Research Bulletin

Evaluation of the Practice Guide for Intervention (PGI): Assessments of quality assurance and supervisee outcomes

Chee Seng Chong, Sarah Cassidy & Mark Howard

Aims

This study aimed to assess how staff practice as assessed by Interview Observations (IO) and Practice Reviews (PR) may be associated with community supervision outcomes. This study also examined if compliance with these Quality Assurance (QA) activities at the Community Corrections Officer (CCO) and office level may have an impact on outcomes.

Methods

IO and PR outcomes for 761 CCOs who had an active caseload between July 2020 and June 2021 were examined. Descriptive statistics were used to examine the trends in completion and the distribution of scores on these assessments. Supervision outcomes (reoffending and supervision failure) for a sample of 3,818 supervisees who commenced a community supervision episode with the above identified CCOs within the study timeframe were also examined. Separate models were fitted to examine whether IO and PR assessment scores, and compliance with these activities were predictive of supervisee outcomes.

Results

We observed an upward trend in QA assessments completed over the timeframe of study, with close to 25% of CCOs fully meeting mandatory delivery requirements. Compliance with these requirements varied between offices with about half of the CCOs fully meeting requirements across most offices. CCOs generally received high ratings on IOs and PRs with large proportions of CCOs receiving the maximum possible rating, indicating ceiling effects.

We found indications that only IO scores were associated with supervision outcomes. Higher Rapport Building scores were marginally associated with lower odds of reoffending. Conversely, higher Intervention Focussed scores were associated with greater odds of supervision failure. This counterintuitive finding suggests that while proficiencies assessed by QA processes are intended to indicate better staff practice, these do not necessarily translate into improved supervision outcomes. Compliance with QA activities was not observed to be associated with outcomes.

Conclusion

While scores on the QA assessments generally suggest high quality staff practice, the ceiling effects in the data limited the interpretation of our results and hold implications for the utility of the QA processes in identifying and supporting staff professional development. Further studies examining the relationship between CCOs' practice and outcomes, and how quality of practice may be more accurately assessed will be beneficial to the development of further training opportunities aimed at improving service delivery and supervisee outcomes.

INTRODUCTION

The Practice Guide for Intervention (PGI) is a framework of manualised exercises that Corrective Services NSW Community Corrections Officers (CCOs) selectively work through with people under their supervision. These exercises are designed to provide a structured platform to address a range of criminogenic needs and responsivity factors. The PGI was developed based on Risk-Need-Responsivity (RNR) principles (Bonta & Andrews, 2007), which emphasises: 1) aligning the intensity of the program with the supervisee's risk of re-offending, 2) addressing the criminogenic needs of the supervisee, and 3) delivering behaviour change content that is tailored to the supervisee's learning style, capabilities, and motivation. When all three principles are applied, especially in community corrections settings, recidivism reductions of up to 50% have been reported (Andrews & Bonta, 2006).

Since its implementation in 2016, Corrections Research Evaluation and Statistics (CRES) had conducted several evaluation studies focusing on the initial stages of PGI rollout and early outcomes (see Howard et al., 2019 for a summary of these studies and Thaler et al., 2019 on implementation of the PGI). While these studies generally found positive staff perceptions and uptake of the PGI, a number of them highlighted initial implementation challenges associated with program fidelity and quality of service delivery. For example, one study found that CCOs tended to favour more general and process-oriented exercises that were not always aligned with the supervisees' criminogenic needs or responsivity factors (Chong et al., 2020), while another found a large proportion of CCOs identified more with traditional program brokerage roles (enabling access to programs via referrals) over their putative roles as agents of behaviour change under the new PGI model of supervision (Tran et al., 2019). In a more recent survey examining staff use of the PGI as it entered a more stable phase of continuous delivery, Cassidy and colleagues (2023) found that a small percentage of CCOs do not consistently implement PGI exercises during supervisory sessions.

Considering these challenges, it is perhaps unsurprising that an early outcome evaluation study by NSW Bureau of Crime Statistics and Research (BOCSAR) found the PGI had limited impact on reoffending outcomes (Ooi, 2020a and 2020b). It is well established that programs are most successful when staff are well trained in delivering them according to their design intentions (Andrews et al., 1990a; Andrews, et al., 1990b; Cullen, 2002; Gandreau, 1996; Lowenkamp et al., 2006; Chadwick et al. 2015; Robinson et al., 2012). It is therefore of critical importance that CCOs are given opportunities to receive feedback to further enhance skill development and encourage best practice.

In recognition of this, Corrective Services NSW Community Corrections has introduced a number of Quality Assurance (QA)[1] procedures that aim to increase not only the fidelity of PGI delivery but also the quality of relationships between CCOs and their supervisees. These activities include

[1] These procedures are now referred to as Continuous Improvement in Practice (CIP) activities. The term QA was retained to reflect the operational context and method as they applied to the current study.

Interview Observations (IOs) and Practice Reviews (PRs), which provide CCOs with regular opportunities for feedback and professional development. To conduct an IO, an observer (which is usually the Team Leader (TL) of the CCO) completes a checklist assessing how well the officer demonstrated certain skills in a nominated supervision session. The completed checklist then forms the basis of a strengths-based feedback discussion about the officer's strengths and areas for development. Four main skill areas are considered during an IO: Rapport Building, Intervention Focussed, Cognitive Techniques, and Prosocial Modelling.

Administration of a Practice Review (PR) involves a TL conducting a desktop review of the case plan and case notes of up to three supervisees with medium and above risk of reoffending on an officer's caseload. This review is conducted against a checklist of items which assess four core skills: Clear and Professional Records, Initial Case Plan Review, Case Plan Update Review, and Intervention Evidence. The outcomes of the review, including strengths and upskilling goals are discussed with the officer and an action plan is developed to meet the identified goals.

PRs were introduced in 2018 and associated policy indicated that these must occur at least twice a year with all CCOs carrying a supervision caseload. IOs were also introduced in 2018, and from January 2020, it was mandatory for all officers with a supervision caseload to complete one every six months. These requirements were made more stringent as of January 2022 and all officers are currently required to complete two to four Interview Observations every six months.

