STARR Skill Use and Supervision Outcomes

Jefferson E. Holcomb
Criminal Justice Program, Appalachian State University
Will Hicks
Political Science Program, Appalachian State University
Melissa Alexander
Chief, U.S. Probation Office[1]
Middle District of North Carolina

ALTHOUGH THE COMMUNITY supervision population has declined during the past several years, those on some form of community supervision still represent the largest number of individuals under correctional control, with nearly 3.8 million adults under probation or post-release supervision (Bureau of Justice Statistics, 2023). Research has increasingly demonstrated that despite its diversionary origins and intentions, community supervision can actually contribute to mass incarceration and expand the scope of correctional supervision in the United States (Jacobson et al., 2017; Phelps, 2013). Therefore, it is essential to understand the activities of community supervision officers and to assess officer activities and supervision strategies (Bonta et al., 2008; Labrecque et al., 2023). Otherwise, supervision risks continuing enforcement and surveillance-oriented strategies that, on their own, have demonstrated limited effectiveness in reducing violations or new criminal behavior. Worse, such strategies can lead to increased revocations and a recurring cycle of community and institutional supervision with the concomitant human, social, and financial costs (Horwitz; 2010; Jacobson et al., 2017; Klingele, 2013).

Over the past several decades, a number of supervision training programs have been developed to improve officer use of skills and strategies consistent with evidence-based practice. Typically, these programs are based on core correctional practices of assessing and targeting client risks, focusing interventions on medium- to high-risk individuals, and the use of cognitive-behavioral strategies to improve client motivation and decision-making. Examples of such programs include STICS (Bonta et al., 2019, 2021), EPICS (Labrecque et al., 2015; Smith et al., 2012), Proactive Community Supervision (Taxman, 2008), and STARR (Lowenkamp et al., 2014; Robinson et al., 2012). While extensive research has been conducted on these programs, impact research tends to focus on comparing outcomes between pre-post intervention samples or clients supervised by trained versus untrained officers (Chadwick et al., 2015; e.g. see Bonta et al., 2021; Labrecque et al., 2015; Robinson et al., 2012; Taxman, 2008). Far less common is research that examines the relationship between specific officer intervention skills or activities and supervision outcomes (see Labrecque et al., 2023). This is not surprising, given that examining the "black box" of supervision activities often requires access to observational or other unique data that can be difficult to obtain (but see e.g., Louden et al., 2012; Raynor et al., 2014; Trotter & Evans, 2012). The recent meta-analysis by Labrecque et al. (2023) identified 25 studies published since 1996 examining various

outcomes associated with officer training programs. Specifically, these outcomes included the content of discussion during officer-client interactions, the specific core correctional practice skills used by the officer, and client recidivism. In general, existing research finds that officers who successfully completed training built on core correctional practices (such as STICS, EPICS, and STARR) are far more likely to use evidence-based intervention strategies during interactions with clients. While there was wide dispersion in the confidence intervals for some measures, Labrecque et al.'s (2023) review indicates that clients supervised by program-trained officers have improved supervision outcomes, especially among those supervised by officers with higher levels of program fidelity.

The present study adds to existing research by examining the Staff Training Aimed at Reducing Rearrest (STARR) program that has been widely implemented throughout the U.S. Probation and Pretrial Services system. STARR training seeks to create more constructive interactions between officers and those under their supervision by developing officer skills to help clients improve their decision-making and refrain from future activities that put their supervision, and the community, at risk (Robinson et al., 2012). STARR emphasizes the development and use of supervision skills such as role clarification, effective reinforcement and disapproval,

problem-solving, and teaching and applying the cognitive model. These strategies and practices, largely built on the principles of cognitive-behavioral interventions and motivational interviewing, have been found to be effective in the community supervision of correctional populations and involuntary clients (Bonta et al., 2011; Trotter, 2006).

There has been limited research on the implementation and impact of STARR within the federal probation system. Studies conducted in the early stages of STARR implementation found that STARR was effective at reducing recidivism (Lowenkamp et al., 2014; Robinson et al., 2011; Robinson et al., 2012), and the reduction persisted over a more significant period of time compared to non-STARR supervised cases (Lowenkamp et al., 2014). Viglione et al. (2020) found that surveyed officers reported a high-level knowledge about and support for STARR-based supervision; however, respondents expressed concerns about program implementation similar to those noted in research on the implementation of other officer training programs. Alarid and Jones (2018) found that STARR-supervised clients reported their officers were using key program skills and had generally favorable opinions about their officer's efforts. Viglione and Labrecque (2021) found that policy changes to mandate the use of STARR during client interactions was effective in increasing the use of various STARR skills, though the overall use of STARR was still lower than expected. In their study on drug-testing outcomes in a STARR supervision district, Hicks et al. (2020) found that those supervised by a STARR-trained officer were more likely to have a positive drug test but no more likely to have their supervision revoked for positive tests than non-STARR supervised clients. More importantly, clients supervised by STARR-trained officers were significantly less likely to have their supervision revoked for a new crime (Hicks et al., 2020). This is consistent with a study by Labrecque and Viglione (2021), who found that STARR-supervised clients had more positive drug tests but fewer arrests and revocations of supervision compared to matched clients supervised by an officer not trained in STARR. Importantly, Clodfelter et al. (2016) examined the implementation of STARR in the district in the present study. They found that the training was delivered in a manner consistent with implementation best practices, and officers demonstrated high levels of fidelity to training during early

assessments. However, that study provided little evidence about the actual use of STARR skills or whether those skills were associated with client outcomes (Clodfelter et al., 2016).

