

# An Intermediate Outcome Evaluation of the Thinking for a Change Program



**Department of Rehabilitation and Correction**  
Gary Mohr, Director  
Sara Andrews, Deputy Director of Operations

Written by  
Gayle Bickle, Research Specialist  
Bureau of Research and Evaluation

## **ACKNOWLEDGMENTS**

The author would like to thank the TFAC facilitators for administering surveys to program participants—the study could not have been done without their help. In addition, the assistance of senior researcher Betty Moore and research interns Allison Short, Christopher Keenan, and Tabitha Evans was crucial for administering surveys statewide and compiling the data for the analysis. Finally, the consultation and advice of Assistant Bureau Chief Brian Martin was immensely valuable and appreciated.

## Executive Summary

The research literature on effective offender programming shows that cognitive – behavioral programming creates larger reductions in recidivism than other types of offender programming.<sup>1</sup> In light of this evidence, the ODRC adopted the *Thinking for a Change (TFAC)* program. In 2009, the department encouraged every prison to implement the TFAC program. The program teaches problem-solving skills, particularly when interacting with others, in order to increase rational thinking and lead to pro-social interactions and behaviors. In addition, through cognitive restructuring (aka, cognitive self-change), thought processes are modified to reduce thinking patterns that are conducive to criminal behavior, i.e., antisocial attitudes. This evaluation uses a quasi-experimental, non-random, two group pre-test post-test design, and it explores intermediate outcomes that examine whether the program has influenced participant’s self-assessment of their social problem-solving skills and approaches and their acceptance of criminal attitudes. The Social Problem-Solving Inventory-Revised (SPSI-R) and the Texas Christian University Criminal Thinking Scales (CTS) were used to measure these components of the program. The findings are as follows:

- ❖ Analyzing the SPSI-R data using GLMM (Generalized Linear Mixed Modeling), we find that, compared to a waiting list comparison group, TFAC group completers do significantly better than their comparison group counterparts on every measure, including *positive problem orientation, negative problem orientation, rational problem solving* and associated subscales (*problem definition and formulation, generation of alternative solutions, decision making, solution implementation and verification*), *impulsivity/carelessness style*, and *avoidance style*. Moreover, the level of significance of these findings indicates that TFAC does impact participants’ understanding of social problem solving skills and approaches.
- ❖ Analyzing the TCU-CTS data using GLMM, we find that TFAC group completers do significantly better ( $p < .001$ ) than their comparison group counterparts on all but one criminal thinking scale (measuring *cold-heartedness*). The scales where TFAC completers do significantly better include *entitlement, justification, power orientation, criminal rationality, personal irresponsibility*, and the *total criminal thinking* score. The findings indicate that TFAC participants appear to reduce (or at least not increase) their acceptance of criminal attitudes when compared to non-participants.
- ❖ Interaction effects were found between risk level and TFAC group participants on the SPSI-R survey, indicating the following:
  - Higher risk treatment group members showed significantly more improvement than those in the lower risk comparison group on all but one scale (avoidance style), including *positive problem orientation, negative problem orientation, rational problem solving* and associated subscales (*problem definition and formulation, generation of alternative solutions, decision making, solution implementation and verification*), and *impulsivity/carelessness style*. Lower risk treatment group members showed significantly more improvement on all of the scales than their lower risk comparison counterparts.

- Higher and lower risk treatment group members differed from each other in their level of improvement on different scales (when compared to the lower risk comparison group). The GLMM estimated marginal means show that higher risk participants did better than lower risk participants on the *rational problem solving* scale and all of the associated subscales (*problem definition and formulation, generation of alternative solutions, decision making, solution implementation and verification*).
- Lower-risk participants did better than higher risk participants on scales measuring dysfunctional problem-solving approaches, including *avoidance, impulsivity, and negative problem orientation* (lack of problem solving self-efficacy). While prior research indicates that programs targeting higher-risk offenders have better recidivism outcomes compared to their lower-risk counterparts, one study used the SPSI-R and found that those who recidivated had significantly worse *avoidance, impulsivity, and negative problem orientation* scores than those who did not recidivate. These findings together suggest higher-risk offenders may need more practice developing positive social problem-solving techniques and approaches to gain the confidence needed to approach problems systematically, as challenges rather than threats.
- ❖ Taken as a whole, these findings suggest that program facilitators are doing a good job helping TFAC participants understand the major components of the program. While sample sizes were too small for analysis, the average scores for individual groups varied, suggesting a need to resume the continuous quality improvement process to help maintain the integrity of the program and to identify facilitators who would benefit from additional training in the delivery of cognitive-behavioral programs.
- ❖ These findings show that the TFAC program does help participants develop social problem-solving skills and does reduce their agreement with criminal thinking errors. The next step is to examine the impact program participation has on compliance with prison rules and on post-release criminal recidivism.

## Introduction

In Ohio, the state prison population has increased over 370% in the past three decades, from 13,360 in 1980 to 50,128 today.<sup>i</sup> In addition, over 28% of Ohio prisoners return to an Ohio prison within three years, most of them (84%) for a new offense.<sup>ii</sup> The cost per year of operating such a large system was over 1.5 billion dollars in fiscal year 2013. In recent years, the Ohio Department of Rehabilitation and Correction (ODRC) took bold steps to implement evidence-based programming that prepares inmates for community reentry and a crime-free life. If successful in this effort, the ODRC will reduce the number of ex-offenders committing serious offenses, resulting in enhanced community safety and a reduced state prison population.

The research literature on effective offender programming shows that cognitive – behavioral programming creates larger reductions in recidivism than other types of offender programming.<sup>iii</sup> In light of this evidence, the ODRC adopted the *Thinking for a Change (TFAC)* program, a non-proprietary program developed by Bush, Glick, and Taymans (2002) through the *National Institute of Corrections*. In 2009, the department encouraged every prison to implement the TFAC program, and sought to increase program fidelity by requiring that all of the facilitators receive the four-day facilitator training prior to starting a group. In addition, a continuous quality improvement program was begun, but never fully implemented (i.e., not all facilitators across the state were assessed).

While much research has examined the effectiveness of other cognitive-behavioral programs, there are few outcome evaluations of the TFAC. One study to date does show that TFAC influences problem solving and two studies find it reduces recidivism among program completers.<sup>iv</sup> However, this research was conducted on probationers, and results may be different for prisoners. One meta-analysis (analysis of the findings from many studies) found that in-prison cognitive-behavioral programs reduce recidivism by almost 8 percent, but community based cognitive-behavioral programs with an aftercare component have a greater effect on recidivism (a 10.8 percent reduction).<sup>v</sup> However, another meta-analysis did not find that cognitive-behavioral treatment provided towards the end of the prison sentence was any different in effect from treatment provided in the community.<sup>vi</sup> Program facilitators in the ODRC do try to select higher risk offenders with less than three years left on their sentence to participate in the TFAC, but not all participants meet these criteria. In this study's sample, only about half of the TFAC participants had three years or less left on their sentence.

This study is the first in a three-phase evaluation of the program, which will include:

- 1) an intermediate outcome analysis evaluating whether the program has influenced participant's self-assessment of their social problem-solving skills and approaches, and their acceptance of criminal attitudes;
- 2) an outcome evaluation of whether the program impacts in-prison behavior;
- and 3) an outcome evaluation of whether the program reduces recidivism.

Examining intermediate outcomes is important to identify whether this program can change how offenders think about difficult problems they encounter and their attitudes about criminal rationales. Since cognitive-behavioral programming purports that how we think drives how we behave, we consider this to be a crucial step in evaluating whether this program can have an effect on more observable behaviors in prison and in the community.

## **Cognitive-Behavioral Programming and the TFAC Program**

Cognitive-behavioral programming rests on the assumption that how we think controls our moods, feelings, values, attitudes, and ultimately, our behavior. These programs help to identify the thought processes that lead to negative feelings and deviant behaviors and replace them with processes that lead to positive feelings and pro-social behaviors. Some distorted thought processes (often called criminal thinking errors) that are characteristic of criminal behavior include justifications for the behavior, a lack of empathy leading to misinterpretation of interpersonal interactions, feelings of entitlement, impulsivity, irrational thought-processes, and the like. Through cognitive restructuring (aka, cognitive self-change), thought processes are modified to reduce thinking patterns that are conducive to criminal behavior, e.g., antisocial attitudes. Problem-solving skills, particularly when interacting with others, are taught in order to improve rational thinking and lead to pro-social interactions and behaviors. The method requires the use of social learning techniques, and employs role-playing and modeling. Role-playing is used to illustrate critical ways of thinking, social skills, and problem-solving skill concepts. This is an essential component of the programming—it allows participants to practice their new skills in class as well as outside of class, thus increasing the potential impact of the cognitive-behavioral training.

In the TFAC program, two facilitators model social skills (e.g., active listening) and social problem-solving examples through role-playing, and participants are assigned homework where they practice the steps of the skill staff demonstrated. During the next session, participants role-play their homework scenarios, and facilitators interactively help them identify the positive and negative thought processes and underlying beliefs, attitudes, and feelings influencing anti-social and pro-social behaviors.<sup>vii</sup>

Why is cognitive-behavioral programming widely used in the field? Because research examining findings from several evaluations of different types of treatment show that cognitive-behavioral treatments successfully reduce recidivism rates, although the reduction is quite variable.<sup>viii</sup> The large amount of variation in reducing recidivism begs the question of how cognitive-behavioral treatment (CBT) should be implemented. One study by Lipsey, et al. examined over fifty previous CBT evaluations to identify the factors that are contributing to the different recidivism outcomes for cognitive-behavioral programs. What they found was that "...the only factors independently related to effect size were (a) the risk level of the participating offenders [programs targeting higher risk offenders yielded a larger recidivism reduction], (b) how well the treatment was implemented [measured by low dropout rate, monitoring of quality and fidelity, and adequate CBT training], and (c) the presence or absence of a few treatment elements." In the latter category, including anger control and interpersonal problem solving components in the program were associated with larger effects; including victim impact and behavior modification were associated with smaller effects.<sup>ix</sup>

## **Methodological Design**

This research uses a quasi-experimental, non-random, two group pre-test post-test design to obtain intermediate outcome measures of program effectiveness in increasing social problem solving skills and reducing criminal thinking errors among Ohio prisoners completing the program.

Participants in the TFAC program during 2011 were compared on two survey instruments to a sample of prisoners who were on the waiting list for the program. The surveys used were the Social Problem Solving Inventory-Revised and the TCU Criminal Thinking Scales. The comparison group was drawn from the waiting list in order to account for the effect of motivation to participate in the program. The surveys were administered both before and after program participation for the TFAC completers (average 11.7 weeks), and approximately 8.5 weeks apart for the comparison group. A team of researchers and graduate research interns surveyed the comparison group sample from June through October of 2010.<sup>x</sup> For the treatment group, TFAC facilitators were trained to administer the surveys, and given the same script the researchers used to provide the participants with informed consent. Participation in the research was not required for either the treatment or comparison groups.