The current study

Recently, CRES had conducted a series of studies which focused on the QA processes. These studies were intended to inform the nature of further training and development opportunities for CCOs. In a survey exploring staff perceptions, CCOs generally held favourable views of QA procedures, reporting benefits such as improved interviewing techniques and enhanced quality in written case notes (Chong et al., 2023). This survey was followed by a quantitative study which examined how staff proficiencies in applying different practice skills were associated with perceptions of working relationships between CCOs and supervisees, and how this may in turn be associated with supervision outcomes (Chong et al., 2024). This study found limited associations between IO outcomes and perceived quality of working relationships (measured using the Dual Role Inventory – Short Form (DRI-SF: Gochyyev & Skeem, 2019). The study also found complex associations between quality of relationships and supervision outcomes; for example, while more favourable ratings of Toughness in the relationship marginally predicted lower odds of supervision failure, higher ratings of Trust predicted higher odds of supervision failure. The findings of this study were limited, however, by statistical considerations associated with observed ceiling scores and the extent to which the DRI-SF adequately captured the full extent of these relationships.

The current study is aligned with the aims of the previous studies and examines how outcomes of IOs and PRs are associated with supervisee outcomes. It is hypothesised that better quality staff practice, as indicated by more positive

outcomes on QA assessments, may be associated with more favourable supervisee outcomes such as reoffending and supervision failure. The findings of this study are intended to evaluate if the skills assessed and developed as part of these QA processes have a relationship with supervisee outcomes and thereby, providing insights into the utility and validity of these QA processes as best practice staff development tools. Relatedly, as ongoing training is an important aspect of ensuring quality service delivery (Lowenkamp et al., 2012), poor compliance with QA activities may have a negative impact on supervisee outcomes. This study therefore also examined if compliance with QA activities at both the individual CCO and broader Community Corrections office levels have an impact on outcomes.

To achieve these aims, we first examined if there was evidence of increasing PR and IO assessments completed over a one-year period starting from July 2020 to June 2021 and the distribution of scores on these assessments. This study also examined CCOs' and offices' frequency of engagement with these activities against the mandatory requirements of two IOs and two PRs per CCO per year which were in place over the study timeframe. Multilevel logistic regression models (where applicable) were then fitted to examine the relationships between supervisee outcomes and QA assessment scores. The relationships between outcomes and compliance with the mandatory requirements on frequency of QA activities were also modelled.

METHODS

Sample

Between July 2020 and June 2021, a total of 761 CCOs were identified as carrying an active caseload of supervisees. From these CCOs, we identified[2] a total of 1,278 IO Checklists and 736 PRs that were completed within the study timeframe. These QA assessments were examined using descriptive statistics.

A sample of 3,818 supervisees who commenced a community supervision episode with the above identified CCOs within the study timeframe were identified for further analyses. Where multiple community supervision episodes were identified for an individual, the most recent episode was retained for analysis. As not all CCOs who carried an active caseload completed the minimum frequency requirements for QA activities, the final sample included in these analyses varied.

To examine the relationship between QA ratings and outcomes, only CCOs who were identified to have completed at least one IO and PR assessment within the study timeframe were included. This reduced the sample to 1,852 individuals who were supervised by 353 CCOs. The full sample of 3,818 supervisees and 761 CCOs were retained for the models examining the relationship between QA compliance and outcomes.

[2] This study was able to examine only IO checklists which contained legible and sufficient. This also applies to PRs. Initial data diagnostics indicated that a very small percentage of IO checklists and PR assessments had missing values (about 4%).

Data

Self-report and administrative data which included supervisee and staff demographic information were examined in this study. Supervisee demographics were extracted from the Corrective Services NSW Offender Integrated Management System (OIMS) which is the central operational database that maintains a range of information on all people who are managed by Corrective Services NSW. Supervisees' age at the start of the index community supervision episode, age at first conviction, gender, supervision length, Indigenous status[3], and risk of reoffending were extracted for use as covariates. Risk of reoffending was estimated using the Level of Service Inventory - Revised (LSI-R, Andrews & Bonta 2000). The LSI-R is a widely used actuarial assessment tool designed to evaluate an individual's criminogenic needs across ten domains. An aggregate score from these domains is then derived to estimate risk for recidivism which is categorised into five levels: Low, Low / Medium, Medium, Medium / High, and High. Staff demographic data including name and office location where the CCO was based at was obtained from Corrective Services NSW Human Resources Management for data triangulation purposes.

Interview Observation Checklist

The IO Checklist is an assessment tool used by an observer to assess a CCO on their proficiency in utilising four core skills in their interactions with their supervisees: Rapport Building; Intervention Focussed; Cognitive Techniques; and Prosocial Modelling (see Chong et al. 2024 for descriptions of these skills). The IO Checklist consists of 21 items. Five rating options were given to each item: "N – no opportunity to apply skill"; "M – missed opportunity"; "D – skills developing"; "E – skills enhancing"; and "P – skill present and clearly observed".

Practice Review

In performing a PR[4], TLs may review their CCO's work with up to three medium and above risk supervisees against a set of 24 items[5] which assesses four core skills: Clear and Professional Records, Initial Case Plan Review, Case Plan Update Review, and Intervention Evidence.

'Clear and Professional Records' assesses the quality of CCOs' written case notes and include considerations of whether case notes provided sufficient information and was clearly written. 'Initial Case Plan Review' focusses on the first case plan that CCOs develop with individuals who commenced a supervision episode and assesses whether CCOs understood their supervisees needs and that case plans were structured to address those needs. 'Case Plan Update' assesses CCOs' ability to reflect on their work with their supervisees and to continually update case plans so that they remain relevant to their supervisees' needs. 'Intervention Evidence' assesses whether documented case notes demonstrated CCOs' proficiency in encouraging

[3] For the purposes of this report, we use the term 'Aboriginal' to refer to all First Nations Australians including Aboriginal and Torres Strait Islander peoples.

[4] There are different types of PRs which are completed by different staff members. This study examined only Primary PRs which are generally completed by the CCO's Team Leader.

[5] Across the study period, two versions of the PR were used. In 2020, PRs consisted of 25 items while an updated version released in 2021 had 24 items. While the wording of some questions has changed, both versions measured the same core skills.

behaviour change and if they were responding appropriately to their supervisees' responsivity factors and level of engagement.

In each PR, TLs are required to rate only three of the four core skills. TLs may either rate items related to 'Initial Case Plan Review' or 'Case Plan Update' depending on whether the case plan examined was an initial or existing case plan. Six response options were given to each item on the PR. These options were identical to those on the IO but with the addition of "N – no skill evidenced"[6].