The present work extends the growing literature on STARR by examining two key questions. First, what is the actual use of specific STARR skills as reported by federal probation officers in the district under study? Second, what is the relationship between the use of specific STARR skills and supervision outcomes?

Sample and Data Measures

Our sample includes 3243 client-terms of supervision in a single federal court district between 2011 to 2019. The sample includes only completed terms of supervision in which the supervision was revoked or ended successfully. This includes 2,938 separate clients, 287 of whom experienced more than one term of supervision. As noted in Table 1, the average term of supervision was slightly more than 950 days, with a minimum of 44 and maximum of 3213 days under supervision.

Outcomes for client terms were measured according to whether a term ended in successful termination (1) or revocation (0). We use a multilevel logit model to explore factors related to the likelihood that a client term ends successfully rather than in revocation. The primary variable of interest, officer STARR skill use, is measured in two ways using data in the federal client management software, Probation and Pretrial Services Automated Case Tracking System (PACTS). First, STARR % is the percentage of supervision contacts over a client's term that included at least one STARR skill during the interaction as reported in the officer's case management notes. Second, STARR type is a measure of the specific skill used during an interaction. The type of STARR skill(s) used in a specific interaction is largely determined by an officer's perception of a client's progress under supervision, a client's attitude and behavior, and other situational factors. Similar to prior studies on assessments of probationer-officer interactions, we collapse different skills into categories consistent with their primary purpose (e.g., Louden et al., 2012; Manchak et al., 2014). STARR is built upon several core activities (see Viglione & Labrecque, 2021). These include sharing essential information about the officer-client relationship, shaping client behavior activities, and building skills to empower and increase prosocial decision-making by clients. Therefore, we separate

STARR skills consistent with their primary function. *Informational Skill* is a measure of the use of Role Clarification, which is the primary informational STARR skill used at the beginning of nearly all supervision terms. Consistent with most cognitive-behavioral approaches, officer efforts to shape or structure client behavior using STARR may be categorized as offering either positive reinforcement or communicating some form of disapproval or reminder about the client's legal status and behavioral expectations. Because these may be in reaction to different client behavior and attitudes, and thus associated with different outcomes, we elected to separate shaping behavior skills into a *Shaping Skill Positive* (the total number of Positive Reinforcement contacts by the supervision officer over a client's term of supervision) and *Shaping Skill Negative* (the total number of Effective Authority, Effective Disapproval, or Effective Punishment contacts). The significance of this methodological approach is discussed in the concluding section. Finally, STARR seeks to empower clients with the tools to improve their own decision-making. Officers are encouraged to build these into their regular interactions with clients rather than in reaction to client behavior. The attribute *Skill Building* is a measure of the number of Problem Solving, Teaching the Cognitive Model, and Applying the Cognitive Model skill contacts reported by officers.

Furthermore, our analysis includes some key covariates noted by prior research as critically important considerations. Client risk (Risk) represents a client's risk level according to their Post-Conviction Risk Assessment (PCRA) score. Consistent with prior studies using PCRA data (e.g., Cohen et al., 2016; Starr & Cohen, 2021), scores were collapsed into quartiles (1 = low, 2 = medium, 3 = medium high, and 4 = high risk) for each client-term of supervision. We segment client's race into three mutually exclusive and exhaustive categories: White, Black, and Other. Sex is measured as female (1) or male (0). We measure client's age as age in years at the beginning of a specific term of supervision, and Supervision Time is the number of days that the client was under supervision. Descriptive statistics for variables used in study analyses are reported in Table 1. As noted in Table 1, 27 percent of client-officer contacts had a reported use of at least one STARR skill. Remarkably, this is the same proportion of client contacts using a STARR skill found by Viglione and Labrecque (2021)

in a different federal district. It is worth noting that the proportion of STARR contacts reported by Viglione and Labrecque (2021) was the proportion following a policy change to increase the use of STARR skills in that district, as only 10 percent of client contacts involved a reported STARR skill use before the policy mandate (Viglione & Labrecque, 2021).

Results

Initial examination of study data raises several interesting observations. First, as reported in Figure 1, officers reported using Positive shaping skills far more frequently than Negative skills. This is consistent with the widely accepted view that cognitive-behavioral interventions should use a higher proportion of positive reinforcing activities compared to negative reinforcement or punishments to be most effective (Bonta & Andrews, 2017). Prior studies of STARR use in other federal districts have reported similar patterns of specific skill use (see Viglione & Labrecque, 2021; Viglione et al., 2020). The sharp increase in the use of all skill types between 2011 and 2014 can largely be attributed to the beginning implementation of STARR in 2011 and its continued development and expansion within the district. By 2015, nearly all officers in the district had been fully trained on the STARR program and the vast majority were deemed to be STARR "proficient" (Clodfelter et al., 2016).[1] However, beginning in 2015, the use of specific STARR skills changed. For example, the use of Positive Shaping Skills generally leveled off and, with minor fluctuations, remained fairly consistent through 2019. While the use of Negative Shaping Skills increased consistently throughout the period under study, this increase was more modest after 2015. The sharp decline in the use of Skill Building after 2014 is most notable. By the end of the study period, Skill Building interactions appear to have become relatively rare and a mere fraction of their previous frequency. Additional comments about this finding will be highlighted in the Discussion and Conclusion section.