### ***Sample Selection and Attrition***

The comparison group is composed of inmates who were in prison during May of 2010, on the waiting list for the TFAC program, and scheduled to be released *after* October of 2010. They were sampled from the TFAC waiting lists for 25 of the 31 Ohio prisons that were running the program.<sup>xi</sup> The number randomly selected from each institution waiting list was based on the number of inmates who had started the program in that institution during the previous year, while also estimating the number of inmates who would start the program in the coming year, given the number of staff recently trained to facilitate the program. This led to the identification of 1,476 comparison group subjects. While this is a large number, we expected a lot of drop-off in post-test participation.<sup>xii</sup>

Table 1 (next page) shows the sample attrition pattern for the comparison and treatment groups. Looking first at the comparison group, 91 (six percent) of the 1,476 comparison group sampled<sup>xiii</sup> were not able to participate in the pre-test due to administrative reasons,<sup>xiv</sup> leaving 1,385 inmates who could participate in the research. While most participants took both the SPSI-R and the CTS, the final groups taking the two surveys were not exactly the same due to some surveys being incomplete or otherwise unusable, or the failure to administer one of the surveys after the program (in the case of the treatment group).

A total of 956 participated in the comparison group CTS survey (955 for the SPSI-R), for a participation rate of 69%.<sup>xv</sup> Of those who took the initial CTS, 608 took the post-test and 133 (14%) dropped out for administrative reasons. Finally, 561 of CTS survey participants had both pre- and post- tests that were usable.<sup>xvi</sup> For the second round of SPSI-R surveys, 602 of those who took the SPSI-R pre-test participated in the post-test, while 131 (14%) dropped out for administrative reasons. Finally, 557 of the SPSI-R participants had both pre- and post-test surveys that were usable. While the drop-out rate was relatively high (about one-fifth of participants), it may not have affected results because comparison group members who did not complete a second survey did not significantly differ on either the SPSI-R or the CTS pre-test scale scores from those who completed both surveys.

The treatment group was comprised of those who participated in *Thinking for a Change* during calendar year 2011 from 21 different prisons.<sup>xvii</sup> There were 748 known subjects and 28 of these (3.7%) refused to participate in the pre-test survey, leaving 720 SPSI-R pre-test surveys for

the TFAC group (96.3% of the starting sample).<sup>xviii</sup> A total of 656 (91%) had usable pre-test SPSI-R surveys.<sup>xix</sup> There were 285 post-test SPSI-R surveys completed. Of these, 256 program participants had usable pre- and post-test SPSI-R surveys.<sup>xx</sup> Notably, most of the attrition of the treatment sample was due to program failure (18%), administrative termination (6%), unusable/unmatched surveys (19%) and survey administration failure (15%), while only about 5% refused or did not report for the survey.

**Table 1: Sample Attrition Pattern**

	<b>Treatment Group – SPSI-R</b>	<b>Treatment Group - CTS</b>	<b>Comparison group</b>	
<i>Starting sample</i>	748 (SPSI-R)	729 (CTS)	1,476	
<i>Administrative drop-out<sup>1</sup></i>			91 (6% of starting)	
<i>Remaining sample</i>			1,385	
<i>No shows/Refusals</i>	28 (3.7% of sample)	28 (3.8% of sample)	430 (31%) for SPSI-R & 429 (31%)for CTS	
<i>Total taking initial surveys</i>	720 (96.3% of sample)	701 (96.1% of sample)		
<i>Number in comparison group or unusable<sup>2</sup></i>	64 (36 in comparison group)	33 (27 in comparison group)	<b>SPSI-R</b>	<b>CTS</b>
			955 (69% of remaining)	956 (69% of remaining)
<i>Total after removing unusable initial surveys</i>	656 (91.1% of total taking)	668 (95.3% of total taking)	918 (96% of total taking)	929 (97.2% of total taking)
<i>Program Failure Termination<sup>3</sup></i>	133 (18% of 720)	130 (18% of 701)		
<i>Program termination unknown<sup>3</sup></i>	43 (6%)	43 (6.1%)		
<i>Program Administrative Termination</i>	45 (6%)	42 (6%)		
<i>Failure to Administer/ Lost in Mail</i>	111 (15%)	124 (19%)		
<i>Administrative drop-out from comparison group</i>			131 (14% of 955)	133 (14% of 956)
<i>No Shows/Refusals</i>	39 (5%)	16 (2%)	222 (23%)	215 (22%)
<i>Total taking post-test surveys</i>	285	313	602	608
<i>Number of unusable and unmatched surveys<sup>4</sup></i>	29 (10%)	70 (22%)	43 (7%)	47 (8%)
<i>Number of usable surveys/ matched surveys</i>	256 (90% of post-tests)	243 (78% of post-tests)	557 (93% of post-tests)	561 (92% of post-tests)

<sup>1</sup>For the comparison group, includes segregation placement, current/past participation in the program, medical/dental treatment, out to court, transferred to other institution, early release.

<sup>2</sup>Surveys were unusable if missing too many responses or if all answers were coded in an obvious pattern, e.g., all responses were the same. Surveys were also unusable if the treatment participant was in the comparison group survey sample in order to avoid survey exposure effects.

<sup>3</sup>Program termination type is unknown for 43 individuals, despite repeated attempts to obtain it. Other data suggest that up to six of the 43 were administratively terminated (due to institutional transfers or prison releases, which occurred prior to survey administration). Most of these cases involved female participants from FPRC.

<sup>4</sup>All pre-test surveys were matched to the post-test survey for the same individual.

Of the 701 CTS pre-test surveys for the TFAC group, 668 had usable pre-test surveys. There were 313 post-test CTS surveys completed. Of these, 243 cases had usable pre- and post-treatment surveys. Further analysis (not shown) found that treatment group members who did not participate

in the second survey did not significantly differ from those completing both surveys on either the SPSI-R or the CTS pre-test scales.<sup>xxi</sup>

### ***The Surveys***

This study uses the Social Problem Solving Inventory-Revised (SPSI-R) to measure changes in social problem-solving, a survey instrument measuring aspects of social problem-solving skills and approaches. The SPSI-R has been used in several research studies, including a few studies of prisoners and probationers. For example, an evaluation of a TFAC program for probationers found significant improvements in the total social problem-solving scale of program completers, relative to those who dropped the program as well as those in the comparison group. In addition, an evaluation of another social learning program addressing interpersonal problem-solving found significant improvement in the treatment group for vulnerable adult prisoners, e.g., scores on the SPSI-R scales show they were less negative and less avoidant in their approach to problems than higher-functioning prisoners.<sup>xxii</sup>

The SPSI-R is a 52-item instrument that provides a global indicator of social problem-solving ability and measures two problem-orientation dimensions (positive and negative), and three problem-solving styles (rational, impulsivity-carelessness, and avoidance), with subscales for rational problem solving. It uses a 5-point Likert scale from “Not at all true of me” to “Extremely true of me” to measure agreement with statements. All scales and subscales have been shown to have adequate to high (.69 to .95) internal consistency for different age groups, good test-retest reliability, and moderate to high structural and concurrent validity. Below is a summary of scale measures, as described in the SPSI-R manual:<sup>xxiii</sup>

**Positive Problem Orientation:** higher scores indicate the person is likely to approach problems as challenges, believes they can solve them, believes that solving problems requires commitment, is less likely to experience “emotional distress” when facing problems, and is more successful at solving them.

**Negative Problem Orientation:** higher scores indicate the subject is more likely to approach problems as threats, does not believe she has the ability to solve problems, becomes “frustrated and upset” when encountering them, is more likely to feel distress when confronted with problems, and is less successful at solving them.

**Rational Problem Solving:** higher scores indicate the person is more likely to employ a systematic approach to problem solving, use rational approaches to solve problems, and effectively solve problems. Four subscales comprise this larger scale:

- 1) *Problem Definition and Formulation:* higher scores indicate the person is more likely to collect as much information as possible about a problem, identifying obstacles and creating a specific problem-solving goal.
- 2) *Generation of Alternative Solutions:* higher scores indicate the subject is more likely to generate several possible solutions when solving problems.
- 3) *Decision Making:* higher scores indicate a greater tendency to assess the possible positive and negative consequences of different solutions, and then decide which course of action to take to maximize the positive consequences.

4) *Solution Implementation and Verification*: Higher scores reflect a greater tendency to enact their chosen solution to a problem, and then examine the positive and negative impact of the solution and their feelings about it. If they find the result is poor, they revisit the problem-solving process.

Impulsivity/Carelessness Style: higher scores indicate the person is more likely to approach problems unsystematically, review few alternative solutions, fail to consider the impact of their actions, and act on the first thing that comes to mind. They tend to be less effective at solving problems.

Avoidance Style: those with higher scores are more likely to delay solving problems, hope the problems will resolve themselves, and try to get others to take responsibility for the resolution. They tend to be less effective at solving problems.

The SPSI-R survey is a good measure of whether participants understand and possibly internalize essential components of the TFAC program. The TFAC manual describes the program as a “problem-solving program embellished by both cognitive restructuring and social skills intervention.”<sup>xxiv</sup> Since problem-solving is at its core, TFAC provides steps to learn good problem-solving skills in social situations in sessions entitled “Stop and Think,” “Problem Description,” “Getting Information to Set a Goal,” and “Choose, Plan, Do and Evaluate.” In addition, social skills sessions addressing difficult situations like responding to someone who is angry include elements of good social problem-solving. Sessions like “stop and think” address ways to reduce *impulsiveness*. The scales for *avoidance*, *positive problem orientation*, and *negative problem orientation* can be viewed as measures of self-efficacy, effort, and confidence when approaching problems, which we hope will occur as a result of learning the skills needed for effective social problem-solving.

The TFAC program also focuses on helping participants identify for themselves the beliefs, attitudes, feelings, and thinking that lead to negative or criminal behaviors, and replacing this criminal thinking with alternative thinking. Thus, the program’s authors explain how thinking controls our actions, and include sessions like “Recognizing the Thinking that Leads to Trouble” and “Finding New Thinking.” To measure changes in criminal thinking, this study employs the Texas Christian University Criminal Thinking Scales (CTS) survey. This survey uses a four-point Likert scale from “disagree strongly” to “agree strongly” to measure levels of agreement with statements. The survey provides a total score for criminal thinking, and also contains scales for Entitlement (sense of privilege and identifying desired things as needed things), Justification (minimizes seriousness; believes crime is caused by social injustice), Personal Irresponsibility (blames someone/something else), Power Orientation (uses aggression to control and manipulate others), Cold Heartedness (callous attitude reflecting minimal involvement in relationships), and Criminal Rationalization (negative attitude towards authority figures and belief that the legal system is unfair).<sup>xxv</sup>

### SPSI-R Analysis

It is preferable that the comparison and treatment groups do not differ on pre-test survey scores; this provides more confidence that differences between the groups on post-test surveys are due to exposure to the program rather than initial group differences. The non-random design of this sample makes such differences more likely. The last column in Table 2 (below) shows the treatment group did not differ significantly (t-test,  $p < .05$ ) from the comparison group on any of the pre-test SPSI-R standardized scale scores. In fact, examination of the means show that the treatment group scored the same or slightly worse on all of the scales as the comparison group, although this difference was not significant (note that lower scores on the Negative Problem Orientation, Impulsivity/Carelessness Style, and Avoidance Style scales indicate better results).

