

Evaluation of the Practice Guide for Intervention (PGI): Triangulating perceptions of dual role relationships and their associations with staff practice and supervisee outcomes

Chee Seng Chong, Sarah Cassidy, & Mark Howard

Aims

The study aimed to assess the dual role relationships between Community Correction Officers (CCOs) and their supervisees, characterised by a unique blend of care and control, under the current Practice Guide for Intervention (PGI) delivery phase and how these may be associated with CCOs' proficiency in applying interviewing skills in sessions. Another objective of this study was to examine if relationship quality is predictive of supervisee outcomes.

Methods

Between March 2022 and August 2023, CCOs across the state were asked to invite their supervisees to complete a Dual Role Inventory – Short Form (DRI-SF) as part of Interview Observations (IO). CCOs and third-party observers of those sessions also completed parallel versions of the DRI-SF. Responses to the DRI-SF were then triangulated with outcomes of the IO assessments as well as reoffending and other supervision outcome data.

Results

CCOs and supervisees were generally positive of the quality of their dual role relationship. This positivity was also reflected by the observers of these sessions. DRI-SF ratings generally indicated that CCOs demonstrated supervisory styles that were characterised by high Caring / Fairness and Trust but low on Toughness. Large proportions of supervisees and observers assigned maximum scores on the DRI-SF, indicating ceiling effects. Some significant but weak associations between DRI-SF scores and ratings of CCOs' proficiency in utilising interviewing skills were found. Only CCO DRI-SF ratings were significantly associated with supervision failure. Greater odds of supervision failure were observed when CCOs perceived higher Trust in the relationship. It is possible that higher Trust may be indicative of greater openness and honesty of supervisees in disclosing potentially detrimental information that could lead to revocation of their parole orders. We also found a marginal association where lower Toughness was associated with lower rates of supervision failure.

Conclusion

Ceiling effects in the data had severely limited the interpretation of our results and hold implications for the utility of the DRI-SF as well as current Quality Assurance (QA) processes in identifying and supporting best practice for CCOs. Further studies examining how CCOs' proficiencies may be more accurately assessed and developed through the current QA processes will benefit the continuation and expansion of staff development and training initiatives.

INTRODUCTION

Corrective Services NSW Community Corrections introduced the PGI in 2016 to enhance the delivery of targeted and individualised behaviour change interventions to people who are supervised in the community. The PGI consists of 56 exercises that CCOs selectively work through with people under their supervision to help address a range of criminogenic and responsivity factors. These exercises are designed to enhance the behaviour change content of supervision sessions and promote adherence to Risk Need Responsivity principles (e.g., Bonta & Andrews, 2016) by providing a platform for CCOs to help supervisees address their criminogenic needs in a structured cognitive-behavioural format that is responsive to the supervisees' learning styles and abilities.

Since its implementation, several PGI process evaluation studies focusing on the initial stages of rollout have been conducted (see Howard et al., 2019 for a summary of these studies). While these studies have generally found positive staff perceptions and uptake of the PGI, an early outcome evaluation study by NSW Bureau of Crime Statistics and Research (BOCSAR) found that the PGI had limited impact on reoffending outcomes (Ooi, 2020a and 2020b). A caveat of these studies is that they were conducted during the early stages of PGI implementation, and reoffending outcomes may have been impacted by initial challenges associated with program fidelity and the quality of service delivery. For example, a process evaluation study by Corrections Research, Evaluation and Statistics (CRES) found that although the PGI was frequently delivered in sessions during the early implementation phase, CCOs often chose to deliver generalist and process-oriented exercises that were unrelated to the supervisees' case plans (Chong et al., 2020). Furthermore, 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). Differences in orientations that officers hold towards their roles can have an influence on how they interact with their supervisees and apply behaviour change techniques in sessions (Ricks & Loudon, 2014).

Behaviour change interventions are most effective when staff are well trained and programs are delivered as intended (Chadwick et al. 2015; Robinson et al., 2011). While program integrity and fidelity can partially be monitored and maintained through the delivery of manualised programs such as the PGI where CCOs are guided in the use of appropriate content and methods, the work that individual CCOs do with those under their care and the components associated with successful or effective supervisor-supervisee relationship are often harder to observe (Bonta et al. 2008).

In recent years, there is a growing recognition that the appropriate use of interviewing skills and cognitive techniques which help supervisees address their thinking patterns can improve supervisee outcomes. For example, several studies have shown that prosocial modelling and reinforcement of positive behaviour were effective in reducing reoffending rates across community supervision in both adult and juvenile settings (Gendreau 1996; Trotter & Evans, 2012; Trotter, 2022). A meta-analysis also identified that the use of cognitive techniques in sessions was associated with a large effect size in reducing reoffending rates (Andrews & Dowden, 2004).

Besides cognitive behavioural techniques, the quality of the relationship between supervising officers and their supervisees is another key element that is associated with variations in supervisee outcomes (Bonta et al., 2011; Trotter & Evans, 2012; Weaver & Mcneill, 2012; Skeem et al., 2007, Kennealy et al., 2012; Dowden & Andrew, 2004). While the importance of building quality

working relationships is well established in other areas of research such as psychotherapy, the nature of the relationship between a supervising officer and their supervisee in criminal justice settings is different from a traditional therapeutic one. In therapeutic relationships a strong therapeutic alliance, which is characterised by a good collaborative relationship between the therapist and the client, a strong agreement on treatment goals and the tasks to achieve those goals, and the development of a positive emotional bond (Bordin, 1979), is consistently associated with positive treatment outcomes for the client (Flückiger et al. 2018, Horvath et al., 2011). However, in these traditional therapeutic relationships, clients and therapists are equal partners who enter a two-way relationship voluntarily (Cornelius-White, 2018), whereas those who are involved with the criminal justice system are often required to participate in mandated case management or treatment programs. In these mandated treatment relationships, the dynamic between the officer and supervisee is unbalanced with supervising officers wielding substantial control over their supervisees. Officers are required to not only provide care for their supervisees from a rehabilitative perspective, but also have control over the supervisee through a surveillance and compliance monitoring role. Hence, classic notions of the therapeutic alliance are argued to focus too narrowly on the therapeutic aspects of the relationship and rarely capture the control part of the relationship that is unique in mandated treatment programs (Skeem et al., 2007).

The recognition of the unique role that supervising officers play in providing both care and control have led to the conceptualisation of the dual role relationship, as assessed through psychometric tools such as the Dual Role Inventory - Revised (DRI-R) and its variants. The DRI-R is a well validated psychometric measure (Skeem et al., 2007) that assesses three factors associated with the quality of dual role relationships: Caring /

Fairness, Trust, and Toughness. The factor of Caring / Fairness assesses aspects associated with therapeutic alliance, Trust is associated with the extent to which officers and supervisees trust one another and the openness about disclosing information. The factor of Toughness taps into the more negative aspects of the relationship such as the officers' disciplinary orientation and their expectations for their supervisees to be compliant and self-sufficient. Evidence suggests that dual role working relationships that are characterised as "firm, fair and caring" promote better supervisee outcomes than relationships that are strictly therapeutic or authoritarian (Kennealy et al., 2012; Skeem & Manchak, 2008).

The current study

Given the importance of the role that CCOs play, knowledge about what happens during supervision and whether proficiency in applying practice skills augments the perceived quality of dual role relationships and supervision outcomes can help inform the nature of further training and skills development for CCOs. In recognition of this, Corrective Services NSW Community Corrections has introduced a number of QA¹ procedures that aim to increase not only the fidelity of PGI delivery but also the quality of working relationships between CCOs and their supervisees. One example is the Interview Observation (IO) process, where a Senior CCO or Team Leader observes and rates how well CCOs apply four main skills in supervision sessions against a checklist (IO Checklist), for the purposes of feedback and professional development. These skills are rapport building (e.g., use of affirmations), intervention focuses (e.g., identifying the aim of the PGI exercise used), cognitive techniques (e.g.,

¹ Since the time of study, these procedures are now referred to as Continuous Improvement in Practice (CIP) activities. The term QA was retained to reflect the operation context and method as they applied to the current study.

helping the offender develop alternative thoughts or behaviours) and pro-social modelling (e.g., reinforcing prosocial behaviour and attitudes).