Supervision outcomes were modeled using multilevel logistic regression to account for the fact that some individuals had multiple terms of supervision. The two-level model accounts for individuals under supervision (Level 2) who are clustered in terms of supervision (Level 1).

Regarding the primary outcomes of interest, the larger the share of supervision contacts that involve STARR skills, the more likely a term will end successfully. As reported in

TABLE 1.
Descriptive Statistics

Variable	Mean	S.E.	Minimum	Maximum
Success (DV)	0.6685	0.0083	0	1
Risk	2.2091	0.022	1	4
STARR %	0.2746	0.0034	0	1
Informational	0.9892	0.0171	0	12
Skill Building	1.1042	0.0472	0	30
Shaping Positive	2.3259	0.0559	0	29
Shaping Negative	1.0607	0.0335	0	23
White	0.3577	0.0084	0	1
Black	0.6247	0.0085	0	1
Other	0.0176	0.0023	0	1
Female	0.1341	0.006	0	1
Violent Crime	0.0796	0.0048	0	1
Drugs	0.3605	0.0084	0	1
Weapons/Firearms	0.3256	0.0082	0	1
Sex Offenses	0.0126	0.002	0	1
Public Order	0.0086	0.0016	0	1
Obstruction/Escape	0.0157	0.0022	0	1
Immigration/Customs	0.0111	0.0018	0	1
Financial Offenses	0.1428	0.0061	0	1
Traffic/DWI	0.0435	0.0036	0	1
Age	39.3213	0.1914	17.8219	84.5808
Supervision Time (days)	966.4468	8.9787	44	3213

Note: Data include 3,243 client-terms.

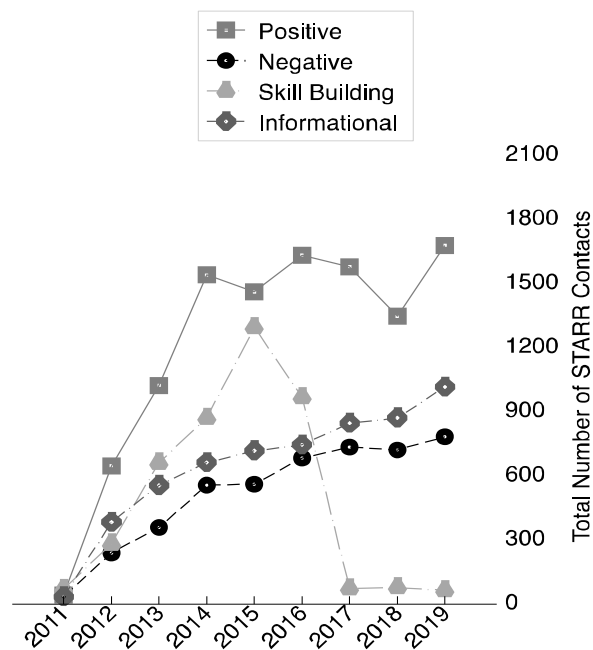**FIGURE 1:**
STARR Skill Use Over Time by Type

Table 2, holding other covariates to their median values, a term of supervision in which 0 contacts included STARR skills had a 0.70 probability of ending successfully. A term in which all contacts included STARR skills had a 0.85 probability of ending successfully. Thus, increasing STARR skills from 0 to 100 percent increased the probability of a successful term of supervision by 15 percentage points.

STARR skill type was also associated with supervision outcome. As noted in Table 2, increased Positive Shaping Skills (.089) were associated with successful terms of supervision, while more frequent use of Negative Shaping Skills (-.24) and Skill Building (-.059) skills were associated with an increased risk of revocation. The use of Informational skills was not associated with supervision outcome.

Not surprisingly, clients whose risk level is higher are significantly less likely to experience a successful expiration of their term

of supervision than clients whose risk level is lower. Holding other covariates to their median values, the likelihood of a successful term of supervision for the lowest risk group is 0.88, while the likelihood of successful term of supervision for the highest risk group is 0.27. Thus, the lowest risk group is 61 percentage points more likely to experience a successful term of supervision than the highest risk group. The federal risk assessment instrument used in the present study (PCRA) has been subject to considerable analysis (Administrative Office of the U.S. Courts, 2018; Cohen et al., 2016; Johnson et al., 2011; Lowenkamp et al., 2013). While the PCRA has recently been revised (i.e., PCRA 2.0; see Serin et al., 2016), the current study highlights the predictive validity of the original measure and the continued importance of including a validated measure of risk in the assessment of supervision outcomes.

The significant relationship between client risk and supervision outcomes warrants additional discussion. In particular, Figures 2-4 (pages 12 and 13) visually display the interactive relationship between the number of contacts involving a particular Skill Type (Positive Shaping, Negative Shaping, and Skill Building), Client Risk, and supervision outcome.[2] Each figure highlights how risk level shapes the relationship between different STARR skills and supervision outcomes. For example, Figure 2 reveals that increasing Positive Shaping Skills is statistically more meaningful for high-risk clients than for low-risk clients. However, increasing Negative Shaping Skill Use for low-risk clients has a more significant impact on this relationship than for high-risk clients. These findings will be discussed more in the concluding section.