Supervision outcomes

Two outcome variables, reoffending and supervision failure, were examined in this study. Reoffending was defined as any new convictions within one year from the start of the index community supervision episode. This was extracted from NSW Bureau of Crime Statistics and Research (BOCSAR) Re-offending Database (ROD), which contain records for the outcomes of every person with a court appearance in NSW. Supervision failure was defined as any revocation or breach of parole orders within one year from the start of the index community supervision episode; relevant data on supervision outcomes were extracted from OIMS.

Compliance

According to service standards, all CCOs with an active caseload of supervisees must complete at least two IOs and PRs per year. Two measures were generated to indicate each office and CCOs' compliance with these standards. A CCO was considered fully compliant if they completed at least two IOs and PRs over the study timeframe. At the office level, full compliance was achieved if every CCO with a caseload based in that office was fully compliant. Compliance scores ranged from 0 to 1, with 0 indicating no compliance and 1 for full compliance. As initial data diagnostics indicated that there were no discernible differences in PR and IO completion rates for those with partial compliance, these assessments were collapsed in deriving a single measure of compliance.

Procedure

IO Checklists and PRs[7] could either be completed using an electronic form or handwritten on paper which was then scanned and uploaded as an image or pdf file. To reduce the need for manual data entry and chances for input error, a script was written in R to recognise and convert handwritten input to text using Object Character Recognition (OCR) implemented by pdftools and tesseract packages in R (Jeroen, 2022a; Jeroen, 2022b).

Triangulation between IO checklist, PRs, supervisee demographics, officer demographics and supervision outcomes were performed by matching CCO and supervisee names. As names were entered manually, omissions and variations in spellings can have a negative impact on data linkage. To circumvent some of these issues, a name matching algorithm using Jaro-Winkler distance (stringdist package; Van der Loo, 2014)

[6] The rating of 'N – no skill evidenced' is applicable to situations where there was no evidence that a skill has been attempted or applied, whereas 'M – missed opportunity" is applicable to situations where there was some evidence observed, but the CCO more often missed the opportunity to apply the skill than not.

[7] At the time of study, the Practice Review Database which allows systematic input of PRs into a structured database has not been implemented.

and fuzzy logic principles was written in R. A final check on the name matched output was manually conducted to ensure that the algorithm's performance was acceptable. In total about 4% of IO and PR assessments could not be matched due to use of name initials or omissions. These unmatched assessments were included in the descriptive analysis but were dropped when deriving aggregate CCO assessment scores for modelling purposes.

Analytical plan

Descriptive statistics were used to examine the frequency of QA activities and distribution of IO and PR rating scores. Each response option on the IO Checklist and PR assessment was given a numeric value. The rating of "M – missed opportunity" was given a numerical score of 1 and "P – skill present and clearly observed" a score of 4[8]. An aggregate rating score for each core skill was derived by taking the mean score of all items related to the core skill. The rating of "N – no opportunity to apply skills" was removed when calculating mean scores. Where missing data was observed, mean scores were generated from the remaining available items. A higher order Total IO score was derived by taking the mean of the aggregate score across all items.

The relationship between supervisee outcomes and CCOs' IO and PR assessment scores were examined using logistic regression models. As outcomes of people who were supervised by the same CCO may be more similar than those who were supervised by a different CCO and at a different office, the need for a mixed effects or multilevel model was first assessed. This was achieved by examining if the inclusion of these variables as random variables resulted in a better fitting model which accounted for significantly more variance than a model without these variables. Variance explained by these random variables were assessed through calculations of Intraclass Correlation Coefficient (ICC). Unless otherwise specified, CCO and office location were entered as random effects with CCO (level 2) nested under office (level 3).

A forward selection procedure was used to determine the model of best fit. This selection process involved examining if increasingly complex models were significantly better fit for the data than preceding simpler models. The modelling process commenced with a baseline model without predictors which was then compared to a mixed effects model with only CCO and office as random effects (unconditional means model).

Depending on whether the random effect variables produced a significantly better fitting model, the baseline or the unconditional means model was then updated to include only covariate variables. This model, which will be referred to as the covariates only model, included variables that may be associated with supervisee outcomes but were not of primary interest to the study. These covariates were supervisee level variables (level 1); age at index community supervision episode, age at first conviction, LSI-R risk of reoffending, gender, length of supervision and Indigenous status. A final model was then fitted by updating the covariates only model to include IO and PR

[8] For PRs, the additional response option of "N – no skill evidenced" was given the lowest possible score of 1 and the numerical values of the other response options were adjusted accordingly.

scores, which were aggregated at the CCO level (level 2 variables). The comparison between the final and covariates only models gave indications of whether PR and IO scores provided significant predictive value beyond the covariate variables.

Model comparison was performed through a series of likelihood ratio tests and by examining changes to the Akaike Information Criterion (AIC) which takes model complexity into account when evaluating model fit. Mixed effect models were fitted using the lme4 package in R (R Core Team, 2024; Douglas et al., 2015). Reoffending and supervision failure outcomes were modelled separately. These models were fitted to a subsample of 1,852 supervisees who were under the supervision of 353 CCOs with available IO and PR scores.

The same modelling process was repeated to examine the relationship between office and CCO compliance with QA activities and supervisee outcomes. While the previous models were fitted to a subsample of supervisees, the models examining the relationship between QA compliance and outcomes were fitted to the entire set of available data.

RESULTS

Completion rates of IO and PR assessments

Interview Observations

A total of 1,278 IO Checklists were completed within the study timeframe, which amounts to an average of 107 per month. The least number of IO Checklists were completed in Jan 2021 (35 completed) while the highest was in May 2021 (180 completed). Figure 1 shows that there was a general increase in the number of IO checklists completed over time. On average about 99 IO checklists were completed monthly over the first six months of the study time frame and this had increased to an average of 114 over the last six months.

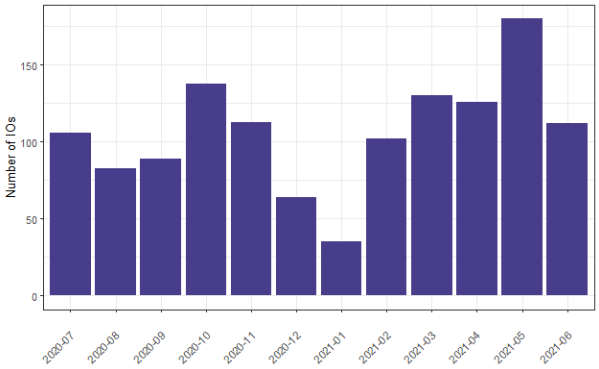

Figure 1. Number of IO checklists completed.