The aim of the current study was to understand the quality of dual role relationships that CCOs have with their supervisees under the PGI model of supervision, and how this relates to staff practices as assessed by current QA processes. Another aim of this study was to understand whether perceptions of relationship quality was associated with supervisee outcomes. To achieve these aims, CCOs² across the state were asked to complete a Dual Role Inventory – Short Form (DRI-SF) as part of the IO process. The DRI-SF is a shorter version of the DRI-R, which reduces the DRI-R from a 30-item to a 9-item self-report scale (Gochyyev & Skeem, 2019). The supervisees and observers of the IO session were also asked to complete parallel forms of the DRI-SF to assess relationship quality from different perspectives. Ratings on these parallel forms were triangulated with ratings on the IO checklist and supervisee reoffending and supervision failure outcomes.

METHODS

Data

This study examined self-report and administrative data which include supervisee and staff demographic information. 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, Indigenous status, and Level of

Service Inventory – Revised (LSI-R) estimated risk of reoffending were extracted for use as covariates. The LSI-R is a well validated actuarial risk assessment instrument which is designed to identify an individual's criminogenic needs and assess their risk of general recidivism. Risk is categorised into five levels (Low, Low / Medium, Medium, Medium / High, and High). Staff demographic data including name, gender, position title and office location was obtained from Corrective Services NSW Human Resources Management for data triangulation purposes.

Interview Observation Checklist

Ratings of the proficiency in which CCOs applied their skills in supervision sessions were extracted from the IO Checklist. The IO Checklist was developed by Corrective Services NSW Community Corrections for use in IOs which was introduced as part of a set of QA processes intended to enhance the quality of interviews and to promote a culture of continuous professional development (see Chong et al. 2023 for more details on these QA processes). At each IO, an observer (Team Leaders or Senior CCOs) attends a supervision session to assess a CCO on their proficiency in utilising four core skills: Rapport Building; Intervention Focussed; Cognitive Techniques; and Prosocial Modelling, in their interaction with their supervisees. These core skills were assessed over 21 items of the IO checklist.

DRI-SF

Ratings of the quality of the dual role relationship between supervisees and their officers were assessed using the DRI-SF (Gochyyev & Skeem, 2019). Each of the factors on the DRI-SF are assessed using three items and each item is scored on a 7-point Likert-type scale to indicate the frequency at which the supervising officer displayed the target attribute (1 = never; 7 = always). Factorial scores were derived by summing up the relevant items. Factorial scores were further

² This includes every officer (CCO, Senior CCO or Trainee) who carries an active caseload of supervisees.

aggregated into a higher order Total score which served as an indicator of the general quality of the assessed dual role relationship. As the Toughness factor assesses aspects of the relationship that are inversely associated with quality, items on this scale were reverse coded prior to the calculation of the Total score. In the current study, three versions of the DRI-SFs (Supervisee, Observer and CCO) were created to assess the perceptions of each of these rater groups. These versions were created by reframing the questions to the perspective of the person completing the form³.

Supervision outcomes data

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 these supervision outcomes were extracted from OIMS.

Sample

After triangulating all three versions of the DRI-SF, we identified that our dataset involved ratings of 2,107 unique IO sessions. However, due to the manual process in which these data were collated and the voluntary nature of supervisees' participation, not all sessions had ratings from all rater groups. Of the 2,107 sessions, 997 sessions were rated by at least two raters and 95 by three.

³ As an example, the item "I treat my supervisee fairly" is rephrased as "The supervisor treats me fairly" in the supervisee form and as "The supervisor treats the supervisee fairly" in the observer form.

Supervisees

The 2,107 supervisory sessions examined in this study involved 1,130 people who were under supervision by Corrective Services NSW Community Corrections between the 1st of March 2022 and the 31st of August 2023. About 85% identified as male and 15% as female. At the start of the observation period, the average age of supervisees was 37.5 years (SD = 10.9) and by the end of the observation period around 39% of them had less than 10 supervision sessions with their current supervisor. About 78% of supervisees had previously been supervised by Community Corrections prior to the index episode. Given that completion of the DRI-SF was voluntary, DRI-SF ratings were received from 189 of the 1,130 supervisees, yielding a final total of 266 supervisee DRI-SFs⁴.

CCOs & Observers

The 2,107 rated sessions also involved a total of 742⁵ CCOs (72% identified as female, 28% as male). Not all CCOs completed a CCO DRI-SF; about 76% (565 CCOs) completed at least one CCO DRI-SF, resulting in a final total of 1,201 CCO DRI-SFs. A total of 399⁶ IO observers (70% identified as female, 30% as male) completed an average of four observer DRI-SFs and Interview Observations each, resulting in a final total of 1,810 observer DRI-SFs and 1,671 Interview Observation Checklists.

Procedure

Between March 2022 and March 2023, CCOs across the state were asked to invite their supervisees to complete a DRI-SF form as part of the IO process.

⁴ Some supervisees completed more than one DRI-SF.

⁵ We were unable to identify 12 CCOs using HR data due to insufficient information provided either by supervisees or observers.

⁶ Eight observers could not be matched for the same reasons given above.

To maintain anonymity and privacy, supervisees were given a QR code to access their copy of the form which they were asked to complete in their own time after the supervision session. The form was hosted on Alchemer for the duration of the observation period and was available online until August 2023. The CCO and the observer of the IO sessions were also asked to complete their respective DRI-SF forms.

IO Checklists, CCO and Observer DRI-SF could either be completed using an electronic form or on pen and 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, 2022). The digitized text was then compiled with the other electronic forms to generate a structured dataset for analysis.

Further post processing of the dataset was required to enable the triangulation between the IO checklists and the three DRI-SF versions. On all DRI-SF versions, respondents were required to record the name of the CCO. As names were entered manually, omissions and variations in spellings can have a negative impact on the linkage between ratings and individual demographic details. To circumvent some of these issues, a name matching algorithm using Jaro-Winkler distance (stringdist package; Van der Loo, 2014) and fuzzy logic principles was written in R. A final check on the name matched output was then manually conducted to ensure that the algorithm's performance was acceptable.

Analytical Plan

Initial data diagnostics indicated that a very small percentage of received DRI-SF contained incomplete data (3% for the CCO and Supervisee DRI-SF forms,

and 4% for the Observer form). Factorial and Total scores across all versions of the DRI-SF were not normally distributed and showed substantial negative skew. Descriptive results and other measures of association therefore used non-parametric statistics.

Where missing data were observed, mean scores were generated from the available items. To examine how well supervisees, supervisors and observers agreed when rating the same IO session, agreement between scores was determined by calculating the Intraclass Correlation Coefficient (ICC). Ideally, this analysis should be conducted on instances where the IO was rated by all three raters, however, given the small sample size (95 sessions), we used a less stringent criteria to include all sessions which were rated by at least two raters (997 sessions). This less stringent criteria meant that there were missing data which excluded the use Cohen's Kappa or its variants, hence the choice of ICC. ICCs were derived using the psych package (Revelle, 2022) in R. ICC estimates and their 95% confidence intervals were calculated based on single rater, absolute agreement, one-way random effects model, on the rationale that inter-rater reliability was the primary interest of this analysis, and scores were obtained from different sets of raters (see Koo & Li, 2016).

Further examination of the pair-wise agreement between rater groups was conducted using Fleiss' Kappa (irr package; Gamer et al., 2012). Kappa values are bounded between -1 and 1, with 1 indicating perfect agreement, -1 indicating perfect disagreement, and 0 indicating independence or chance level agreement. Given the homogeneity in the observed data, it is expected that Kappa values will tend towards 0 due to inflated calculations of percent chance agreement (Feinstein & Cicchetti, 1990); as such, other descriptive statistics such as percentage agreement and scatter plots were included to support the analysis.