Black clients were less likely to complete a term of supervision successfully, but this was not statistically significant. Client gender, however, was related to supervision outcome, as women were over 7.5 percentage points more likely to experience a successful term of supervision relative to men. Older clients were more likely to successfully conclude their term of supervision than younger clients, and the longer the Supervision Time, the more likely a client-term was to end successfully.

Findings on the relationship between Initial Offense and supervision outcome were interesting. Compared to clients who committed violent crimes, clients with DWI/Traffic offenses were the most likely to experience a successful term of supervision, and Sexual offenses were the least likely to conclude successfully. While terms of supervision for all initial offenses other than Sexual offenses were more likely to be completed successfully than terms for Violent offenses, this was statistically significant only for Financial Crimes and DWI/Traffic offenses. Terms for Sex offenses, however, were substantially less likely to end successfully than those for Violent offenses. Notably, terms of supervision for Sex offenses were 40 percentage points less likely to experience successful terms of supervision than terms of supervision for Violent offenses. Prior research on federal sexual offending clients indicates that this category includes a heterogeneous group with significant differences in recidivism rates across specific sex offenses and risk scores (Cohen & Spidell, 2016). The present study did not differentiate type of sex offenses. Furthermore, the present study combines revocation for new offenses and technical violations. Therefore, it

TABLE 2.
STARR Use and Supervision Outcomes

Variable	Coef.	S.E.
Risk	-1.0044**	0.0791
STARR %	0.8644*	0.342
Informational	-0.0721	0.0568
Skill Building	-0.0593**	0.0207
Shaping Positive	0.0888**	0.0221
Shaping Negative	-0.2440**	0.035
Black	-0.0427	0.1186
Other	0.2894	0.4787
Female	0.4305*	0.1898
Drugs	0.3019	0.1969
Weapons/Firearms	0.3520+	0.1922
Sexual Offense	-1.7367**	0.4813
Public Order	0.8621	0.5905
Obstruction/Escape	0.2326	0.4303
Immigration/Customs	1.4981+	0.7772
Financial Offenses	0.8623**	0.2469
Traffic/DWI	2.9343**	0.5179
Age	0.0321**	0.0059
Supervision Time (days)	0.0018**	0.0001
Constant	-0.1144	0.3513
Offender Variance Component	0.0608	0.3042
N	3243	

Note: Outcome represents successful expiration of term (1) versus revocation for any reason (0). Reference categories are violent crime=1 for initial offense, White=1 for race, and male=1 for sex. Coefficients are estimated via multilevel logit, with variance components fit to each client. STARR skills are measured as counts. Risk is measured by dividing client risk based on PCRA into quartiles, 1 = lowest and 4 = highest.
*+ p<0.10, * p<0.05, ** p<0.01*

is possible that the higher revocation for terms of supervision for sex offenses may be due, in part, to more punitive responses to supervision violations and not to higher recidivism rates (see Cohen & Spidell, 2016).

Discussion and Conclusion

The present study sought to assess the relationship between use of STARR skills and supervision outcomes. More importantly, evidence about the *type* of STARR skill used and supervision outcomes could provide helpful insights into officer behavior and supervision outcomes. Results indicate that officers in the district under examination are using a range of STARR skills in interactions with those under their supervision, though perhaps not as consistently as desirable (see also, Viglione & Labrecque, 2021). Numerous studies have found that officers frequently do not fully incorporate training skills in their supervision practice, or at least not in a consistent and sustained manner (Bonta et al., 2019; Gleicher, 2020; Viglione, 2017, 2018). Furthermore, results provide strong evidence of a relationship between type of STARR skill and supervision outcomes. Positive Shaping skill use is associated with successful supervision outcomes and Negative Shaping skill use with supervision failure. However, Figures 2-4 highlight the important interaction between client risk level, STARR skill use, and supervision outcomes. These demonstrate that the use of particular STARR skills is associated with supervision outcome to a varying degree depending on client risk level. Put simply, the relationship between STARR Skill Use and supervision outcomes varies depending on the specific STARR skill and client risk level.

Unfortunately, the current data preclude an assessment of whether findings merely reflect officer behavior in reaction to client behavior or whether the use of directional skills had a preventive impact on future behavior and case outcomes. Officer actions likely influence client behavior but are also influenced by that behavior. This represents a potential problem of endogeneity that is difficult to overcome in social science research. It is likely that both relationships impact the present findings, but traditional means to differentiate causal effects in endogenous environments, such as instrumental variable analysis, were not possible with available data (e.g., Jones & Gondolf, 2002; Rhodes, 2010). Our inclusion of a validated measure of client risk in the model reduces concerns about omitted variable bias, but possible reverse causation or bidirectional

relationship is clearly a possibility with studies on the impact of behavioral programs in real-world settings. Methodologically, the decision to separate shaping skills based on their positive and negative characteristics appears to have merit. Preliminary analyses that collapsed all shaping skills into a single measure revealed a marginal relationship that was not statistically significant. Furthermore, model fitness statistics suggested that the

single measure of shaping skills included attributes that were working in opposite directions and masking the actual relationship.[3] Future studies on such activities should consider the likely endogenous and directional nature of particular officer activities and how that might impact interpretation of study findings.