**Table 2: Comparison and Treatment Group Differences on Pre-Survey SPSI-R Scales: Means, Standard Deviation, T-score, & Significance Level**

SPSI-R Scale	Group	N	Mean	Std. Deviation	Std. Error	t-score	Sig. (2-tailed)
<b>Positive Problem Orientation, Pre-survey</b>	Treatment	256	99.8789	15.73522	0.98345	-	0.658
	Comparison	557	100.3986	15.13486	0.64128		
<b>Negative Problem Orientation, Pre-survey</b>	Treatment	256	97.5898	17.32978	1.08311	-	0.199
	Comparison	557	95.939	16.25493	0.68874		
<b>Problem Definition &amp; Formulation, Pre-survey</b>	Treatment	256	94.7852	16.7584	1.0474	-	0.113
	Comparison	557	96.7792	16.38772	0.69437		
<b>Generation of Alternative Solutions, Pre-survey</b>	Treatment	256	97.9609	17.52193	1.09512	-	0.158
	Comparison	557	99.8133	16.92273	0.71704		
<b>Decision Making, Pre-survey</b>	Treatment	256	93.7578	17.16848	1.07303	-	0.729
	Comparison	557	94.2065	17.14545	0.72648		
<b>Solution Implementation &amp; Verification, Pre-survey</b>	Treatment	256	96.6758	16.54893	1.03431	-	0.215
	Comparison	557	98.2316	16.67533	0.70656		
<b>Rational Problem Solving, Pre-survey</b>	Treatment	256	95.0195	17.60402	1.10025	-	0.22
	Comparison	557	96.6409	17.22138	0.72969		
<b>Impulsivity/Carelessness Style, Pre-survey</b>	Treatment	256	102.3516	19.2733	1.20458	-	0.141
	Comparison	557	100.2334	18.53995	0.78556		
<b>Avoidance Style, Pre-survey</b>	Treatment	256	98.7578	13.90598	0.86912	-	0.361
	Comparison	557	97.8043	13.5988	0.5762		
<b>Total SPSI-R , Pre-survey</b>	Treatment	256	99.625	16.40684	1.02543	-	0.155
	Comparison	557	101.3698	15.84694	0.67146		

It is also important to test for other differences between the two groups so that those characteristics may be controlled for in the analysis, since they might differentially impact survey scores. Table 3 below presents demographic and other data, and shows that there are no significant differences between the two groups on *reading grade level* or *race*. However, there are significant differences in *age group* ( $p < .01$ ; the treatment group is more likely to be middle aged), *gender* ( $p < .001$ ; the treatment group is less likely to be male), and *risk level*<sup>xxvi</sup> ( $p < .001$ ; the treatment group is more likely to have a higher risk of recidivism).

**Table 3: Comparison and Treatment Group Differences on Demographic and other Variables: Means, Standard Deviation, t-score, Chi-Square, & Significance Level**

Variable	Group	N (%)	Mean	Standard Deviation	Standard Error	t-score	Chi-Square	Sign.(2-tailed)
<b>Reading Grade Level</b>	Treatment	256	9.6691	2.28542	0.14284	-0.634		0.526
	Comparison	557	9.7804	2.40254	0.1018			
<b>Race</b>	<b>African-American</b>	Treatment	129 (50.4%)				0.025	0.874
		Comparison	224 (51%)					
	<b>Caucasian/other</b>	Treatment	127 (49.6%)					
		Comparison	273 (49%)					
<b>Gender</b>	<b>Male</b>	Treatment	202 (78.9%)				25.82	.000**
		Comparison	510 (91.6%)					
	<b>Female</b>	Treatment	54 (21.1%)					
		Comparison	47 (8.4%)					
<b>Age Group</b>	<b>Young Adult (17-39 years)</b>	Treatment	163 (63.7%)				12.91	.005*
		Comparison	381 (68.4%)					
	<b>Middle-Aged (40-55 years)</b>	Treatment	89 (34.8%)					
		Comparison	147 (26.4%)					
	<b>Elderly (56-80 years)</b>	Treatment	4 (1.6%)					
		Comparison	29 (5.2%)					
<b>Risk Level</b>	<b>High Risk</b>	Treatment	125 (48.8%)				18.126	.000**
		Comparison	185 (33.2%)					
	<b>Low Risk</b>	Treatment	131 (51.2%)					
		Comparison	372 (66.8%)					

\* $p < .01$ ; \*\* $p < .001$

Prior research finds that two of these variables differentially impact some scale scores on the SPSI-R: middle aged adults tend to have better social problem-solving ability than younger and older adults, and women score higher (worse) on the negative problem orientation scale while men score higher (better) on the positive problem orientation scale.<sup>xxvii</sup>

Since a pre-test survey was conducted, we were able to examine the impact of these independent variables on the pre-test survey SPSI-R scales in multivariate analyses (not shown). The effect of being in the treatment versus the comparison *group* variable was not significantly associated with scale outcomes in any of the models. However, *race* (African American vs. others) was a statistically significant variable for most of the pre-survey scale outcomes and *reading grade level* was significant for six of the outcomes. Given this impact of *race* and *reading level* on the pre-survey outcome variables, they are retained in the analysis, along with the *gender*, *age group*, *risk level*, and *TFAC group*. Prior research that addresses the possible impact of *reading grade level* or being *African-American* on these scales was not found, but those with higher reading levels may be able to understand the survey questions better or use better overall thought processes, while African-Americans may experience different structural or cultural conditions that influence problem-solving skills and criminal thinking.<sup>xxviii</sup>

### ***Multivariate Analysis for the SPSI-R***

To adjust for potential differences between the treatment and comparison groups, multivariate analysis was used. Multivariate analysis was employed because it allows us to determine which individual characteristics (independent variables) are most strongly related (statistically significant) to an outcome variable, net of the effects of other independent variables. This helps to ensure that the relationship between treatment group participation and the SPSI-R scale is *not* due to the effects of other variables. Generalized Linear Mixed Modeling (GLMM) was used because this method overcomes several possible problems related to non-random sample selection and pre-test post-test research designs, such as violations of normality and assumptions of homogeneity of variances and covariances,<sup>xxix</sup> nested data,<sup>xxx</sup> and inter-correlation problems due to repeated measures of data for the same person.<sup>xxxi</sup> With GLMM, statistical models<sup>xxxii</sup> were analyzed: 1) using statistics suitable to the probability distributions of data (e.g., logistic, log-linear Poisson, or linear regression); 2) accounting for the covariance effects of which institution the group was in; and 3) using an interaction variable in a second set of models. All of the models analyzed here were found to be significant (F-value,  $p < .001$ ), meaning the variables as a group have an effect on the outcome variable that is significantly different from “no effect,” and it is very likely this relationship is not due to chance.

Table 4 (next page) presents the coefficient estimates (the amount of change in scale scores as a result of change in the independent variables) of each variable’s impact on the scale score and their significance levels, net of other variables in the model. The findings show that being in the *TFAC group* has a statistically significant impact on every scale, with *TFAC group* participants doing better than their *comparison group* counterparts. Moreover, the level of significance of the

**Table 4: SPSI-R Scales: Coefficient Estimates (Est.), Standard Errors (SE) and Significance Levels (p) from Generalized Linear Mixed Models**

Model Term	Positive Problem Orientation <sup>a</sup>	Negative Problem Orientation <sup>ac</sup>	Problem Def. & Formulation <sup>a</sup>	Generation of Alternative Solutions <sup>a</sup>	Decision Making <sup>a</sup>	Solution Implementation & Verification <sup>a</sup>	Rational Problem Solving <sup>a</sup>	Impulsivity / Carelessness Style <sup>bc</sup>	Avoidance Style <sup>bc</sup>	Total Score <sup>d</sup>
	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)
T4C Group	.049*** (.010)	-.026** (.008)	.083*** (.011)	.064*** (.010)	.066*** (.008)	.069*** (.011)	.082*** (.011)	-.049*** <sup>c</sup> (.008)	-.018* (.008)	5.163*** (.824)
Male	-.017 (.016)	-.022 (.012)	-.010 (.015)	-.014 (.016)	-.023 (.014)	-.031* (.015)	-.022 (.016)	-.008 (.011)	-.021 (.011)	-.058 (1.168)
High risk	-.008 (.009)	.004 (.008)	-.013 (.010)	-.032 (.010)	-.014 (.008)	-.008 (.010)	-.010 (.010)	.002 (.008)	.001 (.008)	-.237 (.786)
African American	.014 (.010)	-.014 (.008)	.022* (.010)	.017 (.010)	.022** (.008)	.022* (.010)	.020 (.010)	-.016* (.007)	-.008 (.008)	.931 (.796)
Young adult <sup>e</sup>	-.009 (.025)	.031 (.020)	-.017 (.026)	-.044 (.026)	.023 (.021)	-.014 (.025)	-.010 (.026)	.015 (.019)	.031 (.019)	-1.913 (1.964)
Middle aged <sup>e</sup>	-.032 (.026)	.070*** (.021)	-.029 (.027)	-.047 (.026)	.003 (.021)	-.018 (.026)	-.018 (.027)	.062** (.020)	.062** (.020)	-5.675** (2.044)
Reading level	.003 (.002)	-.004** (.002)	.004 (.002)	.002 (.002)	.004* (.002)	.001 (.002)	.002 (.002)	-.008*** (.002)	-.005** (.002)	.492** (.161)
Pretest score	.006*** (.000)	.007*** (.000)	.006*** (.000)	.006*** (.000)	.006*** (.000)	.005*** (.000)	.006*** (.000)	.006*** (.000)	.006*** (.000)	.674*** (.025)
Intercept	4.038*** (.044)	3.920*** (.039)	3.977*** (.044)	4.084*** (.044)	4.044*** (.047)	4.086*** (.044)	3.960*** (.000)	4.037*** (.034)	4.034*** (.039)	31.631*** (3.411)

<sup>a</sup> GLMM fit with a Gamma distribution and log link

<sup>b</sup> GLMM fit with a Poisson distribution and log link

\*p<.05; \*\*p<.01; \*\*\*p<.001; Variables found to have significant coefficients are also shaded yellow.