Associations between DRI-SF ratings and IO Checklist outcomes were examined using Spearman's rho. This analysis was conducted on a subset of 1,671 Interview Observations⁷ which were rated on at least one DRI-SF. Each item on the IO Checklist was given one of five possible ratings; "N - no opportunity to apply skill"; "M - missed opportunity"; "D - skills developing"; "E - skills enhancing"; and "P - skill present and clearly observed". Mean scores were derived for each core skill by assigning a numeric score to each rating. The rating of "M - missed opportunity" was given a numerical score of 1 and "P - skill present and clearly observed" a score of 4. The rating of "N - no opportunity to apply skills" was removed when calculating mean scores. A Total IO score was derived by taking the mean of the aggregate score across all items.

A primary objective of this study was to examine if ratings of dual role relationships have predictive validity for supervisee outcomes. It is noted that there is hierarchy in the data where multiple supervisees can be supervised by the same CCO and groups of CCOs who operated from the same Community Corrections Office. An implication of this nested data structure is that the outcomes of people who were supervised by the same CCO or at one office may be more similar than those who were supervised by a different CCO or at a different office. Mixed effects models were therefore fitted to the data to determine if there were any statistically significant relationships between ratings of dual role relationships and outcomes, while isolating the impact that other higher order factors such as those associated with features of the supervisor and office location may have on outcomes.

As part of the modelling procedure, a forward selection procedure was used to determine the model of best fit. This involves starting with a null

model and examining how model fit changes as variables were added in iteratively. The model of best fit was considered found when model fit did not significantly improve with the inclusion of additional variables into the model equation. In the current analysis four models were fitted to each version of the DRI-SF using the stats and lme4 packages in R (R Core Team, 2024; Douglas et al., 2015).

The first model fitted was a null model to establish a baseline where it was assumed that none of the variables considered in this study had an effect on outcomes. A covariate only logistic regression model which included other variables that are known to be associated with supervisee outcomes but were not of primary interest to the study was then fitted. The covariate only model included supervisee level variables such as age at index community supervision episode, age at first conviction, LSI-R risk of reoffending, gender, Indigenous status and number of supervision sessions attended with their primary supervisor⁸. A third model, the DRI-SF model, which extended the covariate model to include DRI-SF factorial ratings was then fitted. A comparison of model fit between the DRI-SF and the covariate only model gave indications of whether DRI-SF ratings provided any predictive value beyond the covariates. Finally, a full model which modified the DRI-SF model to include supervisor and location as nested random effects was fitted. Model comparison was performed by examining if the increasingly complex models were significantly better fit for the data than the simpler ones. This was examined through a series of likelihood ratio tests and the Akaike Information Criterion (AIC) which takes model complexity into account when evaluating model fit.

⁷ About 80% of IOs with DRI-SF ratings.

⁸ This controls for exposure as likelihood of failure may accumulate as a function of supervision length. This also controls for cases where supervisees had multiple CCOs due to variations in their orders or office of supervision.

Of the 1,130 supervisees who were involved in IO sessions where DRI-SF ratings were obtained, data from 380 supervisees were excluded from this final analysis due to an inability to match the ratings to either supervisor or supervisee demographic data resulting from missing or inaccurate name spellings. This resulted in a final dataset of 750 supervisees. The two outcomes of interest were modelled separately, once for each rater group. As not all sessions were rated by each of the rater groups, the models for the Observer DRI-SF consisted of a sample of 687 supervisees while the models for the CCO DRI-SF consisted of a sample of 489 supervisees. No models were fitted for the DRI-SF supervisee version due to limited power associated with small sample size (n = 65).

RESULTS

Following is an overview of the findings of this study. In sections involving the DRI-SF, the Toughness factor was reverse coded in all analyses and tables presented so that a higher score reflects perceived lower Toughness. As such, all DRI-SF results presented in this section can be interpreted

so that higher scores are indicative of more favourable ratings of the supervisory relationship.

What is the quality of dual role relationships between CCOs and supervisees?

High internal consistency was observed for the DRI-SF Total scores across all forms ($\alpha = .9, .8$ and $.8$; for the Supervisee, CCO, and Observer forms respectively). Consistency was high for all factors, except Toughness; see Appendix A for a detailed list of factorial consistencies.

Table 1 shows descriptive statistics for the factorial and DRI-SF Total scores across the three DRI-SF versions. For the factors of Caring / Fairness and Toughness, the median scores on the Supervisee and Observer forms were equivalent to the maximum score as more than half of the respondents gave the highest possible score on all items in these factors. Table 1 also shows that on average, observers and supervisees tended to give higher ratings of DRI-SF Total scores (58.3 and 58.1 respectively) when compared to CCOs (54.4).

Table 1. DRI-SF median and average scores across all DRI-SF versions.

Measure		DRI-SF Version		
		CCO	Supervisee	Observer
Median Score (Range of possible scores)	Total (9-63)	55	62	59
	Caring / Fairness (3-21)	20	21*	21*
	Trust (3-21)	18	20.5	19
	Toughness (3-21)	17	21*	21*
Average Aggregate Factorial Score	Total	54.4	58.1	58.3
	Caring / Fairness	19.6	19.5	20.1
	Trust	17.5	18.9	18.3
	Toughness	17.3	19.7	19.9

* median scores were equal to the maximum score

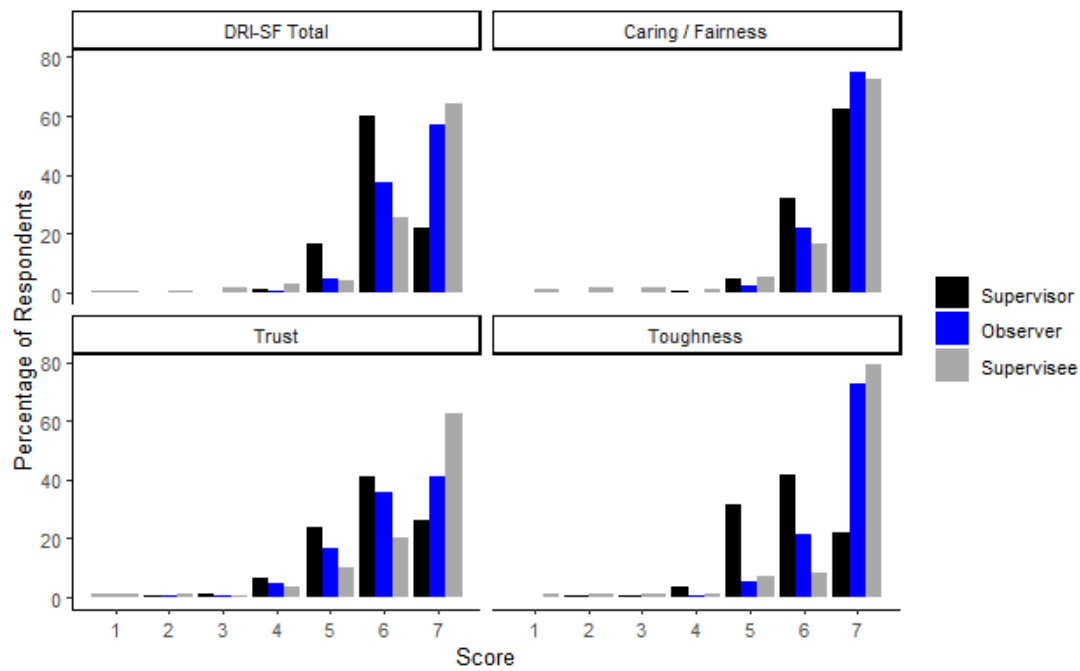


Figure 1. Distribution of item level mean factorial and DRI-SF Total scores across three versions of the DRI-SF.