Practice Reviews

A total of 736 PRs were completed within the study time frame. On average, about 61 PRs were completed per month. The lowest number of completions was observed in Jan 2021 (19 PRs) and the highest in May 2021 (107 PRs). Figure 2 showed some indication that more PRs were completed in the second half of the study time frame (average of 59 in the first six months and 74 in the last six months).

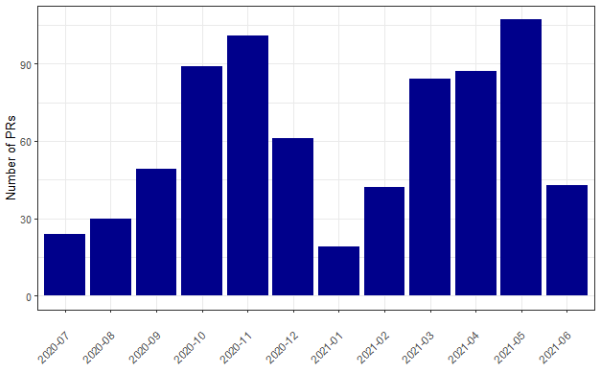

Figure 2. Number of PRs completed.

Distribution of IO and PR scores

Interview Observations

Figure 3 shows that the distributions of scores on IO Total and core skills were negatively skewed with 4, the highest possible score achievable, as the mode score. This was particularly salient for Prosocial Modelling where a score was 4 was given in about 65% of IO Checklists. This was followed by Rapport Building (48%), Cognitive Techniques (47%), and Intervention Focussed (30%). Given the high prevalence of ceiling scores, it was not surprising that a large proportion of IO Checklists (25%) received an average Total score of 4. This means that in one of four IO Checklists assessed, all items were given the most favourable score.

Practice Reviews

Figure 4 shows negatively skewed distributions for all PR core skills except for Intervention Evidence. The largest skew was observed for the domain of Record Keeping where TLs assigned the highest possible score in about 44% of assessments[9]. This was followed by Case Plan Update and Initial Case Plan Review where ceiling level scores were observed in about 32% and 11% of assessments respectively. In contrast, the distribution of scores in the Intervention Evidence domain was more flatly distributed, with an average score of 2.8 as the mode response (8% of assessments).

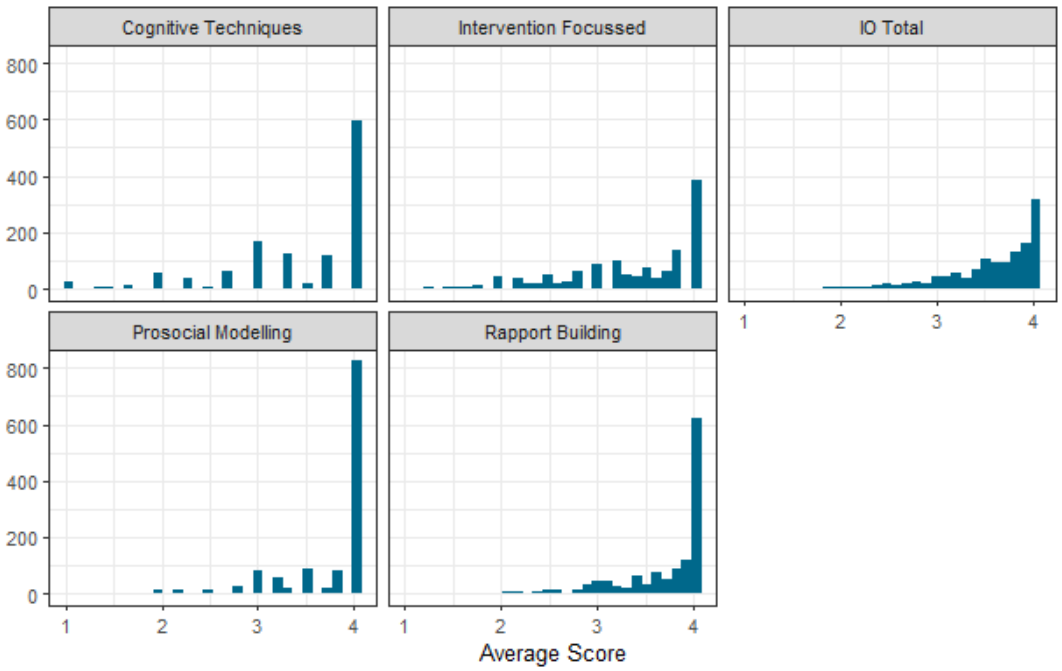

Figure 3. Distribution of average IO core skills and total scores.

[9] Each PR assessment includes reviews of CCOs' work with up to three supervisees under their supervision. To provide a more granular view of TL ratings, these descriptive statistics considered ratings for each supervisee separately. Ratings were collapsed at the CCO level for subsequent modelling purposes.

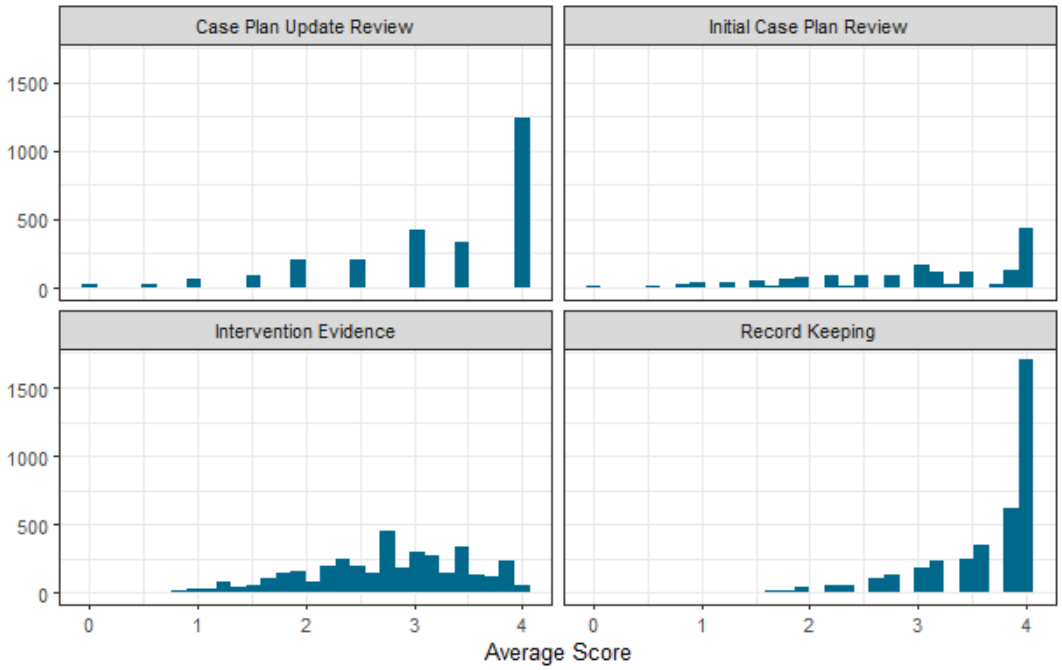

Figure 4. Distribution of average PR core skills.