<sup>c</sup> Lower scores are better outcomes for these scales

<sup>d</sup> GLMM fit with a Normal distribution and an identity link

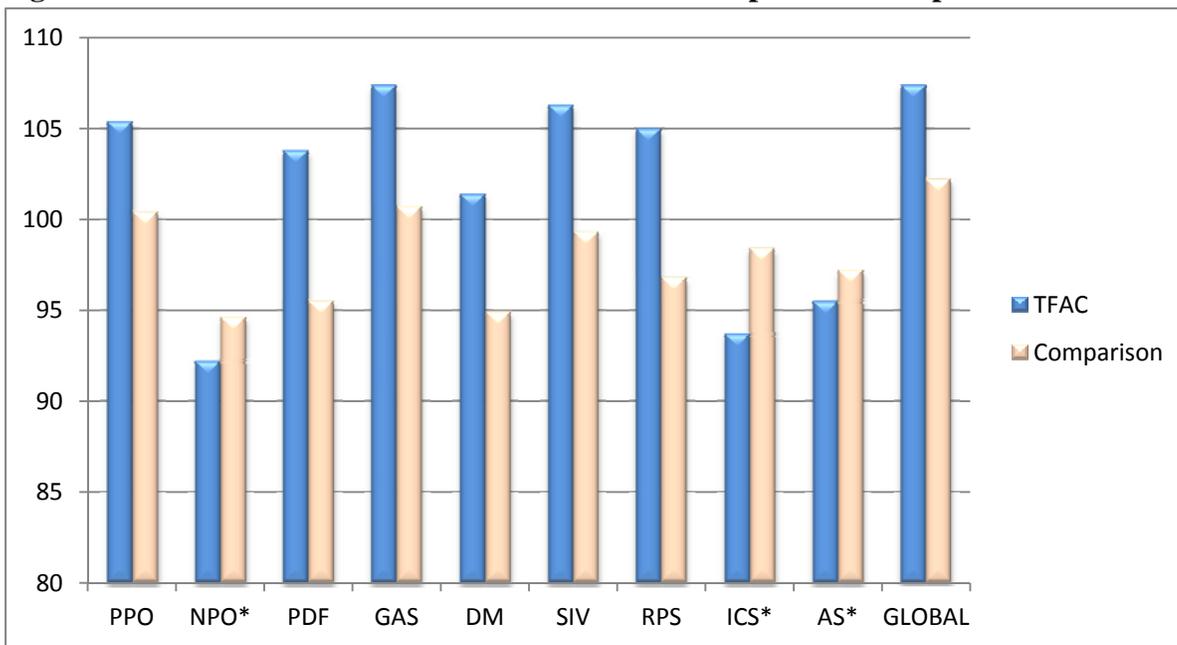
<sup>e</sup> Compared to elderly adults

effect is very high for 8 of the 10 scales; specifically, there is only a very small likelihood ( $p < .001$ ; less than .1%) that the relationship is due to chance.

While the other variables were included in the models primarily as control variables, some findings are interesting and contrary to that of previous research. For example, the *middle aged group* has a significantly higher (worse) *negative problem orientation* ( $p < .01$ ), and score significantly higher (worse) on the *impulsivity/carelessness style* scale and *avoidance style* scale than *older adults*. *Middle aged* subjects also scored significantly worse on the *global measure of problem-solving* than did *older adults*. This is contrary to prior findings that middle-aged adults do better on social problem-solving measures than do other age groups. In addition, the only scale in which *males* were significantly different from females was the *solution implementation & verification* scale, with *males* doing worse than females. *African Americans* scored significantly better than people of other racial groups on four of the scales or rational thinking subscales: *problem definition and formulation*, *decision making*, *solution implementation & verification*, and *impulsivity/carelessness*. Finally, those with higher *reading levels* did better on five of the scales than those with lower *reading levels*. Recidivism risk was not significantly related to any of the scales, and only approaching significance ( $p = .07$ ) on the *decision making* scale.

In order to better understand the impact of being in the *TFAC group* on the SPSI-R scores, estimated adjusted means scores for the group variable are presented in Figure 1 below.

**Figure 1: Estimated Means of the Treatment and Comparison Groups: SPSI-R Outcomes**



\* Lower scores on these scales indicate better results.

These estimated means are net of the effects of other variables in the model. The highest mean differences between the TFAC and comparison groups are found for the subscale *problem definition and formulation* (PDF; 8.3 points) and the *rational problem solving* scale (RPS; 8.2 points). Treatment group members also performed better than the comparison group on the other scales by the following margins: 6.7 points for *generation of alternative solutions* (GAS), 6.5 points for *decision-making* (DM), 7 points for *solution implementation &*

verification (SIV), 4.7 points for *impulsivity/carelessness style* (ICS), 2.4 points for *negative problem orientation*, 1.7 points for *avoidance style* (AS), and 5.2 points for the *global social problem solving score*. Notably, the smallest group differences were for the scales that measure negative approaches to problem solving (negative problem orientation (NPO); problem avoidance (AS), and impulsivity (ICS)).

### ***Multivariate Analysis with Treatment Group by Risk Interaction***

Some of the literature on effective programming finds that higher risk offenders tend to benefit more than lower risk offenders from correctional programs to reduce recidivism, and that the higher the risk, the more programming is necessary to effect a difference or even to ensure they do not react negatively.<sup>xxxiv</sup> For example, recent research found that a relatively low-dose 100-hour program actually made high risk/high need offenders significantly worse than a comparison group of high risk and need offenders who received no treatment.<sup>xxxv</sup> Conversely, if low-risk offenders are provided high intensity programming, the result can be a greater likelihood of recidivism.<sup>xxxvi</sup> The TFAC program would be considered a very low-dose program; if run as intended, it is a 33 hour program lasting approximately three months. However, the program requires homework that might extend the learning period by as much as an hour per session, bringing the total closer to 50 hours. It is possible that participants also received other prison programming during the same timeframe between pre- and post-testing. We did not examine the other in-prison programming because much available data are not reliable. However, we can confidently examine the effect of the interaction between TFAC participation and offender risk on SPSI-R scale scores to see if the program works differently for high versus low risk offenders.

Table 5 (next page) presents the models that include group by risk interaction terms. The first row presents the findings for high risk offenders who participate in the treatment group (compared to low risk offenders in the comparison group), and shows that all scales show greater improvement for the high risk treatment group, and all but one scale (*avoidance style*) is statistically significant. The second row presents findings for low risk offenders in the treatment group and indicates that all scales show statistically significant improvement over lower risk comparison group offenders. The third row shows that high risk offenders in the comparison group scored significantly worse on four of the scales (GAS, DM, SIV, RPS) and significantly better on two (NPO, PDF) compared to the low risk comparison group.

Findings for the control variables remained very similar to the model without the interaction term with one exception: the rational problem solving scale reached significance for African-Americans when compared to other races (primarily ‘white’).

To improve our understanding of how high risk subjects compare to low risk subjects, Figure 2 (page 14) presents the estimated means for the TFAC group by risk interaction terms.

**Table 5: SPSI-R Scales: Models Including the Interaction Effect of TFAC by Risk: Coefficient Estimates (Est.), Standard Errors (SE) and Significance Levels (p) from Generalized Linear Mixed Models**

Model Term	Positive Problem Orientation <sup>a</sup>	Negative Problem Orientation <sup>bc</sup>	Problem Def. & Formulation <sup>a</sup>	Generation of Alternative Solutions <sup>a</sup>	Decision Making <sup>b</sup>	Solution Implementation & Verification <sup>b</sup>	Rational Problem Solving <sup>a</sup>	Impulsivity / Carelessness Style <sup>bc</sup>	Avoidance Style <sup>ac</sup>	Total Score <sup>b</sup>
	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)	Est. (SE)
T4C Group x High Risk <sup>1</sup>	.044*** (.014)	-.022* (.011)	.078*** (.015)	.059*** (.014)	.061*** (.011)	.063*** (.011)	.083*** (.015)	-.040*** (.011)	-.010 (.011)	.046*** (.010)
T4C x Low Risk <sup>1</sup>	.042** (.013)	-.049*** (.011)	.065*** (.014)	.046*** (.014)	.048*** (.011)	.045*** (.011)	.059*** (.014)	-.062*** (.011)	-.027** (.010)	.052*** (.010)
Comparison x High Risk <sup>1</sup>	-.013 (.012)	-.020* (.009)	-.028* (.013)	-.029* (.012)	-.029** (.010)	-.024** (.009)	-.028* (.013)	-.008 (.009)	-.010 (.009)	.000 (.009)
African American	.014 (.010)	-.012 (.008)	.023* (.010)	.018 (.010)	.023** (.008)	.022** (.008)	.021* (.010)	-.015* (.007)	-.009 (.008)	.009 (.007)
Young adult	-.009 (.025)	.03 (.020)	-.016 (.026)	-.044 (.025)	.023 (.021)	-.012 (.021)	-.010 (.026)	.016 (.019)	.031 (.019)	-.019 (.018)
Middle aged	-.032 (.026)	.074*** (.021)	-.028 (.027)	-.047 (.026)	.003 (.021)	-.016 (.021)	-.017 (.027)	.062** (.020)	.061** (.020)	-.055** (.019)
Reading level	.003 (.002)	-.004** (.020)	.004 (.002)	.001 (.002)	.003* (.002)	.001 (.002)	.002 (.002)	-.008*** (.002)	-.005*** (.002)	.005** (.002)
Male	-.017 (.016)	-.022* (.011)	-.01 (.015)	-.014 (.015)	-.024 (.013)	-.031* (.014)	-.022 (.016)	-.008 (.011)	-.021 (.011)	-.001 (.011)
Pretest score	.006*** (.000)	.007*** (.000)	.006 *** (.000)	.006*** (.000)	.006*** (.000)	.005*** (.000)	.006*** (.000)	.006*** (.000)	.006*** (.000)	.007*** (.000)
Intercept	4.039*** (.045)	3.933*** (.036)	3.981*** (.044)	4.088*** (.043)	3.996*** (.035)	4.094*** (.036)	3.964*** (.044)	4.040*** (.034)	4.040*** (.039)	3.924*** (.032)

<sup>a</sup> GLMM fit with a Gamma distribution and log link

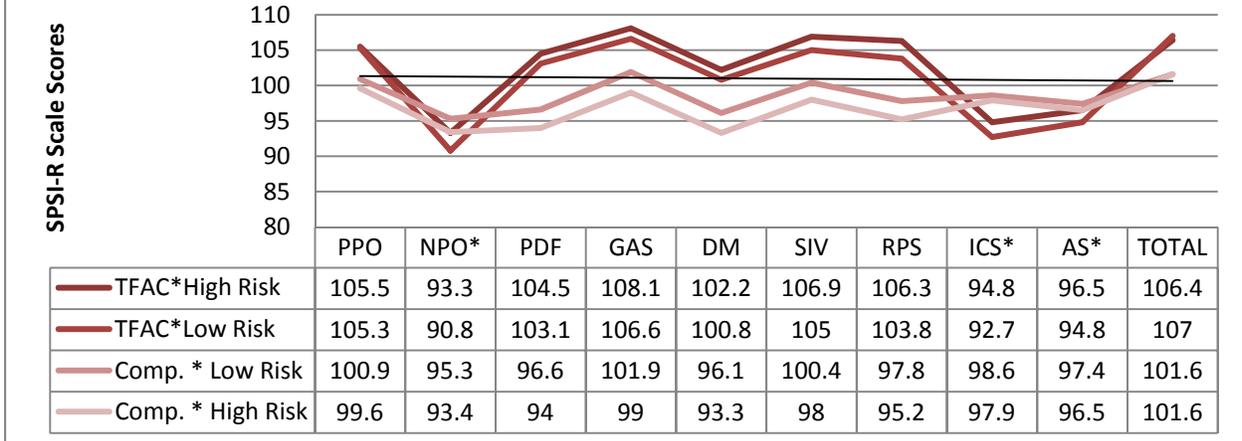
<sup>b</sup> GLMM fit with a Poisson distribution and log link

\*p<.05; \*\*p<.01; \*\*\*p<.001; Variables found to have significant coefficients are also shaded.