Figure 1 shows the distribution of item level means for DRI-SF Total and factor scores across all DRI-SF versions. Ceiling effects were common among supervisees. Across all factors, at least 62% of supervisees gave their CCOs the most favourable rating; DRI-SF Total (64%), Caring / Fairness (73%), Trust (62%) and Toughness (79%). Observers gave

similarly high ratings: DRI-SF Total (57%), Caring / Fairness (75%), Trust (41%) and Toughness (73%). In comparison, across all factors except for Caring / Fairness, a smaller proportion of CCOs rated themselves on the most positive score, with a score of 6 as the mode response.

Table 2. Correlations between DRI-SF factorial and overall scores across all DRI-SF versions.

DRI-SF Form		DRI-SF Scores			
		Total	Caring / Fairness	Trust	Toughness
CCO	Caring / Fairness	.55	-		
	Trust	.69	.41	-	
	Toughness	.65	.29	.27	-
Supervisee	Caring / Fairness	.78	-		
	Trust	.68	.69	-	
	Toughness	.74	.46	.35	-
Observer	Caring / Fairness	.60	-		
	Trust	.76	.50	-	
	Toughness	.60	.34	.37	-

Consistent with the direction observed in the item level means, a series of Spearman rank-order correlations (see Table 2) found that within each version, all factorial ratings were significantly correlated with DRI-SF Total scores with large effect sizes ($r_s > 0.5$, $p < .001$ for all factors). Table 2 also shows that across all DRI-SF versions, the largest correlations between factors were generally observed for factors of Trust and Caring/Fairness; CCO ($r_s = .41$; $p < .001$); supervisee ($r_s = .69$; $p < .001$); observer ($r_s = .50$; $p < .001$). On the other hand, the reverse coded Toughness factor scores generally had the weakest positive correlations with other factors.

What are the similarities and differences in perceptions of the dual role relationship when rated by the supervisee, the CCO, and the observer?

While the previous analysis examined the associations between factors and Total DRI-SF scores within each rater group, the following analysis examined how well the different groups agreed when rating the same IO session. The average measure ICC for all raters on the Total DRI-SF scores was .20 with a 95% confidence interval from 0.16 to 0.24. By convention, an ICC of less than 0.5 is considered to indicate poor agreement. However, a smaller ICC was expected from the current analysis due to the use of one-way models, the requirement for absolute agreement and the observed lack of variability and negative skew in rating scores (Koo & Li, 2016; Mehta et al., 2018).

Further examination of the pair-wise agreement between rater groups was conducted using Fleiss' Kappa. However, given the skewness in the data, it is expected that Kappa values will tend towards 0 which suggests chance level agreements in rating despite the general high scores assigned by all rater groups. This is a known artefact of Kappa statistics when there is high levels of agreement due to skewness or low variance in the data (see Derksen

et al., 2024, and Feinstein & Cicchetti, 1990); as such, interpretation of the degree of agreement based on the generally accepted consensus of Kappa thresholds were therefore not applicable to the current analysis and further descriptive statistics and visualisations were provided to help understand patterns of agreement between these raters.

Agreement between supervisees and CCOs

A series of Fleiss' Kappa showed that there was chance level agreement between supervisees and CCOs; DRI-SF Total: $k = -.1$; Caring / Fairness: $k = .01$; Trust: $k = -.0$; Toughness: $k = -.2$. Figure 2 gives a series of scatter plots showing the agreement in scores between supervisee and CCOs. Linear regression lines of best fit were included to help visualise the degree of correlation between ratings. Black datapoints indicate where absolute agreement was observed (both raters selected the same rating score), red points indicate where supervisees rated the relationship higher than their CCOs, and blue points indicate instances where supervisees rated the relationship lower than their CCOs. Across all ratings, the datapoints were generally congregated in the upper right corner of the figures which indicated that most CCO and supervisee pairs highly rated their relationship.

Absolute agreement was observed in about 38% of all ratings (black datapoints). Where absolute agreement failed to be observed, 76% of these cases were due to instances where supervisees assigned a higher rating than their CCOs. The greatest degree of agreement was observed in the factor of Caring / Fairness (58% of ratings) and the least in Toughness (22%). Nevertheless, some consistency in ratings between supervisees and CCOs was observed with 65% of paired ratings falling within one score of one another (e.g., a supervisee gave a rating of 7 while the CCO gave a rating of 6). However, no significant correlations were found between supervisee and CCO ratings; $r_s < .1$ for all factors.

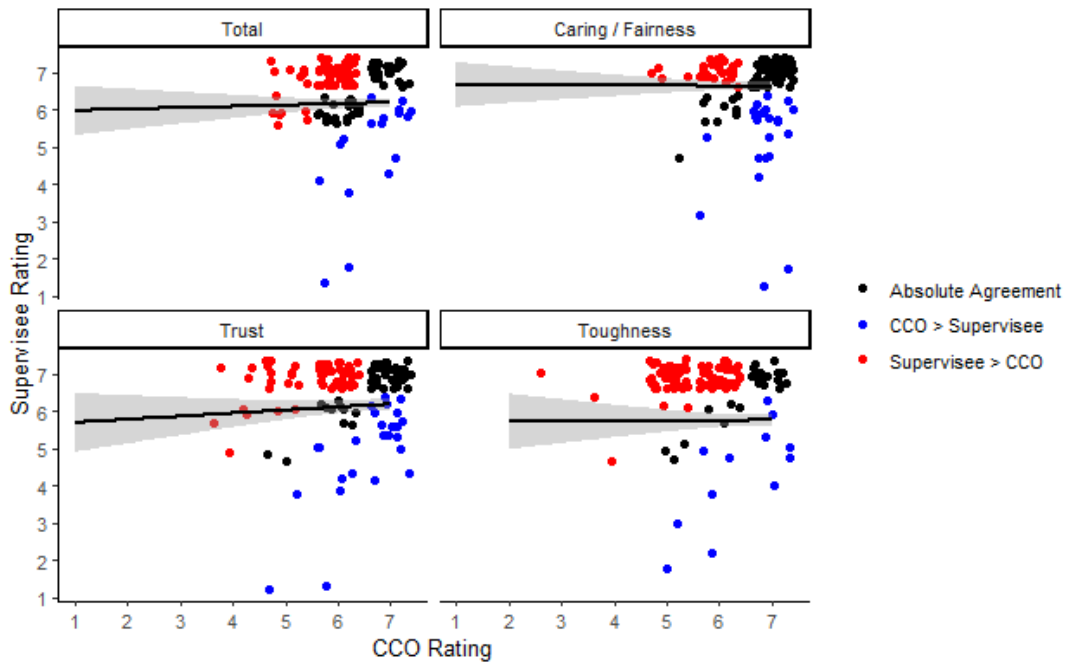


Figure 2. Agreement between supervisees and CCOs (n = 116 IOs).

Agreement between supervisees and observers

A series of Fleiss' Kappa statistics indicated chance level agreement between supervisees and observers; $k = .0$ for the factors of Toughness and Trust, and $k = .1$ for DRI-SF Total and Caring /

Fairness. Figure 3 shows clustering of responses which suggests that both groups generally assigned high rating scores and on average, absolute agreement was observed in about 58% of ratings with the highest level of agreement observed in the Toughness factor (66%) and lowest in Trust (45%).