An interesting finding is the changing use of Skill Building over time and that the use of these STARR skills was associated with

FIGURE 2.
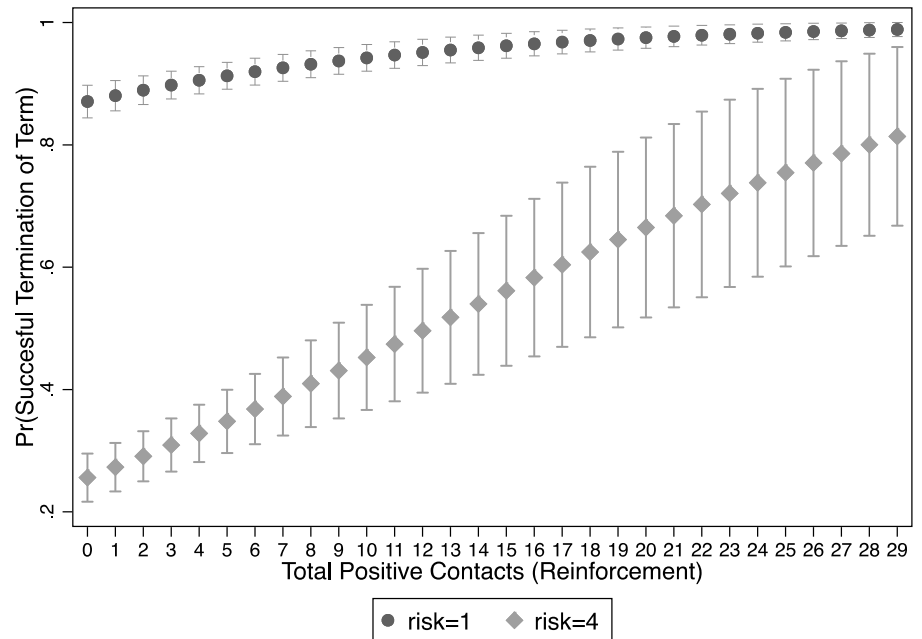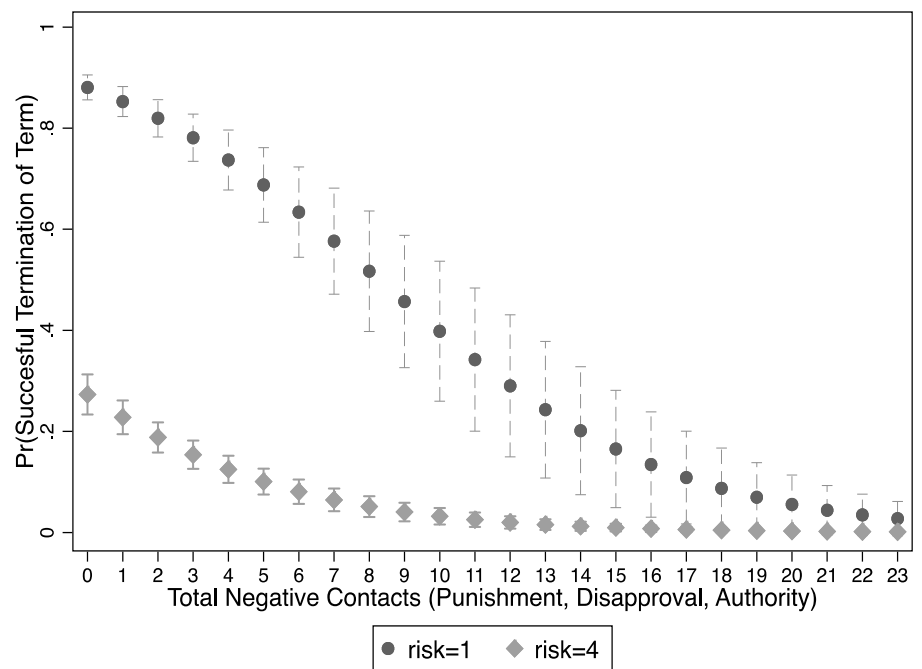**Positive Shaping Skill Use and Supervision Outcomes**