Compliance with IO and PR assessments

For this study, a metric was derived to indicate CCOs' and offices' compliance with mandatory requirements related to engagement in QA activities. Compliance scores range from 0 to 1, with 0 indicating no compliance and 1 indicating full compliance.

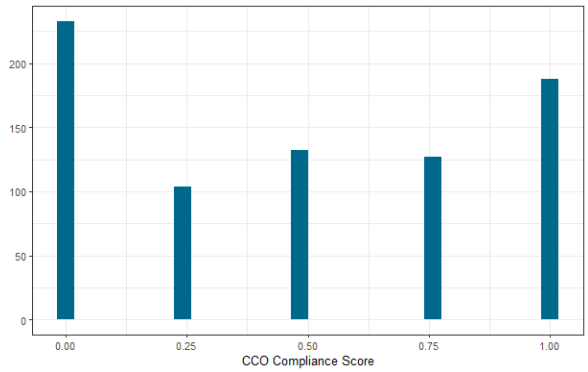

Figure 5. Histogram of CCO compliance scores.

CCO compliance

Figure 5 shows that the distribution of CCO compliance scores tended towards the two ends of the scale, which suggests that the majority of CCOs either did not comply (30%) or fully complied (24%) with engagement requirements. Of those with partial compliance, most (33% of entire sample) completed at least half of the required QA activities.

Office compliance

In contrast to Figure 5, Figure 6 shows that the distribution of office level compliance scores was congregated at the centre of the distribution. The compliance scores of the majority of offices ranged from 0.2 to 0.6 with the largest number of offices (30%) receiving a score of 0.5.

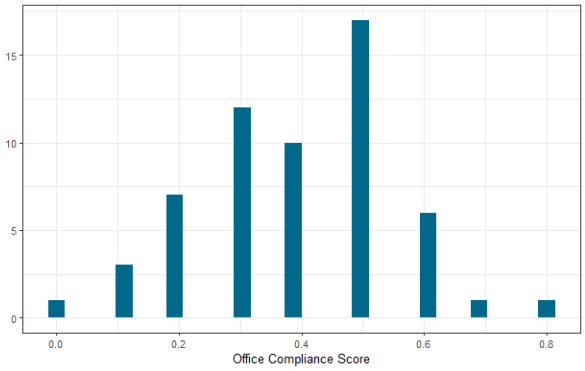

Figure 6. Histogram of office level compliance scores.

Relationship between IO and PR scores on outcomes

To examine the relationship between QA ratings and outcomes, only CCOs who were identified to have completed at least one IO and PR assessment within the study timeframe were included. This reduced the sample to 1,852 individuals who were supervised by 353 CCOs.

Reoffending

The need for a mixed effects model structure was first assessed by comparing the fit between the baseline model against the unconditional means model which included CCO and office as random effects, with CCOs (level 2) nested within office (level 3). The ICC of the unconditional means model was 0.04 which suggests that about 4% of variance in supervisees' likelihood of reoffending can be attributed to officer and office level differences. Despite the low ICC, Table 1 shows that the unconditional means model was a significantly better fit than the baseline mode, hence mixed effects models were fitted to all subsequent models in this set of analysis.

Table 1 also shows that the covariates only model which included supervisee predictors (level one) as covariates was a significantly better fit than the unconditional means model. Next, the full model which included IO and PR assessment scores as level two predictors showed a marginal improvement over the covariate only model (p = .05). The small change in AIC values between the covariates only and full model suggests that gains in model fit achieved by the full model was not justified by the increase in its complexity of including IO and PR assessment scores. However, given that there was a marginal tendency towards improvement, the full model was further explored while acknowledging that model diagnostics support the simpler covariates only model as the most parsimonious model.

The covariates only model explained about 21.1% of the total variance in supervisees' likelihood of reoffending. Table 2 shows the contribution of supervisee, officer and office level differences to the total variance explained. About 15.8% of the variance explained (i.e., of the 21.1% of variance) was attributed to the CCO and office level; 4.2% of variance was attributed to unobserved differences between Community Corrections offices (level 3) and about 11.6% to unobserved differences between CCOs (level 2). These unobserved differences may stem from other unmeasured variables such as CCOs' experience or other office level differences. The remaining 84.1% of variance explained was attributed to the fixed effects which comprised of supervisee level covariate variables.

Table 1. Fit statistics for reoffending outcomes.

Model	AIC	Chi-square	Degrees of Freedom	Significance
Baseline	2383.5	-	-	-
Unconditional means	2378.8	8.76	2	$p < .05$
Covariates only	2152.6	244.12	9	$p < .05$
Full	2152.8	13.88	7	$p = .05$

The contribution of PR and IO scores can be estimated from Table 2 by examining differences in the proportions of variance explained when PR and IO scores were incorporated in the full model. The full model explained about 21.0% of the total variance in supervisees' likelihood of reoffending. In comparison to the covariates only model, the inclusion of PR and IO scores in the full model showed reduced variance attributed to the CCO level (5.1%) and a corresponding increase in variance attributed to the fixed effects variables (89.3%). This suggests that by modelling IO and PR scores within the full model, the amount of unobserved CCO level impact on outcomes was reduced. A slight increase in variance explained attributed to differences at the office level was also noted.

Table 2. Proportion of variance explained.