<sup>c</sup> Lower scores are better outcomes for these scales

<sup>1</sup>These categories of interaction effects are being compared to comparison group members who are low risk.

**Figure 2: Estimated Marginal Means for SPSI-R:  
Group by Risk Interaction**



\*Lower scores indicate better outcomes for these scales

The darkest line in Figure 2 connects mean scores for the *high risk* TFAC group members, the next darkest for the *lower risk* TFAC group member, and so on. Clearly, those in the TFAC group performed better on most of the scales than did those in the comparison group, regardless of the risk level. Also, higher risk subjects in the TFAC group had better scores than lower risk TFAC group subjects for five of the scales, and had worse scores for four of the scales. Notably, the scales in which the higher risk treatment group subjects did better were the rational problem solving scale along with its subscales. Lower risk TFAC group subjects did better on scales measuring dysfunctional problem-solving approaches or techniques.

### TCU CTS Analysis

In order to examine the Criminal Thinking Scale (CTS) pre-test survey score differences between the treatment and comparison groups, t-tests were conducted (see Table 6). The findings show that there were significant differences between the treatment and comparison groups on five of the seven scales, including *entitlement*, *justification*, *criminal rationalization*, *personal irresponsibility*, and the *total* CTS score. In addition, in every case the TFAC group did better (had lower scores) on the scales than did the comparison group.

**Table 6: GLMM Comparison and Treatment Group Differences on Pre-Survey CTS Scales: Means, Standard Deviation, t-score, & Significance Level of the Coefficients**

CTS Scale	Group	N	Mean	Std. Deviation	Std. Error of mean	t score	Sig. (2-tailed)
<b>Entitlement, Pre-survey</b>	Treatment	243	17.43	5.5712	.3574	2.749	.006*
	Comparison	561	18.66	5.9367	.2507		
<b>Justification, Pre-survey</b>	Treatment	243	18.25	5.2959	.3397	3.001	.003*
	Comparison	557	19.59	6.8170	.2878		
<b>Power Orientation, Pre-survey</b>	Treatment	243	25.21	7.1674	.4598	.775	0.439
	Comparison	557	25.66	7.6589	.3234		
<b>Cold-Heartedness, Pre-survey</b>	Treatment	243	22.27	6.4109	.4113	1.025	.306
	Comparison	557	22.76	6.0981	.2575		
<b>Criminal Rationalization, Pre-survey</b>	Treatment	243	29.64	10.8113	.6934	4.192	.000**
	Comparison	557	32.43	7.5361	.3182		
<b>Personal Irresponsibility, Pre-survey</b>	Treatment	243	19.94	6.3408	.4068	3.810	.000**
	Comparison	557	21.90	6.8764	.2903		
<b>Total Criminal Thinking , Pre-survey</b>	Treatment	243	22.12	4.7571	.3052	3.726	.000**
	Comparison	557	23.5	4.8302	.2039		

\*p<.01; \*\*p<.001

The treatment and comparison samples are slightly different for those who took the CTS and those who took the SPSI-R because there were a few individuals who had usable pre- and post-test SPSI-R scores who did not have usable pre- and post-test CTS scores, and vice-versa. Nevertheless, the groups were quite similar, and the findings for the demographic and other differences between the two groups were the same; that is, there were significant differences between the treatment and control groups for *gender*, *age group*, and *risk level* (results not shown). While there were not differences between the groups on *race* or *reading level*, the impact of these variables on the pre-test survey CTS scales in multivariate analysis showed that race had a statistically significant relationship with five of the pre-test scale outcomes, and reading grade level was significant for four of them. Consequently, the final models include those variables.

### ***Multivariate Analysis of the Criminal Thinking Scales Survey***

GLMM analysis found that all of the models analyzed were statistically significant (F-test, p<.001). The results of the GLMM coefficient estimate analysis are shown in Table 7 (next page). Significant findings are shaded. The results show that, controlling for the effects of other variables in the model, TFAC group subjects performed better than the comparison group on all

but one of the CTS scales (Cold Heartedness). In all cases where the finding is significant, the level of significance is very high, giving us more confidence that the relationship is not due to chance ( $p < .001$ ).

**Table 7: TCU Criminal Thinking Scales: Coefficient Estimates (Est.), Standard Errors (SE) and Significance Levels (p) from Generalized Linear Mixed Models<sup>a</sup>**

Model Term	Entitlement	Justification	Power Orientation	Cold Heartedness	Criminal Rationalization	Personal Irresponsibility	Total Score
	Est. (SE)	Est. (SE)	Est. (SE)				
<b>TFAC Group</b>	-.064*** (.019)	-.095*** (.023)	-.093*** (.019)	-.011 (.017)	-.069*** (.020)	-.083*** (.020)	-.062*** (.013)
<b>Male</b>	.008 (.029)	.051 (.040)	.01 (.037)	.078** (.025)	.078 (.131)	.046 (.028)	.021 (.023)
<b>High risk</b>	-.005 (.018)	.006 (.020)	.001 (.017)	-.029 (.016)	-.004 (.016)	.012 (.018)	.001 (.012)
<b>African American</b>	-.05** (.018)	.029 (.020)	-.035* (.017)	.013 (.016)	-.037* (.016)	-.040* (.018)	-.024* (.011)
<b>Young adult<sup>b</sup></b>	.033 (.047)	.073 (.058)	.078 (.051)	-.053 (.040)	.045 (.059)	.022 (.046)	.013 (.033)
<b>Middle aged<sup>b</sup></b>	.029 (.048)	.045 (.059)	.064 (.051)	-.067 (.042)	.021 (.058)	.009 (.048)	.001 (.033)
<b>Reading level</b>	-.009* (.004)	-.011** (.004)	-.004 (.003)	-.002 (.003)	-.004 (.003)	-.006 (.004)	-.004 (.002)
<b>Pretest score</b>	.034 *** (.001)	.028*** (.002)	.027*** (.001)	.027*** (.001)	.021 *** (.001)	.031*** (.001)	.033*** (.001)
<b>Intercept</b>	2.344*** (.066)	2.435*** (.076)	2.502*** (.065)	2.616*** (.058)	2.824*** (.087)	2.434 (.067)	2.385*** (.049)

<sup>a</sup> GLMM fit with a Gamma distribution and log link

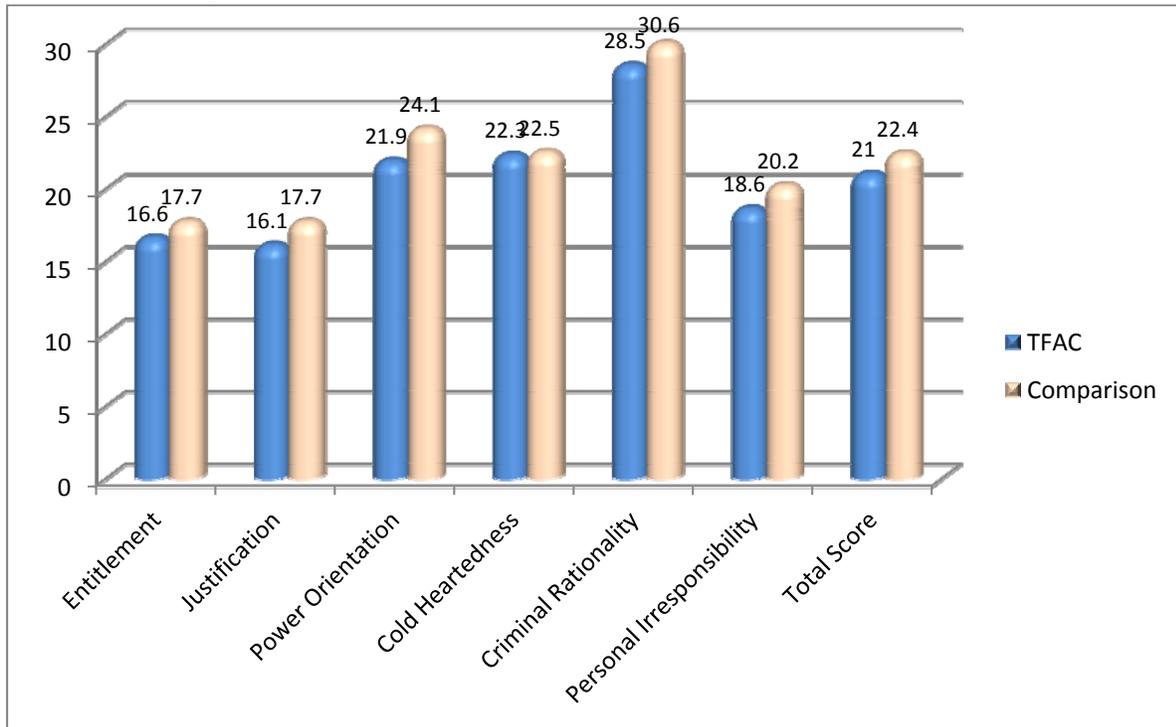
\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ ; Variables found to have significant coefficients are also shaded yellow.