FIGURE 3.
Negative STARR Skill Use and Supervision Outcomes

negative supervision outcomes (see Figure 4). It is unlikely that such skills actually increase revocations. A more likely explanation is that officers are engaging in Skill Building in response to negative client attitude and behavior. This is despite the fact that STARR training emphasizes such structuring skills should be part of normal supervision activities and interactions, rather than in reaction to particular circumstances. If Skill Building is more likely to be used by officers during interactions presenting attitudes or behavior contrary to successful supervision, such skills will likely be associated with supervision terms at increased risk for supervision failure. The impact that the dramatic decline in the use of these skills (see Figure 1) had on present findings is unclear. However, such a change is certainly noteworthy. A review of STARR training modules and comments by district supervisors support the observation that Skill Building interactions tend to be more time intensive and are a more challenging skill for officers to use effectively. This is supported by prior results for STARR showing that building skills, such as Teaching the Cognitive Model and Problem-Solving, are often the least used skills (Viglione & Labrecque, 2021). Additionally, as with any new program, enthusiasm for use of the skills may have declined over time. The sharp decline may also reflect officer preferences to focus on more intuitive and less time consuming Informational 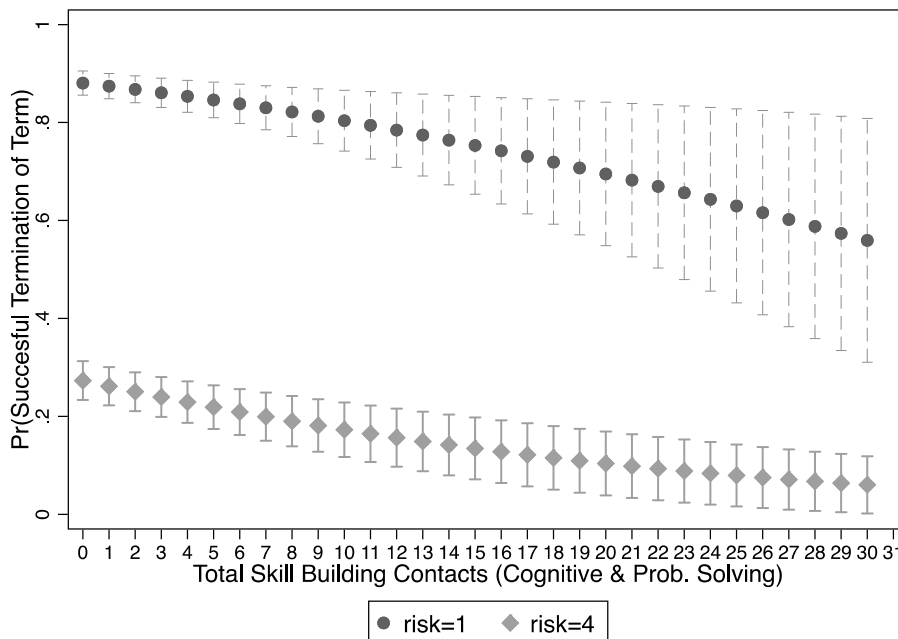and Shaping skill use. The present study was conducted in a district with extensive training, monitoring, and commitment to STARR program fidelity (Clodfelter et al., 2016). However, even in these environments, programs can experience program drift that may dilute program effectiveness. Discussions with the district indicate a renewed emphasis not only on the use of STARR, but particularly on the use of skill building interactions proactively, rather than solely reactively. Evidence of changing patterns of STARR skill use in an organizational setting committed to program fidelity raises questions about such changes in settings without the same level of support, commitment, and accountability. At a minimum, given the extensive research on the effectiveness of cognitive-behavioral interventions, federal districts should pay close attention to possible regression in officer use of Skill Building activities.

Finally, it is worth noting that all study data were pre-pandemic. The pandemic had a significant impact on supervision training and practice (Cohen & Starr, 2021; Mangione & Cohen, 2021; but see Hronick et al., 2021). Extensive research highlights the challenges and critical nature of ongoing monitoring of program implementation during the best of times. The present study suggests that, even in an organizational context supportive of STARR implementation, the use of STARR skills can vary over time. Those involved in the development and training of STARR in their own districts should continue to monitor officer use of specific skills to ensure it is consistent with program fidelity and the potential impact of any changes in that use.

While informative, the present study suffers from important data limitations commonly found in evaluations of officer training on supervision outcomes. First, the study was based on a sample from a single federal district that has implemented STARR consistent with many of the best practices highlighted in the implementation scholarship. This has the advantage of reducing concerns about program implementation, but it may limit the generalizability to other districts where the STARR program has been implemented in a less rigorous or inconsistent manner. Next, relying on officer reporting of STARR skill use, rather than on direct observation, is another source of potential error. Using officer case notes and officer-produced records as sources of data on officer behavior can be problematic. In the present case, this may be mitigated by the district's use of regular fidelity checks and the emphasis on internal data collection and review by district administration. Nevertheless, this remains a possible source of error and bias. Perhaps most importantly, as an ongoing interactive intervention over the course of a term of supervision, the possibility of endogenous causal relationships could not be overcome with available data. We recommend that future research assess the use and impact of officer training in specific cases with repeated observational data. While challenging, such data could help researchers better understand how client behavior impacts officer use of cognitive-behavioral interventions and, more importantly, attempt to untangle the relationships between those interventions, post-intervention client and officer behavior, and supervision outcomes over time. Despite such concerns, the present study indicates that the use of STARR skills in general, and specific skills in particular, is associated with supervision outcome. Future efforts to specify these relationships would provide an important insight into the black box of supervision practices and the impact of officer training programs.

Endnotes

1. Follow-up communications with the chief probation officer from the district confirmed that while all officers continued to participate in training "booster sessions" throughout the study period, all officers were deemed STARR "competent" by the conclusion of the study period. Supporting data available upon request.

FIGURE 4.
Skill Building and Supervision Outcomes

2. In situations where researchers analyze binary outcomes, and do so in reference to a theory that posits an interaction (e.g., the effect of one independent variable depends on the value of another independent variable), it is not uncommon to see researchers include interaction/product terms in their statistical models. Because product terms are the only way to test theories that include the idea of an interaction in the case of ordinary least squares with a continuous dependent variable, researchers often assume the same is true for binary outcome models like the logit and probit. Some statisticians have begun warning researchers that this practice is often misguided and unnecessary (Berry, Golder, & Milton, 2012; Berry, DeMeritt, & Esarey, 2010; Clark & Golder, 2023). Binary outcome models, as in the present study, apply a link function to transform data so that they more appropriately follow the assumptions of linear modeling. In the case of a logistic regression, this link function models the probability an outcome occurs ($Pr(Y=1)/Pr(Y=0)$) by first transforming this probability into a log of the odds ratio. The application of this link function means the entire model is *multiplicative* in the probability an outcome will occur because the model takes the *log* of that probability first. Thus, the impact of the STARR measures reported in Table 2 is based upon the interaction of control variables, including risk, and the predictor variables of interest. Figures 2-4 were created to visualize the interactive relationship between risk, STARR skill, and supervision outcomes.