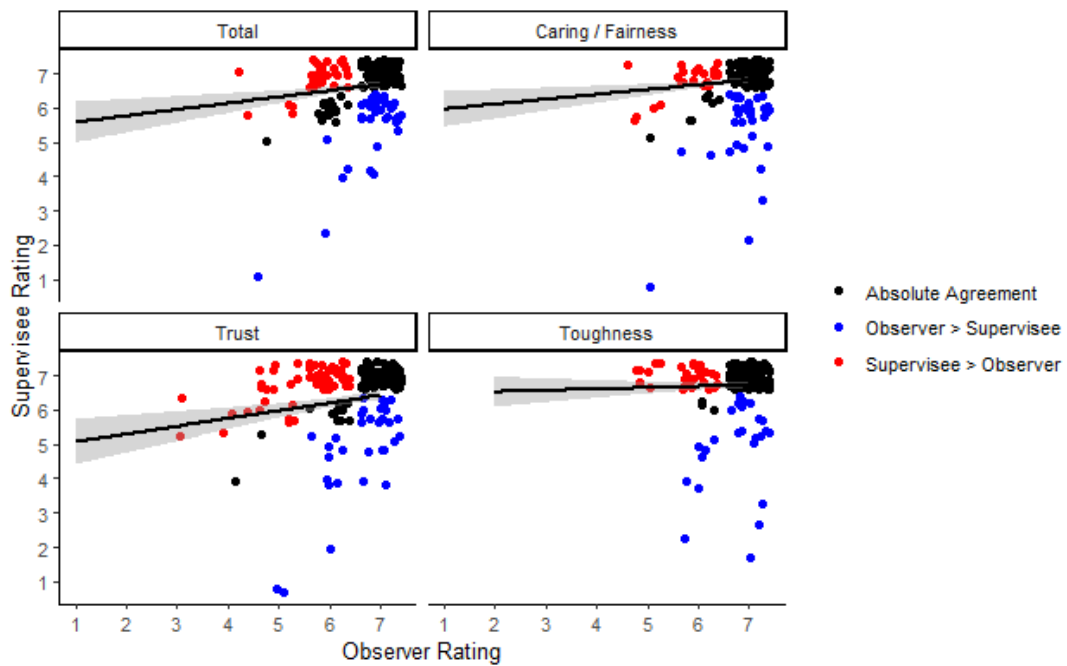


Figure 3. Agreement between supervisees and Observers (n = 158 IOs).

Where disagreement was observed, supervisees gave higher ratings than observers on slightly more cases (57%). In about 76% of disagreements, paired ratings fell within one rating score of one another suggesting some degree of consistency in ratings. Significant weak positive correlations between ratings were observed across most factors; DRI-SF Total: $r_s = .20$; $p < .05$; Caring / Fairness: $r_s = .18$; $p < .05$; Trust: $r_s = .24$; $p < .05$ and Toughness: $r_s = .10$; *ns*.

Agreement between CCOs and observers

Figure 4 shows the agreement between CCOs and observers. The density of datapoints in these figures were higher due to the larger number of sessions that these two groups rated. However, a series of Fleiss' Kappa indicated that there was chance level agreement between these two groups of raters; $k = -.0$ for Toughness, $k = 0.1$ for DRI-SF Total and $k = .2$ for Trust and Caring / Fairness

respectively. On average, absolute agreement was observed in about 45% of cases; with the highest level of agreement observed in the Caring / Fairness factor (63%) and lowest in Toughness (33%). In 80% of disagreements, observers assigned a higher rating score than CCOs (blue datapoints).

Although there was poor absolute agreement between the raters, there was consistency in the ratings where paired ratings between CCOs and observers tended to be similar. In about 76% of disagreements, paired ratings fell within one rating score of one another. Rating scores were also generally clustered on the higher end of the rating scale with weak to moderate positive correlations found between rating scores (DRI-SF Total: $r_s = .30$; $p < .001$; Caring / Fairness: $r_s = .24$; $p < .001$; Trust: $r_s = .38$; $p < .001$ and Toughness: $r_s = .20$; $p < .001$).

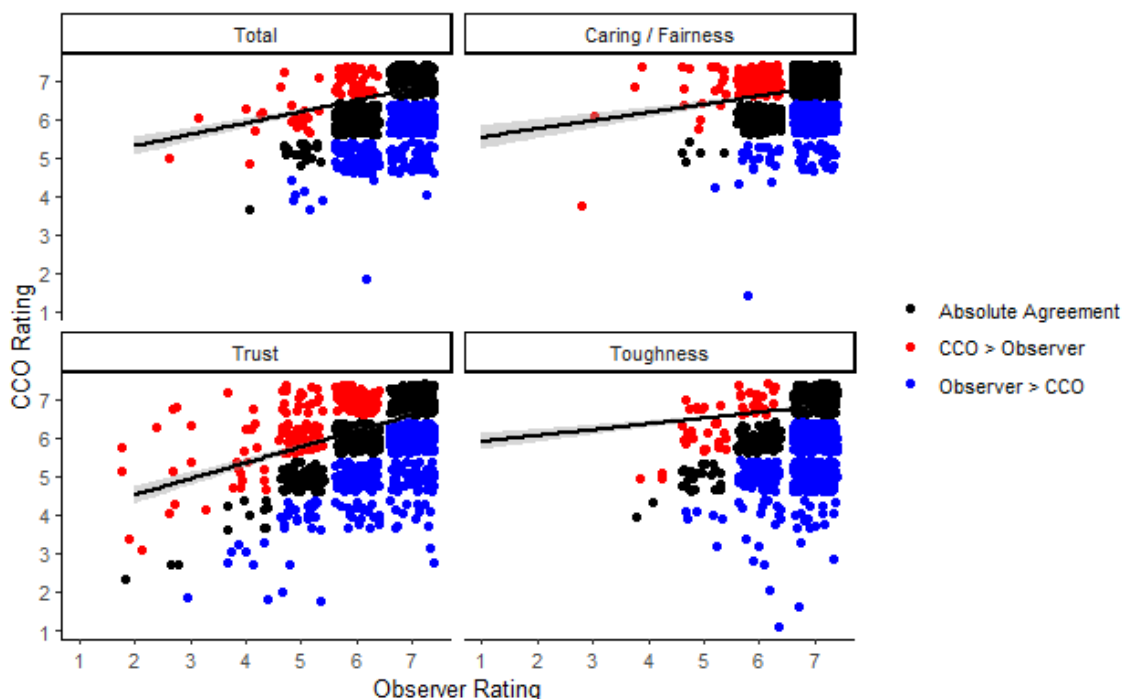


Figure 4. Agreement between CCOs and observers (n = 913 IOs).

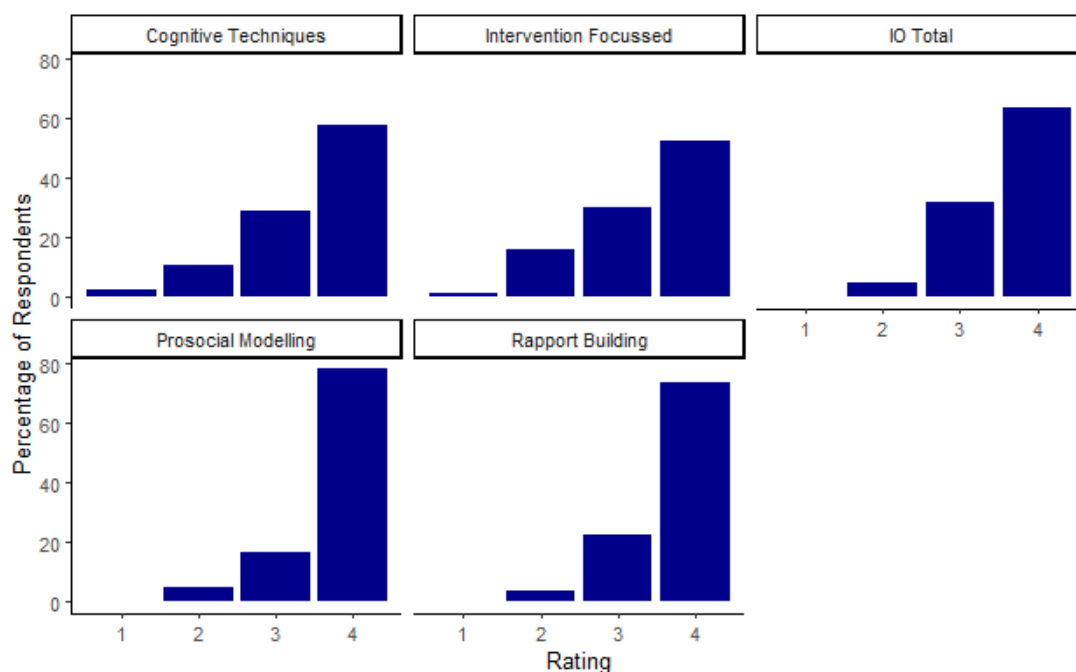


Figure 5. Distribution of mean IO Total and core skill scores (n = 1,671).