Effect	Variance	
	Covariates only model	Full model
Fixed effects	0.74 (84.1%)	0.78 (89.3%)
CCO	0.10 (11.6%)	0.04 (5.1%)
Office	0.04 (4.2%)	0.05 (5.6%)

Table 3. Output of the full model for IO and PR scores on reoffending

	Variables	Levels	Odds Ratio	Significance
Supervisee	LSI-R	Low	1	-
		Medium-Low	1.75 [0.81. 3.76]	
		Medium	3.27 [1.52, 7.03]	$p < .05$
		Medium-High	8.9 [4.00, 19.84]	$p < .01$
		High	15.69 [6.14, 40.10]	$p < .01$
	Age		0.97 [0.96, 0.99]	$p < .05$
	Age at first conviction		0.99 [0.98, 1.00]	$p < .05$
	Gender	Female	1	
		Male	1.69 [1.18, 2.41]	$p < .05$
	Indigenous status	Non-Indigenous	1	
		Indigenous	0.85 [0.67, 1.09]	
	Episode Duration		1.00 [0.97, 1.04]	
IO Checklist	Rapport Building		0.69 [0.45, 1.04]	$p = .08$
	Intervention Focussed		1.45 [1.14, 1.85]	$p < .05$
	Cognitive Techniques		0.95 [0.75, 1.20]	
	Prosocial Modelling		0.92 [0.60, 1.42]	
PR	Record Keeping		1.04 [0.72, 1.50]	
	Intervention Evidence		1.19 [0.90, 1.58]	
	Case Plan Reviews[10]		0.94 [0.74, 1.19]	

[10] To reduce missing data, this variable was derived by collapsing scores on 'Initial Case Plan Review' and 'Case Plan Update Review'

Table 3 shows the results of the full model. Unsurprisingly, many level 1 covariate variables at the supervisee level were significant predictors of reoffending; odds of reoffending were higher for people with higher LSI-R classifications, and who were younger both at the start of the index community supervision episode and when first convicted. Being male was also associated with a 69% increase in odds of reoffending.

Table 3 also shows that there were some associations between IO scores and reoffending. Higher Intervention Focussed scores were significantly related to greater odds of reoffending such that a one-point increase in a CCO's average Intervention Focussed score was associated with a 45% increase in odds of reoffending. Higher Rapport Building scores were marginally associated with lower odds of reoffending; a one-point increase in average Rapport Building scores was associated with a 31% decrease in odds of reoffending. Scores from PRs were not associated with reoffending outcomes.

Supervision Failure

Assessments for the need of a mixed effects model structure indicated that the data was unable to support a nested random effects (three-level) structure. The baseline model was therefore compared against a simpler unconditional means model with only CCO entered as a random effect variable. Table 4 showed that the unconditional means model was not a better fit for the data and the ICC suggests that the random intercept accounted for less than 1% of the variance in supervision failure outcomes. Given that a more complex structure was not supported and that CCO level differences do not appear to explain meaningful variance in the outcome, no further models were pursued.

Table 4. Fit statistics for supervision failure outcomes.

Model	AIC	Chi-square	Degrees of Freedom	Significance
Baseline	1722.3	-	-	-
Unconditional means	1724.2	0.08	1	*n.s.*

Does compliance with QA have a relationship with outcomes?

The following analyses aimed to examine if compliance with QA activities has a relationship with supervisee outcomes and was conducted on the full sample of 3,818 supervisees. PR and IO scores were not included as variables in these models due to missing data for CCOs with compliance scores of zero.

Reoffending

Table 5 shows the model fit statistics examining the relationship between compliance with QA activities and reoffending outcomes. As the unconditional means model was a significantly better fit than the baseline model, mixed effects model structure was applied to subsequent models. The ICC of the unconditional means model was 0.03 which suggests that about 3% of the variance was accounted for by the nested data structure. Table 5 also shows that the covariates only model was a better fitting model than the full model. This suggests that indices of CCO and office level compliance with QA activities did not have a significant relationship with reoffending outcomes.

Table 5. Fit statistics for reoffending outcomes.

Model	AIC	Chi-square	Degrees of Freedom	Significance
Baseline	4996.9	-	-	-
Unconditional means	4984.3	16.54	2	$p < .001$
Covariates only	4493.2	509.12	9	$p < .001$
Full	4495.8	1.37	2	*n.s.*

Supervision Failure

Table 6 shows that the unconditional means model was a better fit than the baseline model, supporting the need for mixed effects model specifications. The ICC for the unconditional means model was 0.04 which suggests that about 4% of the variance was accounted for by the nested data structure. The full model was not a better fit than the covariates only model, suggesting that CCO and office level compliance with QA activities were not significantly related to supervision outcomes.

Table 6. Fit statistics for supervision failure outcomes.

Model	AIC	Chi-square	Degrees of Freedom	Significance
Baseline	3573.8	-	-	-
Unconditional means	3567.2	10.22	2	$p < .01$
Covariates only	3288.5	297.11	9	$p < .001$
Full	3292.3	0.16	2	*n.s.*

DISCUSSION

This study aimed to inform the nature of further training and development opportunities for CCOs by examining how well staff engaged with QA activities and how staff practice as assessed by these processes were associated with supervision outcomes. Over the study timeframe, we observed an upward trend in the number of IO and PR assessments completed. While we observed indications of relatively poor overall CCO and level office compliance with QA activities, the positive trajectory in assessments completed may reflect growing familiarity and improved integration of QA processes as part of standard practice.

This study found ceiling scores across most of the skills assessed by the IOs and PRs, particularly for Rapport Building and Prosocial Modelling in IOs, and Record Keeping in PRs. While this may suggest that TLs were highly positive and confident in their CCO's practice, the tendencies towards ceiling ratings on the IO Checklists and PRs raises implications for the validity of these scales in guiding professional development and informing targeted feedback. When ratings are consistently high or overly favourable, the capacity to identify specific areas for growth is diminished, limiting the potential for meaningful and continuous professional development. It should be noted that the statistical power of the subsequent analyses conducted in the current study was restricted by these ceiling scores and further investigations of the construct validity and the calibration of how these QA assessments are scored may be required. A similar observation was made in a previous study (Chong et.al 2024) which found ceiling scores across the skills assessed by IOs in addition to ratings of the supervisory relationship.

Model diagnostics indicated that incorporating IO and PR scores provided only marginal improvements in predicting reoffending outcomes beyond the effects of supervisee-level covariates. Further examination suggests that

only IO scores were associated with supervisees' likelihood of reoffending. Higher Rapport Building skills was marginally associated with lower odds of recidivism. This finding aligns with existing literature emphasizing the protective influence of high-quality supervisory relationships, where high scores on this measure may tap into positive perceptions of the relational aspects of supervision such as empathy, respect, caring and fairness which have been shown to foster stronger engagement and improved behavioural outcomes among supervisees (Trotter, 2016; Bonta et al., 2008; Skeem et al., 2007).