3. We used two different methods to compare the model we report to one in which we combine positive and negative shaping contacts into a single measure (i.e., total shaping skill use). First, we used a likelihood ratio test, which suggests that separating contact type improves model-fit (LR Chi^2=68.52, p<0.01). Secondly, we also compared AIC and BIC (i.e., Akaike's Information Criterion and Bayesian Information Criterion) measures of model-fit. Lower AIC and BIC values suggest better model-fit and, in both cases, the model that separates shaping skills into positive and negative shaping skills produces lower AIC and BIC values. These tests reveal that although we lose efficiency by using two variables instead of one to measure this type of STARR contact, we gain more accurate insight into how these types of contacts shape supervision outcomes.

References

Administrative Office of the U.S. Courts. (2018). *An overview of the Federal Post Conviction Risk Assessment.* Available at https://www.uscourts.gov/sites/default/files/overview_of_the_post_conviction_risk_assessment_0.pdf

Alarid, L. F., & Jones, M. (2018). Perceptions of offender satisfaction on probation and supervised release with STARR skill sets. *Federal Probation, 82*, 37-41.

Berry, W. D., Golder, M., & Milton, D. (2012). Improving tests of theories positing interaction. *The Journal of Politics, 74*(3), 653-671.

Berry, W. D., DeMeritt, J. H. R, & Esarey, J. (2010). Testing for interaction in binary logit and probit models: Is a product term essential? *American Journal of Political Science, 54*(1), 248-266.

Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct*. Routledge.

Bonta, J., Rugge, T., Scott, T., Bourgon, G., & Yessine, A. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation, 47*, 248–270.

Bonta, J., Rugge, T., Bourgon, G., & Wanamaker, K. A. (2019). A conceptual replication of the Strategic Training Initiative in Community Supervision (STICS). *Journal of Experimental Criminology, 15*, 397-419.

Bonta, J., Bourgon, G., Rugge, T., Scott, T-L., Yessine, A. K., Gutierrez, L., & Li, J. (2011). An experimental demonstration of training probation officers in evidence-based community supervision. *Criminal Justice and Behavior, 38*, 1127-1148.

Bureau of Justice Statistics. (2023). *Correctional populations in the United States, 2021-Statistical tables* (Summary) (NCJ 305542). U.S. Department of Justice.

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.

Clark, W. R., & Golder, M. (2023). *Interaction models: Specification and interpretation*. New York: Cambridge University Press.

Clodfelter, T. A., Holcomb, J. E., Alexander, M. A., Marcum, C. D., & Richards, T. N. (2016). A case study of the implementation of Staff Training Aimed at Reducing Rearrest. *Federal Probation, 80*(1), 30-38.

Cohen, T. H., Lowenkamp, C. T., & VanBenschoten, S. W. (2016). Examining changes in offender risk characteristics and recidivism outcomes: A research summary. *Federal Probation, 80*(2), 57–65.

Cohen, T. H., & Spidell, M. C. (2016). How dangerous are they: An analysis of sex offenders under federal post-conviction supervision. *Federal Probation, 80*(2), 21–32.

Cohen, T. H., & Starr, V. L. (2021). Survey of U.S. Probation and Pretrial Services agencies' adaptations to COVID-19. *Federal Probation, 85*(1), 14–23.

Gleicher, L. (2020). Effective practices in community supervision model: Staff perceptions of the model and implementation. *European Journal of Probation, 12*(3), 157-181.

Hicks, W., Holcomb, J. E., Alexander, M. A., & Clodfelter, T. A. (2020). Drug testing and community supervision outcomes. *Criminal Justice and Behavior, 47*(4), 419-436. doi: https://doi.org/10.1177/0093854819898219

Horwitz, A. (2010). The costs of abusing probationary sentences: Overincarceration and the erosion of due process. *Brooklyn Law Review, 75*(3), 753-790.

Hronick, G., Vernier-Gelven, N., & Starr, V. L. (2021). Risk, need, and responsivity supervision in the pandemic. *Federal Probation, 85*(1), 37–42.

Jacobson, M. P., Schiraldi, V., Daly, R., & Hotez, E. (2017). *Less is more: How reducing probation populations can improve outcomes.* Papers from the Executive Sessions on Community Corrections. Harvard Kennedy School. https://www.hks.harvard.edu/sites/default/files/centers/wiener/programs/pcj/files/less_is_more_final.pdf

Johnson, J. L., Lowenkamp, C. T., VanBenschoten, S. W., & Robinson, C. R. (2011). The construction and validation of the Federal Post Conviction Risk Assessment (PCRA). *Federal Probation, 75*(2), 16-29.

Jones, A. S., & Gondolf, E. W. (2002). Assessing the effect of batterer program completion on reassault: An instrumental variables analysis. *Journal of Quantitative Criminology, 18*(1), 71–98.