Are ratings of dual role relationship associated with interviewing skills?

Figure 5 shows the mean ratings across the IO Total and four core skills. Similar to the pattern observed in DRI-SF ratings, IO checklist scores were negatively skewed. CCOs received an average rating of 4⁹ on the IO Total score in about 63% of IO checklists. Assessments of Prosocial Modelling skills received the highest ratings, with CCOs scoring an average of 4 in about 78% of IO checklists. This was followed closely by Rapport Building (74%), Cognitive Techniques (58%) and Intervention Focussed (53%) skills.

Table 3 shows the correlation between ratings on the different DRI-SF versions and IO rating. Across the three DRI-SF versions, the highest correlations were observed between Observer DRI-SF and IO

ratings. This was expected as both DRI-SF and IO checklists were completed by the same Individual. There were moderate positive relationships between ratings of CCOs' Rapport Building skills and DRI-SF Total ($r_s = 0.34$; $p < 0.01$), Caring / Fairness ($r_s = 0.45$; $p < 0.01$), and Trust ($r_s = 0.30$; $p < 0.01$) scores on the Observer DRI-SF.

Weak positive correlations were observed between the various IO checklist scores and supervisee DRI-SF ratings. The largest correlation found was between the DRI-SF factor of Trust and Intervention Focussed skill ($r_s = 0.16$; $p < 0.05$). This was followed by a small positive correlation between Toughness and Prosocial Modelling ($r_s = 0.15$) which approached significance. Weak positive correlations were also observed between Trust and other core CCO skills such as Prosocial Modelling ($r_s = 0.14$) and Rapport Building ($r_s = 0.12$); and IO Total scores ($r_s = 0.14$). These associations were not statistically significant. No associations were found between the DRI-SF factor of Toughness and IO ratings.

⁹ Rating scores were rounded up. Observers gave CCOs a score of 4, the maximum possible score, on all 21 items in about 20% of the IO checklists examined.

Table 3. Correlations between DRI-SF and IO ratings.

DRI-SF Form		IO Checklist				
		IO Total	Rapport Building	Intervention Focussed	Cognitive Techniques	Prosocial Modelling
DRI-SF Observer	DRI-SF Total	<i>.28</i>	<i>.34</i>	<i>.19</i>	<i>.24</i>	<i>.27</i>
	Caring / Fairness	<i>.37</i>	<i>.45</i>	<i>.26</i>	<i>.30</i>	<i>.35</i>
	Trust	<i>.27</i>	<i>.30</i>	<i>.20</i>	<i>.27</i>	<i>.26</i>
	Toughness	<i>.15</i>	<i>.19</i>	<i>.10</i>	<i>.12</i>	<i>.16</i>
DRI-SF Supervisee	DRI-SF Total	<i>.10</i>	<i>.06</i>	<i>.09</i>	<i>.08</i>	<i>.11</i>
	Caring / Fairness	<i>.06</i>	<i>.07</i>	<i>-.00</i>	<i>.03</i>	<i>.04</i>
	Trust	<i>.14</i>	<i>.12</i>	<i>.16</i>	<i>.03</i>	<i>.14</i>
	Toughness	<i>.02</i>	<i>.02</i>	<i>.03</i>	<i>-.00</i>	<i>.15</i>
DRI-SF CCO	DRI-SF Total	<i>.08</i>	<i>.11</i>	<i>.08</i>	<i>.08</i>	<i>.10</i>
	Caring / Fairness	<i>.10</i>	<i>.13</i>	<i>.09</i>	<i>.11</i>	<i>.12</i>
	Trust	<i>.06</i>	<i>.08</i>	<i>.05</i>	<i>.11</i>	<i>.08</i>
	Toughness	<i>-.00</i>	<i>.01</i>	<i>.01</i>	<i>-.03</i>	<i>.02</i>

Note. The highest correlation in each row is in italics and bold.

Weak positive correlations were also found between IO and CCO DRI-SF ratings. The largest correlation was observed between the DRI-SF factor of Caring / Fairness and Rapport Building ($r_s = 0.16$; $p < 0.01$), followed by Caring / Fairness and Prosocial Modelling ($r_s = 0.12$; $p < 0.01$). A weak positive association was also observed between Cognitive Techniques and Caring / Fairness ($r_s = 0.11$; $p < 0.05$). The factor of Toughness did not appear to have any measurable relationship with IO assessments.

Do ratings of dual role relationship have predictive validity for supervisee outcomes?

To determine whether DRI-SF ratings have predictive validity for supervisee outcomes, a forward selection procedure was used to determine the model of best fit. Four models of increasing complexity were fitted to each DRI-SF version.

These are the null logistic regression model; the covariate only logistic regression model; the DRI-SF logistic regression model which extended the covariate model to include DRI-SF factorial ratings; and the full model which modified the DRI-SF model to include supervisor and location as nested random effects. The null model presents a baseline model where we aim to determine if the inclusion of additional covariates and predictor variables (ratings of DRI-SF factors in this case) may lead to better model fit and less error. No models were fitted for the Supervisee DRI-SF due to the small sample size.

Observer DRI-SF

Reoffending

Table 4 shows the fit statistics for the four models examining the association between Observer DRI-SF ratings and reoffending outcomes. The covariate

only model appeared to be the model of best fit which suggests that Observer DRI-SF factorial ratings, supervisor and office location did not contribute significant predictive value beyond the covariate variables.

Table 4. Fit statistics of the models built during the forward selection procedure for determining the best fitting Observer DRI-SF reoffending model.

Model	AIC	Chi-square	Degrees of Freedom	Significance
Null	891.03	-	-	-
Covariates only	849.46	59.57	9	$p < .001$
DRI-SF	854.12	1.34	3	<i>n.s.</i>
Full	858.12	0.01	2	<i>n.s.</i>

Supervision Failure

Table 5 shows the fit statistics for the four models examining the association between Observer DRI-SF ratings and supervision failure outcomes. Similar to what was observed in Table 4, the covariate only model appeared to be the model of best fit. This suggests that Observer DRI-SF ratings did not have predictive value for supervision outcomes.

Table 5. Fit statistics of the models built during the forward selection procedure for determining the best fitting Observer DRI-SF supervision outcome model.

Model	AIC	Chi-square	Degrees of Freedom	Significance
Null	596.33	-	-	-
Covariates only	522.38	91.95	9	$p < .001$
DRI-SF	527.05	1.33	3	<i>n.s.</i>
Full	530.62	0.43	2	<i>n.s.</i>

CCO DRI-SF

Reoffending

Table 6 shows the fit statistics for the four models examining the association between CCO DRI-SF ratings and reoffending outcomes. The covariate only model appeared to be the model of best fit which suggests that Observer DRI-SF factorial ratings, supervisor and location did not have predictive value beyond the covariate variables.

Table 6. Fit statistics of the models built during the forward selection procedure for determining the best fitting CCO DRI-SF reoffending model.

Model	AIC	Chi-square	Degrees of Freedom	Significance
Null	634.89	-	-	-
Covariates only	619.30	33.58	9	$p < .001$
DRI-SF	621.39	3.91	3	<i>n.s.</i>
Full	625.39	0.00	2	<i>n.s.</i>

Supervision Failure

Table 7. Fit statistics of the models built during the forward selection procedure for determining the best fitting CCO DRI-SF supervision outcomes model.

