depending on referral pathway).
Program Type	Programs within the EQUIPS suite including EQUIPS Foundation, EQUIPS Domestic Abuse, EQUIPS Addiction and EQUIPS Aggression [Source: OIMS]

Table 5 Standardised mean differences between those who completed EQUIPS programs and those who did not complete any program, in raw and weighted forms.

	Standardised differences	
Covariate	**Raw**	**Weighted**
Custodial episode length	.06	-.00
SEIFA IRSD Score	-.01	.00
Accumulated conviction count	-.03	.01
Cumulative time in prison	-.05	.00
Age at referral	.10	.01
LSI-R Risk Category	.02	.00
Referral count (binned)	.22	-.00
Aboriginality	-.06	.00

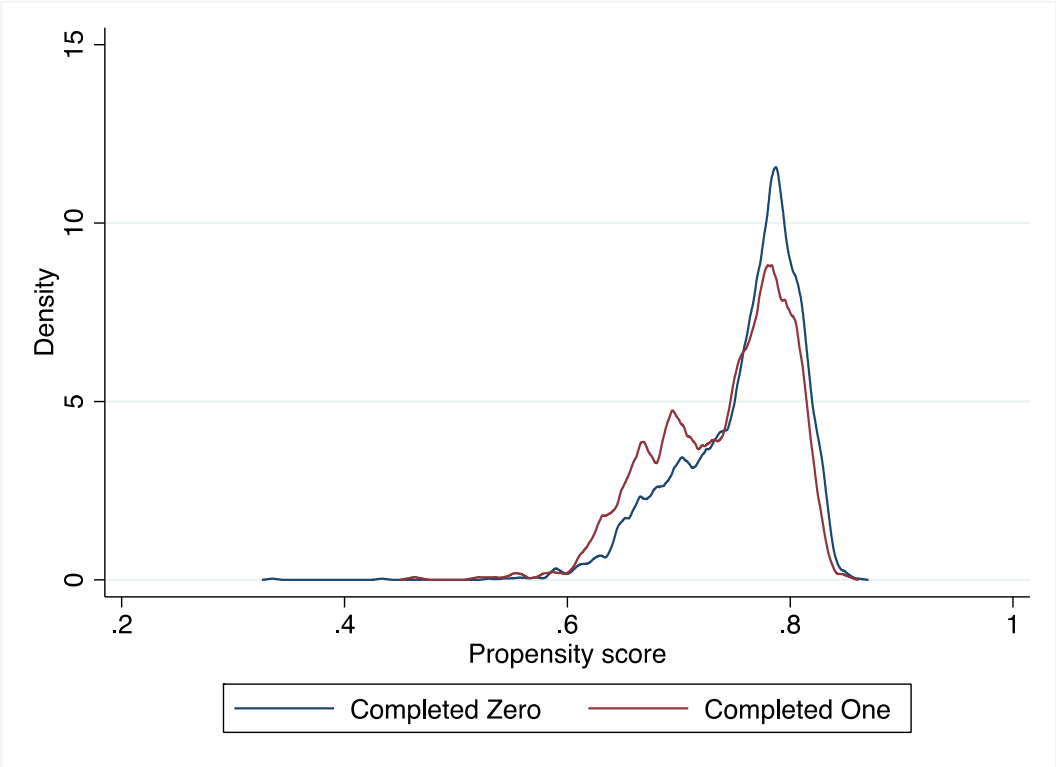

Figure 1. Propensity score overlap plot comparing completers (of one program) and non-participators (those who were referred but did not participate) whilst in custody, after undergoing regression-adjusted IPTW weighting.

Table 6 Standardised mean differences between those who completed two EQUIPS programs and those who completed one, in raw and weighted forms.

	Standardised differences	
Covariate	**Raw**	**Weighted**
Custodial episode length	-.07	-.07
SEIFA IRSD Score	.14	.01
Accumulated conviction count	-.09	.00
Cumulative time in prison	.06	-.03
Age at referral	-.15	-.00
LSI-R Risk Category	-.08	-.04
Referral count (binned)	.35	.00
Aboriginality	-.19	-.02

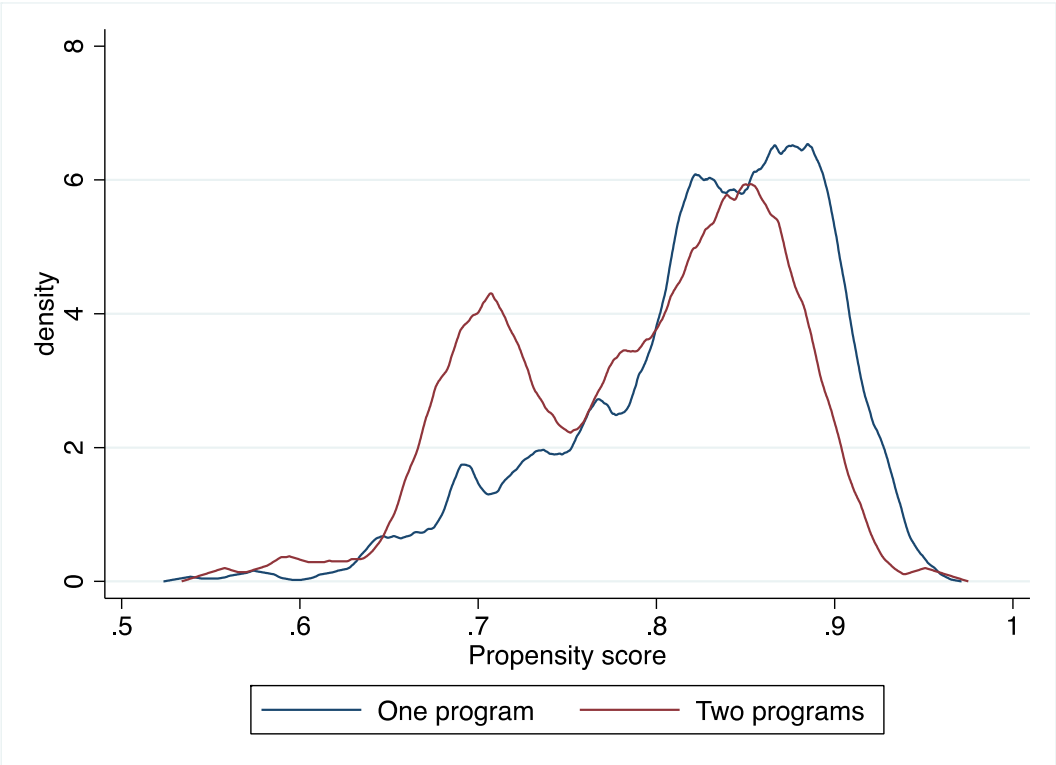

Figure 2. Propensity score overlap plot comparing completers of one program vs completers of two programs whilst in custody, after undergoing regression-adjusted IPTW weighting.