Contrary to expectations, higher Intervention Focussed (IO) scores were associated with higher odds of reoffending. This finding aligns with a previous study (Chong et al., 2024) which found that Intervention Focussed scores were positively correlated with supervisee's perceptions of trust in the dual-role relationship, which in turn was associated with higher odds of supervision failure. It is possible that CCOs who are more proficient in applying Intervention Focussed techniques may encourage supervisees to be more open and honest in disclosing potentially detrimental information that could necessitate a criminal justice response. While seemingly counterintuitive, these results suggest a complex dynamic where although proficiency in Intervention Focussed skills may be associated with stronger supervisory relationships, such relationships do not necessarily translate to more positive outcomes. Furthermore, demonstrations of skill use may be influenced by contextual factors where CCOs may be more likely to focus on intervention techniques when supporting behaviour change with individuals who have more complex needs or responsivity factors.

In general, the mixed effects models of this study found that CCO and office level differences accounted for about 15% (11.6% at the CCO level) of the variance explained by the model predicting reoffending outcomes. These findings were somewhat aligned with a previous study (Galouzis, 2020) which found that about 14.9% of the variance in a model examining rates of re-imprisonment for community supervised individuals were explained by CCO and office level factors (for non-Aboriginal supervisees only, CCO and office level factors were not significant for Aboriginal supervisees). Interestingly, when QA scores were included in the current model, the variance explained at the CCO level dropped from 11.7% to 5.1%, which suggests that although there are some concerns about the validity of these assessments, these scores as a reflection of staff practice may account for some of the impact that CCOs have on outcomes.

Although this study found associations between IO scores and reoffending outcomes, we did not find overall compliance with QA activity requirements, measured at both CCO and office levels, to be significantly associated with either reoffending or supervision failure. A possible explanation is that assessed IO scores were indicative of the existing skills of the supervising officer and the extent of engagement in QA activities at the individual or office level may have limited role in enhancing those skills. This possibility is supported by our observations of ceiling scores across many IO and PR assessments. However, our confidence in these findings is limited by several considerations. The timeframe of our study was limited to a one-year period during an early phase of implementation and mandatory requirements, and a longer

follow-up period may be required for these QA processes to result in measurable outcomes.

Some other limitations are noted. Despite best efforts to triangulate the data, some data points could not be matched due to missing information or linkage errors stemming from the use of name initials, abbreviations and omissions. Although estimates suggest that these linkage errors were not prevalent, they may still have an impact on our findings, particularly on our calculations of compliance with mandatory QA requirements. Unmatched data may inflate counts of non-compliance. This limitation may be mitigated to a certain extent in future studies with the introduction of the Practice Review Database in July 2021. The Practice Review Database is an online system which streamlines the process for conducting and keeping records of PRs and may therefore reduce errors associated with data management and technical challenges of utilising computer vision techniques to extract hand-written information.

Additionally, counts of non-compliance may also be overstated for staff members who left their roles during the period of study and therefore lacked the opportunity to complete the required QA activities. The analyses also did not account for potential changes in a supervisee's supervisory officer or office over the course of the study period. While this simplification allows for more straightforward analysis, it may have introduced error in the statistical estimates.

Conclusions

In summary, the current study indicated that staff showed increasing uptake of the QA processes over the study timeframe and that TLs were generally highly positive about the way that CCOs work with their supervisees, reflected in high ratings on IO and PR measures of quality. There were indications that IO scores were related to supervision outcomes, including some counterintuitive negative associations between reoffending outcomes and Intervention Focussed skills. We also noted that compliance with QA activities was not associated with supervision outcomes. However, our confidence in the current findings was limited by statistical considerations associated with the observed lack of variation in assessment scores. Relatedly, these findings also raise implications for the validity of these QA assessments in supporting staff professional development. Further evidence-based studies examining the utility of these QA processes and explorations of other assessments of staff practice may be beneficial for supporting ongoing staff development.

REFERENCES

Andrews, D. A., & Bonta, J. (2000). *The level of service inventory-revised.* Multi-Health Systems.

Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). LexisNexis/Matthew Bender.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990a). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*(1), 19–52. https://doi.org/10.1177/0093854890017001004

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990b). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, *28*(3), 369–404. https://doi.org/10.1111/j.1745-9125.1990.tb01330.x

Bonta, J., & Andrews, D. A. (2007). Risk-need-responsivity model for offender assessment and rehabilitation. *Rehabilitation*, *6*(1), 1-22

Bonta, J., Rugge, T., Scott, T.-L., Bourgon, G., & Yessine, A. K. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation, 47*(3), 248–270. https://doi.org/10.1080/10509670802134085

Cassidy, S. E., Nahleen, S., Chong, C. S., Thaler, O., & Howard, M. (2023). *Evaluation of the Practice Guide for Intervention (PGI): Staff Perceptions Five Years After Implementation* (Issue 67). Corrective Services NSW.

Chadwick, N., Dewolf, A., & Serin, R. (2015). Effectively training community supervision officers: A meta-analytic review of the impact on offender outcome. *Criminal Justice and Behavior*, *42*(10), 977–989. https://doi.org/10.1177/0093854815595661

Chong, C. S., Cassidy, S. E., & Howard, M. (2024). *Evaluation of the Practice Guide for Intervention (PGI): Triangulating perceptions of dual role relationships and their associations with staff practice and supervisee outcomes* (Issue 65). Corrective Services NSW.

Chong, C. S., Cassidy, S. E., Nahleen, S., Thaler, O., & Howard, M. (2023). *Evaluation of the Practice Guide for Intervention (PGI): Staff perceptions of Quality Assurance processes* (Issue 61). Corrective Services NSW.

Chong, C. S., Tran, N. L., van Doorn, G., & Howard, M. (2020). *Evaluation of the Practice Guide for Intervention (PGI): Relationships between offender needs and PGI use in case planning and supervision practice* (Issue 44). Corrective Services NSW.

Cullen, F. T. (2002). Rehabilitation and treatment programs. In J. Q. Wilson & J. Petersilia (Eds.), *Crime:* Public policies for crime control (pp. 253–290). ICS Press.

Douglas B., Martin M., Ben B., & Steve W. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01.

Gandreau, P. (1996). The principles of effective interventions with offenders. In A. T. Harland (Ed.), *Choosing Correctional Options That Work: Defining the Demand and Evaluating the Supply* (pp. 117–130). Sage.