Model	AIC	Chi-square	Degrees of Freedom	Significance
Null	423.08	-	-	-
Covariates only	373.95	67.12	9	$p < .001$
DRI-SF	369.52	10.43	3	$p < .05$
Full	373.52	0.00	2	<i>n.s.</i>

Table 7 shows the fit statistics for the four models examining the association between supervision outcomes and CCO DRI-SF ratings. From the model fit statistics, the DRI-SF model was identified to be

Table 8. Logistic regressions coefficients and odds ratio for the CCO DRI-SF model.

Variable		b	se	z ratio	p	Odds	95% CI	
							Lower	Upper
Covariates	LSI-R Low	-16.57	933.79	-0.02	n.s.			
	LSI-R Med/Low	-2.61	0.65	-4.01	p <.001	0.07	0.01	0.25
	LSI-R Med	-1.2	0.43	-2.79	p <.01	0.3	0.13	0.7
	LSI-R Med/High	-0.4	0.42	-0.96	n.s.			
	Age at start of index supervision	-0.4	0.02	-2.26	p <.05	0.96	0.93	0.99
	Gender - Male	0.31	0.46	0.68	n.s.			
	Aboriginal	0.33	0.29	1.15	n.s.			
	Age at first conviction	0	0.02	-0.15	n.s.			
	Sessions attended	-0.09	0.02	-3.81	p <.001	0.92	0.87	0.96
DRI-SF	Caring / Fairness	-0.16	0.25	-0.64	n.s.			
	Trust	0.51	0.18	2.8	p <.01	1.66	1.18	2.4
	Toughness	-0.3	0.16	-1.92	p = .05	0.74	0.54	1.01

Note. b represents the estimated change in supervision failure outcomes for a one-unit change in the predictor variable; se represents the degree of uncertainty in the b estimate; odds ratio represents the odds of supervision failure as the predictor variable increases and CI represents the confidence intervals of those odds bounded within the lower and upper limits (LL and UL). Odds ratio were presented for significant variables.

the model of best fit, suggesting that CCO DRI-SF ratings were predictive of whether supervisees breached their parole orders. As can be seen in Table 8, only the factor of Trust was significantly associated with supervision outcomes. The direction of the relationship was contrary to expectations, such that each unit increase in average Trust rating scores was associated with a 66% increase in odds of supervision failure. The relationship between Toughness and supervision outcomes also approached significance ($p = 0.05$). Regression outputs indicated that as ratings of CCO Toughness became more favourable (indicating lower Toughness), the odds of supervision failure decreased, so that each unit increase in reverse-coded scores was associated with a 26% reduction in odds of supervision failure.

DISCUSSION

The current study aimed to assess and examine the quality of the dual role relationship between CCOs and their supervisees under the PGI model of supervision, including how perceptions of relationship quality were associated with CCOs' use of effective interviewing skills and supervision outcomes. Perceptions of relationship quality by CCOs, supervisees, and observers were assessed using the DRI-SF while CCOs' proficiency at using interviewing skills were obtained from observer ratings on the IO Checklists.

It was observed that all three rater groups highly rated the quality of the dual role relationship between CCOs and supervisees. This was particularly salient among supervisee and observers; more than half of them gave the maximum possible score when rating the factors of Caring / Fairness and Toughness. In comparison, CCOs tended to self-rate themselves with slightly lower scores. While the ICC and Kappa statistics generally indicated poor absolute agreement between raters, it is noted that this may be due to the lack of variance in the skewed distribution of scores (see Feinstein & Cicchetti, 1990). Consistency in ratings were observed in the descriptive analysis of these results. Rating scores generally clustered together on the higher end of the rating scale and paired ratings tended to be similar with the majority of ratings falling within a 1-point margin. It was also observed that supervisees generally had a greater degree of agreement with observers than CCOs.

Some significant associations between DRI-SF ratings and IO scores were found. The largest associations were found between IO scores and the Observer DRI-SF, which may be expected given that the same individual completed both ratings. Correlations between the IO scores and factorial rating scores on the other versions of the DRI-SF were very low. However, there were some indications that patterns of association between IO scores and factorial DRI-SF ratings varied depending on the rater group, which could suggest that proficiency in applying a particular core skill may be associated with perceptions of different supervisory styles depending on the rater. For example, CCOs who were observed to have greater proficiencies in applying Prosocial Modelling skills may perceive themselves as having a supervisory style that is characterised by high Caring / Fairness, whereas supervisees may have perceived greater Trust in that relationship.

Examinations of the DRI-SF and IO ratings both indicated consistently high positive ratings which approached ceiling. While this may suggest that all rater groups held uniformly positive perceptions of the dual role relationships that are fostered under the PGI model of supervision, the tendencies towards ceiling ratings on the DRI-SF and IO checklists raises implications for the validity of these scales. For example, the observed skewness may potentially be due to demand effects or bias in self-rating style data collection procedures (see Howard et al., 2019, which identified similar issues). It should be noted that the statistical power of the analysis on these measures is restricted by these ceiling scores and further investigations of the construct validity and the calibration of how these measures are scored may be required.

The current study did not find evidence that DRI-SF ratings held predictive value for supervisee outcomes, with the notable exception of ratings made by CCOs. One possible explanation is that CCO ratings may tap into perceptions of the relationship over a longer term in contrast to the Observer DRI-SF, which rated shorter term in-session behaviours (unfortunately it was not possible to examine the predictive validity of Supervisee DRI-SFs due to the limited sample size). Our results showed that CCO ratings were significantly associated with outcomes related to supervision failure although not reoffending. The direction of some findings were contrary to expectations, whereby higher rates of supervision failure was associated with higher CCO self-ratings on Trust. One interpretation of this finding is that higher Trust may be indicative of greater openness and honesty of supervisees in disclosing potentially detrimental information that could lead to revocation of their parole orders.

The results also indicated that Toughness was marginally associated with supervision failure where a more favourable score (characterised by lower

Toughness) was linked to lower rates of supervision failure. This finding is consistent with conceptualisations of Toughness as more negative aspects of the supervisory relationship, such as punitiveness, indifference to supervisees' views and expectations of compliance (Skeem et al., 2007), as well as other literature indicating poorer outcomes for more authoritarian relationships (Kennealy et al., 2012; Skeem & Manchak, 2008). Items of the DRI-SF similarly appear to focus on negative relational characteristics, such as talking down to the supervisee. From this perspective, however, it is unclear to what extent the Toughness and other factors of the DRI-SF overlap with or capture the 'firm' aspect of the dual role relationship. We note that ratings on the Toughness factor of the DRI-SF often showed poor internal consistency, which indicates that items may reflect more than one construct or respondents may have had different interpretations of certain items. There may be scope for future studies to further develop measures of dual role relationships to more closely capture firmness and consider how it corresponds with Toughness as formulated in the DRI-SF.

Several limitations of this study are noted. Ratings on the DRI-SF and IO were highly positive and approached ceiling on a number of factors. The relationships between these ratings and outcomes were difficult to estimate due to the skewness and low variance in the data. We note that aims of this study included an initial exploration of the validity of data derived from these measures, and less favourable findings in this regard would necessarily impact on the viability of other analyses. As previously mentioned, while this study indicated that factors of Caring / Fairness and Trust were rated highly and valued in the supervisory relationship, available measures provided limited insights into how aspects of Toughness are incorporated into practice and correspond with firmness in the dual role relationship. In addition, voluntary response rates from supervisees were

relatively low, and it was not possible to conduct some analyses utilising this sample.

In summary, the current study indicated that multiple stakeholders including CCOs, supervisees, and observers tended to have highly positive views of aspects of supervisory relationships. There were indications that DRI-SF factors may be related to supervision outcomes, including a counterintuitive negative association between Trust and supervision failure. However, our confidence in the findings is limited by concerns about the validity of the measure in adequately capturing the dual role relationship. To some extent this may also be extended to the IO checklists that are employed by Community Corrections to assess interview skills. We observed patterns of association suggesting that the proficiency with which CCOs applied practice skills corresponded with perceptions of relationship quality. While QA processes such as the IO could therefore be of value in promoting professional development to enhance dual role relationships, our results gave initial indications that the utility of the IO checklist may be limited in this regard by low discrimination power. Further evidence-based studies examining the impact of QA processes on staff practice and the implementation of more valid measures would be beneficial for ongoing initiatives to support CCOs in developing optimal supervisory relationships.