Klingele, C. (2013). Rethinking the use of community supervision. *The Journal of Criminal Law and Criminology, 103*(4), 1015-1069.

Labrecque, R. M., Viglione, J., & Caudy, M. (2023). The impact of community supervision officer training programs on officer and client outcomes: A systematic review and meta-analysis. *Justice Quarterly, 40*(4), 587–611.

Labrecque, R. M., & Viglione, J. (2021). The impact of a community supervision officer training program on client outcomes: A propensity score modeling analysis by officer training dosage. *Criminal Justice and Behavior, 48*(3), 315-331.

Labrecque, R. M., Smith, P., & Luther, J. D. (2015). A quasi-experimental evaluation of a model of community supervision. *Federal Probation, 79*(3), 14-19.

Louden, J. E., Skeem, J. L., Camp, J., Vidal, S., & Peterson, J. (2012). Supervision practices in specialty mental health probation: What happens in officer-probationer meetings? *Law and Human Behavior, 36*(2), 109–119.

Lowenkamp, C. T., Holsinger, A., Robinson, C. R., & Alexander, M. (2014). Diminishing or durable treatment effects of STARR? A research note on 24-month re-arrest rates. *Journal of Crime and Justice, 37*(2), 275-283. doi:10.1080/0735648X.2012.753849.

Lowenkamp, C. T., Johnson, J. L., Holsinger, A. M., VanBenschoten, S. W., & Robinson, C. R. (2013). The federal Post Conviction Risk Assessment (PCRA): A construction and validation study. *Psychological Services, 10*(1), 87–96. https://doi-org.proxy006.nclive.org/10.1037/a0030343

Manchak, S. M., Skeem, J. L., Kennealy, P. J., & Louden, J. E. (2014). High-fidelity specialty mental health probation improves officer practices, treatment access, and rule compliance. *Law and Human Behavior, 38*(5), 450–461. https://doi-org.proxy006.nclive.org/10.1037/lhb0000076

Mangione, C., & Cohen, T. H. (2021). The impact of COVID-19 on treatment and testing. *Federal Probation, 85*(1), 58–61.

Phelps, M. S. (2013). The paradox of probation: Community supervision in the age of mass incarceration. *Law & Policy, 35*(1–2), 51–80. https://doi.org/10.1111/lapo.12002

Raynor, P., Ugwudike, P., & Vanstone, M. (2014). The impact of skills in probation work: A reconviction study. *Criminology and Criminal Justice, 14*(2), 235-249. Doi: 10.1177/1748895813494869

Rhodes, W. (2010). Estimating treatment effects and predicting recidivism for community supervision using survival analysis with instrumental variables. *Journal of Quantitative Criminology, 26*(3), 391–414.

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. doi:10.1080/0735648X.2012.674823

Robinson, C., VanBenschoten, S., Alexander, M., & Lowenkamp, C. T. (2011). A random (almost) study of Staff Training Aimed at Reducing Rearrest (STARR): Reducing recidivism through intentional design. *Federal Probation, 75*(2), 57-63.

Serin, R. C., Lowenkamp, C. T., Johnson, J. L., & Trevino, P. (2016). Using a multi-level risk assessment to inform case planning and risk management: Implications for officers. *Federal Probation, 80*(2), 10–15.

Smith, P., Schweitzer, M., Labrecque, R. M., & Latessa, E. J. (2012). Improving probation officers' supervision skills: An evaluation of the EPICS model. *Journal of Crime and Justice, 35*(2), 189-199,

Starr, V. L., & Cohen, T. H. (2021). The use of location monitoring at the post-conviction stage of supervision. *Federal Probation, 85*(3), 34-41.

Taxman, F. (2008). No illusions: Offender change and organizational climate in Maryland's Proactive Community Supervision efforts. *Criminology & Public Policy, 7*(2), 275-302.

Taxman, F. S., Henderson, C., Young, D. W., & Farrell, J. (2012). The impact of training interventions on organizational readiness to support innovations in juvenile justice offices. *Administration of Mental Health Policy and Mental Health Services Research*. doi: 10.1007/s10488-012-0445-5

Trotter, C. (2006). *Working with involuntary offenders: A guide to practice* (2nd ed.). Los Angeles, CA: Sage Publications.

Trotter, C., & Evans, P. (2012). An analysis of supervision skills in youth probation. *Australian & New Zealand Journal of Criminology, 45*(2), 255-273. (AS U Learn)

Viglione, J. (2017). Street-level decision making: Acceptability, feasibility, and use of evidence-based practices in adult probation. *Criminal Justice and Behavior, 44*(10), 1356-1381.

Viglione, J. (2018). A multi-level examination of organizational context on adult probation officer attitudes towards evidence-based practice. *International Journal of Offender Therapy and Comparative Criminology, 62*(5), 1331-1356.

Viglione, J., Alward, L., & Sheppard, D. L. (2020). Staff Training Aimed at Reducing Rearrest: Probation officer attitudes and experiences. *European Journal of Probation, 12*(3), 238-264.

Viglione, J., & Labrecque, R. M. (2021). Core correctional practices in community supervision: An evaluation of policy increase to mandate probation officer use of skills. *International Journal of Offender Therapy and Comparative Criminology, 65*(8), 858-881.