<sup>b</sup> Compared to elderly

Other results show that those with *higher reading levels* do significantly better (score lower) on the justification scale, *males* do worse than females on the *cold-heartedness* scale, and *African Americans* do better than other racial groups on five of the scales: *entitlement*, *power orientation*, *criminal rationalization*, *personal irresponsibility*, and the *total CTS score*. Risk was not significant for any scale, and including a *risk by group* interaction did not result in a significant interaction between *risk* and *group* (findings not shown). One would expect ‘risk’ to be a significant factor, since criminal thinking is believed to be a major driver of criminal behavior, and prior research reports a significant prediction of criminal behavior by criminal attitudes, as well as criminal attitudes by criminal behavior (through socialization in criminal environments such as prison).<sup>xxxvii</sup>

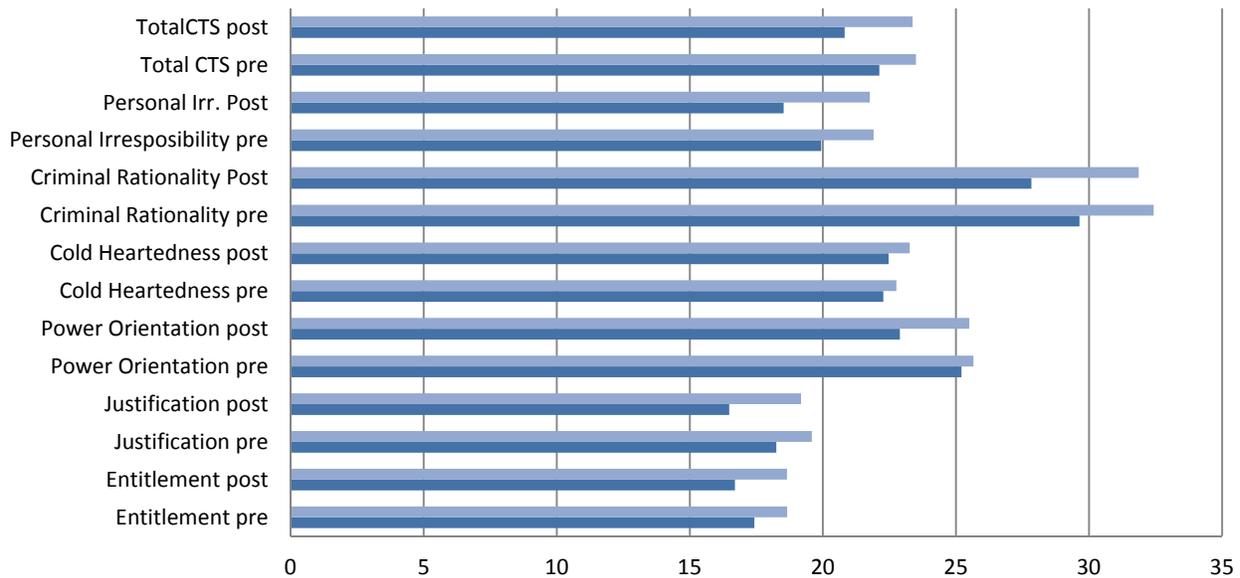
In order to better understand the impact of being in the *TFAC group* on the CTS scales, estimated adjusted means scores are presented in Figure 3 (next page), net of the effects of other variables in the model. The greatest differences between the groups are for the Power Orientation scales (2.2 points) and the Criminal rationality scales (2.1 points).

**Figure 3: Estimated Means of the TFAC Treatment and Comparison Groups for CTS Outcomes**



Recall that the TFAC group scored significantly better (lower) than the comparison group on the CTS scales in the pre-test. In order to visualize the change in the TFAC and comparison group scores between the pre- and post-test, Figure 4 (next page) presents the means for the groups on the two tests. The darker bars represent TFAC means and the lighter bars reflect comparison group means. The table below the chart shows the actual pre- and post- test mean scores for both groups. As can be seen by the chart, the *TFAC group* scores did improve from the pre- to the post-test (ranging from  $-.7$  to  $-2.3$ ) for all but the *cold heartedness* scale, whereas declines for the *comparison group* were minimal (ranging from  $-.1$  to  $-.5$ ). The comparison and TFAC groups both increased their scores on the cold-heartedness scale from the pre- to the post-test.

**Figure 4: Means of Comparison and Treatment Group for Pre- and Post- CTS Scores**



	Entitlement pre	Entitlement post	Justification pre	Justification post	Power Orientation pre	Power Orientation post	Cold Heartedness pre	Cold Heartedness post	Criminal Rationality pre	Criminal Rationality Post	Personal Irresponsibility pre	Personal Irr. Post	Total CTS pre	TotalCTS post
Comparison	18.7	18.6	19.6	19.2	25.7	25.5	22.8	23.3	32.4	31.9	21.9	21.8	23.5	23.4
TFAC	17.4	16.7	18.3	16.5	25.2	22.9	22.3	22.5	29.6	27.8	19.9	18.5	22.1	20.8

## Discussion

The results of this evaluation indicate that prisoners who complete the *Thinking for a Change* program experience a significant improvement in the two major foci of the program: social problem solving skills and criminal thinking errors. In fact, TFAC completers score significantly better than a waiting list comparison group on all of the Social Problem Solving Inventory – Revised scales, including *positive problem orientation, negative problem orientation, rational problem solving* and associated subscales (*problem definition and formulation, generation of alternative solutions, decision making, solution implementation and verification*), *impulsivity/carelessness style, avoidance style*, and the *total social problem solving score*. Program completers also showed significant improvement on most of the TCU Criminal Thinking Scales (*entitlement, justification, power orientation, criminal rationality, personal irresponsibility*, and the *total CTS score*). These findings are more impressive because they include treatment subjects from 21 institutions, not a single program or pilot project.

An interaction effect between *TFAC group* participation and *risk of recidivism* was found for all scales and subscales for the SPSI-R. First, both higher and lower risk treatment group members did significantly better on all but one of the SPSI-R scales than did the lower risk

comparison group. However, the estimated marginal means (which take into account the effect of other variables in the model) show that higher risk treatment group members did better than lower-risk TFAC group members on some scales and subscales, and vice-versa for other scales. Specifically, those in the treatment group who had a higher level of recidivism risk (medium to high risk) improved to a greater extent than low risk treatment participants on the *total* SPSI-R score and the *rational problem solving* scale and associated subscales (*problem definition and formulation, generation of alternative solutions, decision making, solution implementation and verification*). On the other hand, lower risk TFAC group subjects did better than the higher risk treatment group on the *total* SPSI-R score, as well as scales measuring dysfunctional problem-solving approaches and techniques, such as *avoidance, impulsivity, and negative problem orientation* (problem-solving self-efficacy).

Prior research using the SPSI-R on a probation sample found that those who recidivated had significantly higher *avoidance, impulsivity, and negative problem orientation* scores than those who did not recidivate.<sup>xxxviii</sup> Other research finds that prisoners with less effective coping strategies had more disciplinary violations,<sup>xxxix</sup> and the predominant coping strategies for dealing with reentry barriers among those who recidivate are emotion-focused and avoidance-focused, rather than problem-focused strategies.<sup>xl</sup> These findings together suggest that higher risk participants would benefit from developing more positive social problem-solving techniques and approaches. They may need more experience and success performing these techniques than lower-risk offenders in order to gain the confidence needed to see problems as challenges rather than threats to be avoided, and successfully resolve them. As Bandura proposed, “Efficacy expectations determine how much effort people will expend and how long they will persist in the face of obstacles and aversive experiences. The stronger the perceived self-efficacy, the more active the efforts. Those who persist in subjectively threatening activities that are in fact relatively safe will gain corrective experiences that reinforce their sense of efficacy, thereby eventually eliminating their defensive behavior.”<sup>xli</sup>

The finding that the program did not significantly impact the mean scores for the cold-heartedness scale (CTS) for TFAC completers, and that the score increased for both the treatment and comparison groups, is interesting, given that other measures of criminal thinking on the CTS showed a reduction for program completers. Does time in prison increase “cold-heartedness,” which could reduce victim empathy? This finding needs further exploration.

One surprising finding from the CTS analysis was that risk was not significantly related to any of the CTS scales. This is problematic because if the CTS measures criminal thinking, and if criminal thinking is highly associated with criminal behavior, we would expect higher-risk offenders to have higher levels of criminal thinking than lower-risk offenders. One recent study<sup>xlii</sup> also found that none of the CTS scales were significantly related to being classified as a high recidivism risk, or being a “persistent offender” (8 or more lifetime arrests); nor did it have predictive validity, as there was no relationship with arrest recidivism at 6 months. These findings suggest a need to re-examine the efficacy of the CTS for recidivism prediction and identification of the thinking patterns that are related to recidivism.

While we hoped to compare results across groups and institutions, the number of treatment group participants was too small to conduct meaningful statistical analysis. The TFAC program is most often delivered by institutional case managers, whose primary work is to assist prisoners who are having problems. Most of the case managers had not facilitated cognitive-behavioral programs in the ODRC prior to the statewide implementation of the TFAC program. There was some variation among the different TFAC groups in their survey outcomes, suggesting a need to resume the continuous quality improvement process to help maintain the integrity of the program and to assist facilitators as they work to improve the skills needed to successfully facilitate this program. Recently, the department began providing facilitator training that includes a broad range of knowledge and techniques that TFAC facilitators can employ. All case managers are required to take this training. Some supervisors of TFAC facilitators may also benefit from such training, even if they do not facilitate the program. The training could assist them in their assessment of the facilitation skills and knowledge demonstrated by their staff, as well as their ability to coach their staff to improve their performance.

This research is not without problems. First, the participants in the study were not randomly assigned to the treatment and comparison group conditions, making it more likely that some selection bias is occurring in assigning offenders to TFAC groups. We tried to control for motivation to take the program by selecting the comparison group from the program waiting list. We did find that the treatment and comparison groups were similar on most demographic and other factors, the exceptions being gender, age, and risk level, with the treatment group more likely to be female, middle-aged, and have a higher level of risk for recidivism. There were no significant differences on the SPSI-R pre-test scale scores; in fact, treatment group members scored the same or worse than the comparison group. However, the treatment group scored significantly better than the comparison group on several scales of the pre-test CTS. To address these differences, statistical controls for demographic and other variables were used to adjust for potential confounding effects.

Secondly, we could not include those who did not complete the program (and thus, the survey) in the analysis, so we could not determine if different levels of program exposure would have a differential impact. In addition, there was a high degree of dropout from pre- to post-test, especially in the treatment group. This raised concerns about selection bias due to mortality. To address this, we compared the pre-test scores of the treatment and comparison group subjects who left the analysis to their respective group members who took both surveys, and found no differences on any of the pretest scales for either the SPSI-R or the CTS. While this does not inform us about the program impact on those who started but did not complete the program, it suggests that, at least initially, the results for those who took both surveys did not significantly differ from those who only took the pre-test.