Galouzis, J., Meyer, D., & Day, A. (2020). Associations between parole outcomes and characteristics of people under supervision, supervising officer, and supervising office. *Criminal Justice and Behavior, 47*(10), 1228–1243. https://doi.org/10.1177/0093854820946418

Gochyyev, P., & Skeem, J. L. (2019). Efficiently assessing firm, fair, and caring relationships: Short form of the Dual Role Relationship Inventory. *Psychological Assessment, 31*(3), 352.

Howard, M., & Chong, C. S. (2019). *Effects of the Practice Guide for Intervention (PGI) on behaviour change intervention dosage among community-based offenders* (Issue 40). Corrective Services NSW. https://doi.org/10.1080/00221473.1971.10624014

Jeroen, O. (2022a). pdftools: Text Extraction, Rendering and Converting of PDF Documents. R package version 3.1.1. https://CRAN.R-project.org/package=pdftools

Jeroen, O. (2022b). tesseract: Open Source OCR Engine. R package version 5.1.0. https://CRAN.R-project.org/package=tesseract

Lowenkamp, C. T., Latessa, E. J., & Smith, P. (2006). Does correctional program quality really matter? The impact of adhering to the principles of effective intervention. *Criminology & Public Policy*, *5*(3), 575–594. https://doi.org/10.1111/j.1745-9133.2006.00388.x

Lowenkamp, M. S., Robinson, C. R., Koutsenok, I., Lowenkamp, C. T., & Pearl, N. (2012). The importance of coaching: A brief survey of probation officers. *Federal Probation*, *76*(2), 36–40.

Ooi, E. J. (2020a). The impact of the Practice Guide for Intervention (PGI) on recidivism among offenders serving a community-based order. *Crime and Justice Bulletin No. 229*. Sydney: NSW Bureau of Crime Statistics and Research.

Ooi, E. J. (2020b). The impact of the Practice Guide for Intervention (PGI) on recidivism among parolees *Crime and Justice Bulletin No. 228*. Sydney: NSW Bureau of Crime Statistics and Research.

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Robinson, C. R., Lowenkamp, C. T., Holsinger, A. M., VanBenschoten, S., Alexander, M., & Oleson, J. C. (2012). A random study of staff training aimed at reducing re-arrest (STARR): Using core correctional practices in probation interactions. *Journal of Crime and Justice*, *35*(2), 167–188. https://doi.org/10.1080/0735648X.2012.674823

Skeem, J. L., Louden, J. E., Polaschek, D., & Camp, J. (2007). Assessing Relationship Quality in Mandated Community Treatment: Blending Care With Control. *Psychological Assessment, 19*(4), 397–410. https://doi.org/10.1037/1040-3590.19.4.397

Thaler, O., Chong, C. S., Raudino, A., & Howard, M. (2019). *Process evaluation of the Practice Guide for Intervention (PGI): Staff experiences of implementation and continuing service delivery* (Issue 60). Corrective Services NSW.

Tran, N. Le, Thaler, O., Chong, C. S., & Howard, M. (2019). *Process evaluation of the Practice Guide for Intervention (PGI): Staff perceptions of community supervision in the context of change* (Issue 59). Corrective Services NSW.

Trotter, C. (2016). *Working with involuntary clients: A guide to practice*. SAGE Publications. https://doi.org/10.1080/0312407x.2015.1112339

Van der Loo, M. P. (2014). The stringdist package for approximate string matching. *The R Journal, 6*(1), 111.

Other CRES Research Titles

Aug 2025	Identification of criminogenic needs using the PARRCC assessment tool
July 2025	Profiles of criminogenic need among Aboriginal and non-Aboriginal men and women in prison
June 2025	Conducting survey research using inmate digital tablets: Are respondents representative of the inmate population?
Dec 2024	Evaluation of the Practice Guide for Intervention (PGI): Triangulating perceptions of dual role relationships and their associations with staff practice and supervisee outcomes
Nov 2024	Exploring perceptions of correctional climate among people in prison in New South Wales
Nov 2024	Exploring perceptions of organisational fairness and safety attitudes among correctional staff in NSW
Sept 2024	Implementing digital technologies in Prison: Inmates' ongoing experiences of tablet access and connections with family and friends
Sept 2024	Drivers of program participation in NSW correctional centres: A trend analysis of digital tablets and Five Minute Interventions (FMI)
June 2024	Full body x-ray scanners at NSW correctional centres: Scanner activity, contraband, and behavioural outcomes
June 2024	Full body x-ray scanners at NSW correctional centres: Staff and inmate perspectives
April 2024	Evaluation of throughput and participation outcomes for the Intervention Pathways model
March 2024	Evaluation of the Compulsory Drug Treatment Program: Within-treatment change
Feb 2024	Examination of treatment dosage delivered through the Intervention Pathways Model
Dec 2023	Assessment and experience of prison climate in NSW correctional centres
Dec 2023	Five Minute Intervention (FMI) skills acquisition by correctional staff: The role of manager buy-in
Dec 2023	Dynamics of wellbeing and needs satisfaction among people in prison
Dec 2023	Empirical review of the Pathways to Employment (P2E) pilot for women in prison
Dec 2023	Conducting survey research using inmate digital tablets: Lessons learned from research conducted in NSW correctional centres
Oct 2023	Assessing the impact of Five Minute Interventions (FMI) training on behavioural indicators of correctional centre safety and order: An interrupted time series analysis
Oct 2023	Impacts of digital tablets on trends in correctional centre safety and order: A controlled interrupted time series study
Sept 2023	Development and validation of the Triage Risk Assessment Scale: Sexual Offending (TRAS:SO)
Sept 2023	Review of the birth certificate Identity Management Enhancement pilot
July 2023	Implementation of an initiative for people with less than five months to serve in prison: Staff insights on the Short Sentence Intensive Program (SSIP)
June 2023	The Intensive Learning Centre: Staff experiences of implementation and the SILC reforms
April 2023	Five Minute Interventions (FMI): Custodial staff views and experiences of implementing FMI
Dec 2022	Effects of the Workplace Mentor Program on correctional officers' perceptions of workplace culture
Dec 2022	Five Minute Interventions (FMI): Long-term effects of training on custodial staff attitudes towards prisoners, motivation and ability to support rehabilitation, and job stress and satisfaction

Research Bulletin No.68
ISSN 2207 0850

Corrections Research, Evaluation & Statistics
Delivery, Performance & Culture
Corrective Services NSW
Locked Bag 5000
Parramatta NSW Australia 2124

Email:
research.enquiries@correctiveservices.nsw.gov.au