REFERENCES

- Bonta, J., & Andrews, D. A. (2016). *The psychology of criminal conduct*. Routledge.
- Bonta, J., Bourgon, G., Ruge, T., Scott, T., Yessine, A., Gutierrez, L., & Li, J. (2011). An experimental demonstration of training probation officers in evidence-based community supervision. *Criminal Justice and Behavior, 38*, 1127-1148.
- Bonta, J., Ruge, 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.

- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, 16(3), 252.
- 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.
- Chong, C. S., Cassidy, S., Nahleen S., Thaler, O. & Howard, M. (2023). *Evaluation of the Practice Guide for Intervention (PGI): Staff perceptions of the Quality Assurance processes*. Sydney, NSW: 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*. Sydney, NSW: Corrective Services NSW.
- Cornelius-White, J. H., Kanamori, Y., Murphy, D., & Tickle, E. (2018). Mutuality in psychotherapy: A meta-analysis and meta-synthesis. *Journal of Psychotherapy Integration*, 28(4), 489.
- Derksen, B. M., Bruinsma, W., Goslings, J. C., & Schep, N. W. (2024). The Kappa Paradox Explained. *The Journal of Hand Surgery*.
- 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.
- Dowden, C., & Andrews, D. A. (2004). The importance of staff practice in delivering effective correctional treatment: A meta-analytic review of core correctional practice. *International Journal of Offender Therapy and Comparative Criminology*, 48(2), 203–214.
- Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology*, 43(6), 543–549.
- Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, 55(4), 316.
- Gamer, M., Lemon, J., Gamer, M. M., Robinson, A., & Kendall, W. (2012). *Package 'irr'. Various coefficients of interrater reliability and agreement*. <https://cran.r-project.org/web/packages/irr/index.html>
- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works!. *Criminology*, 34(4), 575–608.
- 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., Thaler, O., & Tran, L. (2019). Innovations in a model for enhancing the behaviour change content of supervision with community-based offenders. *Advancing Corrections*, 7, 71–83.
- Howard, M., Tran, L., Thaler, O. & Chong, C.S., (2019). *Blending care and control in delivery of the Practice Guide for Intervention (PGI): An assessment of the quality of dual role relationships between offenders and supervising officers in the community*. Sydney, NSW: Corrective Services NSW.
- Jeroen Ooms (2022). *pdfutils: Text Extraction, Rendering and Converting of PDF Documents*. R package version 3.1.1. <https://CRAN.R-project.org/package=pdfutils>
- Jeroen Ooms (2022). *tesseract: Open Source OCR Engine*. R package version 5.1.0. <https://CRAN.R-project.org/package=tesseract>
- Kennealy, P. J., Skeem, J. L., Manchak, S. M., & Eno Loudon, J. (2012). Firm, fair, and caring officer-offender relationships protect against supervision failure. *Law and Human Behavior*, 36(6), 496.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30.
- Mehta, S., Bastero-Caballero, R. F., Sun, Y., Zhu, R., Murphy, D. K., Hardas, B., & Koch, G. (2018). Performance of intraclass correlation coefficient (ICC) as a reliability index under various distributions in scale reliability studies. *Statistics in Medicine*, 37(18), 2734–2752.
- Ooi, E. J. (2020a). *The impact of the Practice Guide for Intervention (PGI) on recidivism among offenders*

-
- servicing 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/>>.
- Revelle, W. (2022) *psych: Procedures for Personality and Psychological Research*, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych> Version = 2.2.5.
- Ricks, E. P., & Eno Loudon, J. (2014). The relationship between officer orientation and supervision strategies in community corrections. *Law and Human Behavior, 39*(2), 130-141.
- 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.
- Skeem, J. L., Loudon, 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>
- Skeem, J. L., & Manchak, S. (2008). Back to the future: From Klockars' model of effective supervision to evidence-based practice in probation. *Journal of Offender Rehabilitation, 47*, 220-247. <http://dx.doi.org/10.1080/10509670802134069>
- Tran, N. L., 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.* Sydney, NSW: Corrective Services NSW.
- Trotter, C. (2022). *Working with involuntary clients: A guide to practice.* Routledge.
- Trotter, C., & Evans, P. (2012). An analysis of supervision skills in youth probation. *Australian & New Zealand Journal of Criminology, 45*(2), 255-273.
- Van der Loo, M. P. (2014). The stringdist package for approximate string matching. *The R Journal, 6*(1), 111.
- Weaver, B., & McNeill, F. (2012). Managing cases or supporting change? Some implications of desistance research for case management. *EuroVista: Probation and Community Justice, 2*(2), 97-108.

APPENDIX A

Internal consistency of DRI-SF factors across all DRI-SF forms

	DRI-SF Form	Cronbach's Alpha	95% Confidence Interval	Internal Consistency
Supervisee	Trust	.85	.77 – .91	Good
	Caring / Fairness	.92	.88 – .95	Excellent
	Toughness	.77	.61 – .86	Acceptable
	Total	.90	.85–.93	Excellent
Officer	Trust	.80	.78 – .82	Good
	Caring / Fairness	.83	.79 – .86	Good
	Toughness	.43	.34 – .50	Unacceptable
	Total	.80	.78 – .82	Good
Observer	Trust	.81	.78 – .83	Good
	Caring / Fairness	.85	.82 – .87	Good
	Toughness	.39	.26 – .51	Unacceptable
	Total	.81	.79 – .83	Good

Other CRES

Research Titles

Sept 2024	Implementing digital technologies in Prison: Inmates' ongoing experiences of tablet access and connections with family and friends	Sept 2023	Development and validation of the Triage Risk Assessment Scale: Sexual Offending (TRAS:SO)
Sept 2024	Drivers of program participation in NSW correctional centres: A trend analysis of digital tablets and Five Minute Interventions (FMI)	Sept 2023	Review of the birth certificate Identity Management Enhancement pilot
June 2024	Full body x-ray scanners at NSW correctional centres: Scanner activity, contraband, and behavioural outcomes	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 2024	Full body x-ray scanners at NSW correctional centres: Staff and inmate perspectives	June 2023	The Intensive Learning Centre: Staff experiences of implementation and the SILC reforms
April 2024	Evaluation of throughput and participation outcomes for the Intervention Pathways model	April 2023	Five Minute Interventions (FMI): Custodial staff views and experiences of implementing FMI
March 2024	Evaluation of the Compulsory Drug Treatment Program: Within-treatment change	Dec 2022	Effects of the Workplace Mentor Program on correctional officers' perceptions of workplace culture
Feb 2024	Examination of treatment dosage delivered through the Intervention Pathways Model	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
Dec 2023	Assessment and experience of prison climate in NSW correctional centres	Nov 2022	Effects of Words @ Work training on perceptions of offender rehabilitation and job experiences among Corrective Services Industries (CSI) overseers
Dec 2023	Five Minute Intervention (FMI) skills acquisition by correctional staff: The role of manager buy-in	Sept 2022	Offender participation outcomes and predictors of treatment completion in the High Intensity Program Units (HIPUs)
Dec 2023	Dynamics of wellbeing and needs satisfaction among people in prison	Aug 2022	Staff and inmate experiences of prison social climate at Rapid Build correctional centres: A quantitative evaluation
Dec 2023	Empirical review of the Pathways to Employment (P2E) pilot for women in prison	July 2022	Evaluation of the Alternate Sanctions Program (ASP): Within-treatment Change
Dec 2023	Conducting survey research using inmate digital tablets: Lessons learned from research conducted in NSW correctional centres	June 2022	Implementing digital technologies in prisons: Inmate uptake and perceived value of in-cell digital tablets
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		



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

Telephone: (02) 8346 1556
Email: research.enquiries@dcj.nsw.gov.au