Finally, it is possible that having facilitators administer the surveys to the treatment group may have resulted in respondents answering the surveys less honestly than they would if researchers administered the surveys. In particular, it may have influenced pre-test differences

on the CTS survey, where the treatment group scored better on most scales than the comparison group. We did provide training and a script for the facilitators to read that indicated the facilitators would not look at their surveys, but the participant's relationship with the facilitator may have influenced responses despite this promise. However, the changes from pre- to post-test indicate, at the very least, that participants learned how they *should* be solving problems or changing criminal thinking.

Taken together, the findings are very encouraging. The results clearly show that incarcerated TFAC group participants are learning what constitutes good problem-solving approaches and techniques, and are declining in their likelihood of agreeing with "criminal thinking" as defined by the CTS. In other words, the program is providing the intended learning outcomes. These findings also suggest that the facilitators as a group are effectively teaching participants the major components of the program. But, the question remains, was the program intensive enough for participants to internalize and generalize what they learned, and will it impact their future behavior? Reduced recidivism was found for probation samples in prior studies, but no research to date has focused on those receiving the program while in prison. This question will be explored in two future analyses that will examine whether the program significantly impacts in-prison behavior and recidivism after release.

---

<sup>i</sup> D. A. Andrews and James Bonta, *The Psychology of Criminal Conduct*, (Cincinnati, OH: Anderson, 1998), 2<sup>nd</sup> ed.; Nancy Landenberger and Mark Lipsey. "The Positive Effects of Cognitive-Behavioral Programs for Offenders: A Meta-Analysis of Factors Associated with Effective Treatment," *Journal of Experimental Criminology* 1(4) (2005):435-450; Mark Lipsey, Nancy Landenberger, and Sandra Jo Wilson, *Effects of Cognitive-Behavioral Programs for Criminal Offenders: A Systematic Review*, (Nashville, Tenn.: Campbell Systematic Reviews, 2007); David B. Wilson, Leana Allen Bouffard, and Doris MacKenzie, "Quantitative Review of Structured, Group-Oriented, Cognitive-Behavioral Programs for Offenders," *Criminal Justice and Behavior*, 32 (2005): 172-204; Steve Aos, Marna Miller, and Elizabeth Drake, *Evidence-Based Adult Corrections Programs: What Works and What Does Not*. (Olympia, WA: Washington State Institute for Public Policy, 2006); Center for Evidence Based Corrections, *Evidence-Based Practices in Corrections*, (Sacramento, Calif.: Department of Corrections and Rehabilitation, 2006).  
<sup>i</sup> April 22, 2013.

<sup>ii</sup> Konicek, Paul, *ODRC Three-Year Recidivism Rates*, (Columbus, Ohio: ODRC, 2013).

<sup>iii</sup> Andrews and Bonta, *The Psychology of Criminal Conduct*, 1998; Landenberger and Lipsey, "The Positive Effects of Cognitive-Behavioral Programs for Offenders," 2005; Wilson, Bouffard, and MacKenzie, "Quantitative Review of Structured, Group-Oriented, Cognitive-Behavioral Programs for Offenders," 2006; Aos, Miller, and Drake, *Evidence-Based Adult Corrections Programs*, 2006; Center for Evidence Based Corrections, *Evidence-Based Practices in Corrections*, 2006.

<sup>iv</sup> Lori Suzanne Golden, Robert J. Gatchel, and Melissa Anne Cahill, "Evaluating the Effectiveness of the National Institute of Corrections' Thinking For a Change Program Among Probationers," *Journal of Offender Rehabilitation*, 43(2), 2006: 55-73.; Christopher T. Lowenkamp, Dana Hubbard, Matthew D. Makarios, "A Quasi-Experimental Evaluation of Thinking for a Change: A 'Real-World' Application," *Criminal Justice and Behavior*, 36 (2), (2009): 137-146. 2009.

<sup>v</sup> Aos, Miller, and Drake, *Evidence-Based Adult Corrections Programs*, 2006.

<sup>vi</sup> Landenberger and Lipsey, "The Positive Effects of Cognitive-Behavioral Programs for Offenders," 2005: 471.

<sup>vii</sup> Lipsey, et al, 2007, 4-5. Lori Golden, *Evaluation of the Efficacy of a Cognitive-Behavioral Program for Offenders on Probation: Thinking for a Change*, (PhD diss., The University of Texas Southwestern Medical Center at Dallas, 2002), 30-42.

<sup>viii</sup> Robert R. Ross, Elizabeth Fabiano, & Crystal D. Ewles, "Reasoning and Rehabilitation," *International Journal of Offender Therapy and Comparative Criminology*, 32, 1988: 29-35; Wilson, Bouffard and MacKenzie, "A Quantitative Review," 2005.

<sup>ix</sup> Lipsey, Landenberger, and Wilson, *Effects of Cognitive-Behavioral Programs for Criminal Offenders*, 2007.

---

<sup>x</sup> Initially, we thought that most groups met two times per week for 10 to 12 weeks (3 sessions are 'optional'). However, about one-quarter of the sample met in groups that convened once per week. The average amount of time between the pre- and post-test for the twice a week group was about 10 weeks. For the once a week group, it was about 21 weeks. hat served seriously mentally ill inmates (OCF), and four others that did not have very large waiting lists (DCI, MPRC, LORCI & NCCTF). Since the time of sample selection, some institutions have been combined with others, so CMC is now called FMC, and DCI and MPRC have merged into one prison (at least, in terms of administration, n

<sup>xi</sup> Those institutions that were excluded from analysis included one that served seriously ill inmates (CMC), one tot bricks and mortar).

<sup>xii</sup> Some drop-off was expected due to early release, some who started the TFAC program, and others who would not take the follow-up due to lack of interest, failure to understand the purpose of the follow-up survey, administrative reasons (e.g., had a medical appointment), or other reasons.

<sup>xiii</sup> At one institution, the sample was drawn unintentionally from a different program waiting list. Of the 60 inmates who took the survey, only 11 were on the TFAC waiting list prior to the date of the survey; consequently, only those 11 surveys were used in the analysis.

<sup>xiv</sup> Administrative reasons include placement in segregation (n=46), current or past participation in the program (n=29), receiving medical or dental treatment (n=7), being out to court (n=5), and transferred to another institution (n=4).

<sup>xv</sup> The primary loss of participants occurred prior to coming to hear about the survey, and was most evident in institutions with lower security levels. Failure to participate may be due in part to the fact that we excused inmates who were in an earned credit program (for sentence reduction) if taking the survey would impact their earned credit. Of the 479 inmates who did not participate, 149 (31%) were in programs for which they earned credit. However, it is unknown whether the time of survey administration conflicted with that of their earned credit program.

<sup>xvi</sup> Comparing comparison group members who completed both surveys to those who only completed the first survey, we found no initial-survey differences on any of the standardized SPSI-R scales. There was also no significant difference for race, number of prior imprisonments, reading level, and security level. However, there were significant differences in the mean age (with the "no post-test" group being younger; 33.5 vs. 35.6), gender (with women much less likely to take the post-test; 48.5% vs. 62.1% of men), and risk (those completing both had lower risk scores) of those groups. CTS survey results were similar, with no significant difference on the initial survey, gender, race, security level, and reading level. However, there were significant differences in age (with the "no post-test" group being younger), prior imprisonment (those with both surveys were more likely to have a prior imprisonment), and risk (with those with both surveys having lower risk scores).

<sup>xvii</sup> Treatment group participants were surveyed at the second session and the last session of the TFAC program. Drop-off in TFAC participation often occurs after the first session of a group, after participants learn more about group goals and expectations. The first chapter of the TFAC program is primarily an introduction to the program and the cognitive-behavioral approach. All facilitators were trained in the protocol for survey administration. A training document and the informed consent script are included in Appendix A.

<sup>xviii</sup> There were differences in initial numbers for the SPSI-R and CTS samples as a few facilitators forgot to administer both surveys.

<sup>xix</sup> Among the SPSI-R post-test treatment group surveys completed, 29 were not usable due to missing or obviously repetitive responses (e.g., all were identical) or because they did not have a matching pre-test.

<sup>xx</sup> Those in the treatment group who took both the pre- and post- SPSI-R survey were compared on a number of variables with those who only took the pre-survey. No significant differences were found for any of the SPSI-R scales, age, number of priors, minimum sentence, sex race, or marital status. However, there were significant differences in security level, with treatment group members taking both tests more likely to have a higher security level than those taking only the pre-survey. When those of higher vs. lower security levels in the entire treatment group were compared on the survey scales, the only significant difference was that inmates with lower security levels (levels 1 & 2) scored higher on the Avoidance Style scale than those with higher security levels (levels 3, 4, & 5).

<sup>xxi</sup> There was no significant difference between treatment group participants who completed both CTS surveys compared to those who only completed the pre-test survey on race, gender, security level, and reading level. However, there were significant differences on age (with those taking the post-test being older), prior imprisonments (with those taking both more likely to have priors) and risk (with those taking both having a significantly lower risk

---

score). When those with one or more priors were compared to those without any priors, there were no significant differences on any of the CTS scales.

<sup>xxii</sup> Golden, et al., 2006; Jennifer Hayward, Mary McMurrin, & Joselyn Sellen, "Social Problem Solving in Vulnerable adult prisoners: Profile and Intervention, *Journal of Forensic Psychiatry and Psychology*, Vol 19 (2), 2008: 243-248.

<sup>xxiii</sup> Thomas J. D’Zurilla, Arthur M. Nezu, & Albert Maydeu-Olivares, *SPSI-R: Social Problem-Solving Inventory – Revised: Technical Manual*, (North Tonawanda, New York: Multi-Health Systems, Inc), 2002.

<sup>xxiv</sup> Jack Bush, Barry Glick, and Juliana Taymans, *Thinking for a Change: Integrated Cognitive Behavioral Change Program*, (Washington, D.C.: National Institute of Corrections), 2002.

<sup>xxv</sup> Knight, et al., "An Assessment for Criminal Thinking." *Crime & Delinquency*, Vol. 52 (1):159-177., 2006.

<sup>xxvi</sup> Risk scores were obtained from the static risk instrument scored on all prisoners in the department. They were divided into categories of higher vs. lower risk by examining the relationship of the risk score to prior imprisonments. Those who scored three or above were significantly more likely to have prior imprisonments (Chi-square,  $p < .001$ ) than those scoring less than three.

<sup>xxvii</sup> D’Zurilla, et al., "Social Problem-Solving Deficits and Hopelessness, Depression, and Suicidal Risk in College Students & Psychiatric Patients." *Journal of Clinical Psychology*, 54: 1-17, 1998.

<sup>xxviii</sup> The readability of the survey is estimated to be at the fourth grade level. In all, there were fourteen participants who have tested below the fourth grade reading level, 4 (1.6%) in the treatment group and 10 (1.8%) in the comparison group.

<sup>xxix</sup> Initially, we tried to use the repeated measures Generalized Linear Model statistics (similar to MANCOVA), but the Box M statistical test indicated the within group covariance matrices were not equal, which could lead to distorted results. We tested a model where we matched the treatment sample on certain characteristics with individuals in the comparison sample, hoping that sample size equivalency would help resolve possible problems. Still, the Box M test indicated there could be problems. GLMM was used to overcome these problems. The results for the effect of being in the treatment group on the scores of the SPSI-R scales were quite similar between the repeated measures GLM and the GLMM approach. Still, concerns about the accuracy of the statistical estimates generated by the GLM approach led to using the GLMM approach.

<sup>xxx</sup> Nested data refers to variables that could influence individual effects, e.g., the institutions in which the inmates take the TFAC program.

<sup>xxxi</sup> Galwey, Nick W., *Introduction to Mixed Modeling: Beyond Regression and Analysis of Variance*, (Hoboken, NJ : Wiley), 2006.

<sup>xxxii</sup> A statistical model is a formalization of relationships between variables in the form of mathematical equations. A statistical model describes how one or more variables are related to one or more other variables. Modeling can be used for description, prediction, or to assess causality. In this case, we are using the latter approach.

<sup>xxxiii</sup> Donald A. Andrews, Ivan Zinger, Robert Hoge, James Bonta, Paul Gendreau, and Francis T. Cullen (1990) "Does Correctional Treatment Work? A Clinically Relevant and Psychologically Informed Meta-Analysis" *Criminology* 28(3): 369-404; Brian Lovins, Christopher T. Lowenkamp, and Edward J. Latessa. 2009. "Applying the Risk Principle to Sex Offenders: Can Treatment Make Some Sex Offenders Worse?" *The Prison Journal*, 89(3): 344-357; Christopher T. Lowenkamp and Edward J. Latessa. 2005. "Increasing the Effectiveness of Correctional Programming through the Risk Principle: Identifying Offenders for Residential Placement." *Criminology and Public Policy*, 4 (2): 263-290; Lori Brusman Lovins, Christopher T. Lowenkamp, Edward J. Latessa and Paula Smith, "Application of the Risk Principle to Female Offenders." *Journal of Contemporary Criminal Justice*, 23(4): 383-398, 2007.

<sup>xxxiv</sup> Guy Bourgon & Barbara Armstrong, "Transferring the Principles of Effective Treatment into a "Real World" Prison Setting." *Criminal Justice and Behavior* 32(1): 3-25, 2005.

<sup>xxxv</sup> Brian Lovins, Christopher T. Lowenkamp, and Edward J. Latessa. 2009; Christopher T. Lowenkamp and Edward J. Latessa. 2005, p. 277; Brusman Lovins, et al., 2007.

<sup>xxxvi</sup> Paul Gendreau, Tracy Little, Claire Goggin, 1996, "A Meta-Analysis of Predictors of Adult Offender Recidivism: What Works!" *Criminology*, 34: 575-670; Alexander M. Holsinger, 1999, "Assessing Criminal Thinking: Attitudes and Orientations Influence Behavior," *Corrections Today*, 61:22-25; Daniel Boduszek, Christopher G. McLaughlin, & Philip E. Hyland, "Criminal Attitudes of Ex-Prisoners: The Role of Personality, Criminal Friends, and Recidivism," *Internet Journal of Criminology*, ISSN 2045-6743 (downloaded 6/1/2013).

<sup>xxxviii</sup> Golden, et al., 2006.

---

<sup>xxxix</sup> Edward Zamble & Frank J. Porporino, "Coping with Imprisonment," *Canadian Journal of Criminology*, Vol. 26, 403-421.

<sup>xl</sup> Lindsay A. Phillips & Mary Lindsay, "Prison to Society: A Mixed Methods Analysis of Coping with Reentry," Vol. 55(1): 136-154.

<sup>xli</sup> Albert Bandura, "Self Efficacy: Toward a Unifying Theory of Behavioral Change," *Psychological Review*, Vol. 84(2):191-215, 1977.

<sup>xlii</sup> Faye S. Taxman, Anne Giuranna Rhodes & Levent Dumenci, 2011, "Construct and Predictive Validity of Criminal Thinking Scales," *Criminal Justice and Behavior*, Vol. 38 (2): 174-187.

## BIBLIOGRAPHY

Andrews, Donald A., Ivan Zinger, Robert Hoge, James Bonta, Paul Gendreau & Francis T. Cullen, 1990, "Does Correctional Treatment Work? A Clinically Relevant and Psychologically Informed Meta-Analysis," *Criminology*, 28 (3): 369-404.

Andrews, Donald A. & James Bonta, 1998, *The Psychology of Criminal Conduct*. (2<sup>nd</sup> ed.). Cincinnati, Ohio: Anderson Publishing Company.

Aos, Steve, Marna Miller, & Elizabeth Drake, 2006, *Evidence-Based Adult Corrections Programs: What Works and What Does Not*, Olympia: Washington State Institute for Public Policy.

Bandura, Albert, 1977, "Self Efficacy: Toward a Unifying Theory of Behavioral Change," *Psychological Review*, Vol. 84(2):191-215).

Brusman Lovins, Lori, Christopher T. Lowenkamp, Edward J. Latessa, and Paula Smith, 2007, "Application of the Risk Principle to Female Offenders," *Journal Of Contemporary Criminal Justice*, Volume: 23, Issue: 4, Pages: 383-398.

Boduszek, Daniel, Christopher G. McLaughlin, & Philip E. Hyland, "Criminal Attitudes of Ex-Prisoners: The Role of Personality, Criminal Friends, and Recidivism," *Internet Journal of Criminology*, ISSN 2045-6743 (downloaded 6/1/2013).

Bush, Jack, Barry Glick, & Juliana Taymans, 2002, *Thinking for a Change: Integrated Cognitive Behavioral Change Program*.

Chang, Edward C., Thomas J. D'Zurilla, and Lawrence J. Sanna, 2009, "Social Problem Solving as a Mediator of the Link Between Stress and Psychological Well-Being in Middle-Adulthood," *Cognitive Therapy Research*, Vol. 33: 33-49.

D'Zurilla, Thomas J., Edward C. Chang, Edgar J. Nottingham IV, & Lino Faccini, 1998, "Social Problem-Solving Deficits and Hopelessness, Depression, and Suicidal Risk in College Students & Psychiatric Patients," *Journal of Clinical Psychology*, 54: 1-17.

D'Zurilla, Thomas J., Albert Maydeu-Olivares, & Gail L. Kant, 1998, "Age and Gender Differences in Social Problem-Solving Ability." *Personality and Individual Differences*, Vol. 25: 241-252.

D'Zurilla, Thomas J., Arthur M. Nezu, & Albert Maydeu-Olivares, 2002, SPSI-R: *Social Problem-Solving Inventory – Revised: Technical Manual*. North Tonawanda, New York: Multi-Health Systems, Inc.

Galwey, Nick W., 2006, *Introduction to Mixed Modeling: Beyond Regression and Analysis of Variance*, (Hoboken, NJ : Wiley).

Gendreau, Paul, Tracy Little, & Claire Goggin, 1996, "A Meta-Analysis of Predictors of Adult Offender Recidivism: What Works!" *Criminology*, 34: 575-670

Golden, Lori Suzanne, Robert J. Gatchel, & Melissa Anne Cahill, 2006, "Evaluating the Effectiveness of the National Institute of Corrections' Thinking for a Change Program Among Probationers," *Journal of Offender Rehabilitation*, Vol. 43(2): 55-73.

Golden, Lori, 2002, *Evaluation of the Efficacy of a Cognitive-Behavioral Program for Offenders on Probation: Thinking for a Change*. (PhD diss., The University of Texas Southwestern Medical Center at Dallas).

Hayward, Jennifer, Mary McMurrin, & Joselyn Sellen, 2008, "Social Problem Solving in Vulnerable Adult prisoners: Profile and Intervention," *The Journal of Forensic Psychiatry & Psychology*, Vol. 19 (2): 243-248.

Holsinger, Alexander M., 1999, "Assessing Criminal Thinking: Attitudes and Orientations Influence Behavior," *Corrections Today*, 61:22-25.

Knight, Kevin, Bryan R. Garner, D. Dwayne Simpson, Janis T. Morey, & Patrick M. Flynn, 2006, "An Assessment for Criminal Thinking," *Crime & Delinquency*, Vol. 52 (1):159-177.

Konicek, Paul, 2013, *The Ohio Department of Rehabilitation and Corrections' Three-Year Recidivism Rates*. Columbus, Ohio: ODRC.

Landenberger, Nancy & Mark Lipsey (2005). "The Positive Effects of Cognitive-Behavioral Programs for Offenders: A Meta-Analysis of Factors Associated with Effective Treatment," *Journal of Experimental Criminology*, 1 (4): 435-450.

Lipsey, Mark, Nancy Landenberger, and Sandra Jo Wilson., 2007, "Effects of Cognitive-Behavioral Programs for Criminal Offenders: A Systematic Review," *Campbell Systematic Reviews*.

Lovins, Brian, Christopher T. Lowencamp & Edward J. Latessa, 2009, "Applying the Risk Principle to Sex Offenders: Can Treatment Make Some Sex Offenders Worse?" *The Prison Journal*, 89(3): 344-357.

Lowencamp, Christopher T. & Edward J. Latessa, 2005, Increasing the Effectiveness of Correctional Programming through the Risk Principle: Identifying Offenders for Residential Placement, *Criminology & Public Policy*, Vol. 4 (2): 263-289.

Phillips, Lindsay A. & Mary Lindsay, "Prison to Society: A Mixed Methods Analysis of Coping with Reentry," *International Journal of Offender Therapy and Comparative Criminology*, Vol. 55(1): 136-154.

Ross, Robert R., Elizabeth Fabiano, & Crystal D. Ewles, 1988, "Reasoning and Rehabilitation," *International Journal of Offender Therapy and Comparative Criminology*, 32, 1988: 29-35.

Taxman, Faye S., Anne Giuranna Rhodes & Levent Dumenci, 2011, "Construct and Predictive Validity of Criminal Thinking Scales," *Criminal Justice and Behavior*, Vol. 38 (2): 174-187.

Wilson, David B., Leana Allen Bouffard, & Doris MacKenzie, 2005, "A Quantitative Review of Structured, Group-Oriented, Cognitive-Behavioral Programs for Offenders," *Criminal Justice and Behavior*, 32 (2): 172-204.

Zamble, Edward & Frank J. Porporino, "Coping with Imprisonment," *Canadian Journal of Criminology*, Vol. 26, 403-421.

Appendix A  
Training Documents and Informed Consent Script

## **Survey Administration Instructions & Script for the *Thinking for a Change* Evaluation**

### **1. What are the surveys and what are they for?**

There are two surveys we are using to evaluate whether program participants are learning what the program is designed to teach. The *Social Problem-Solving Inventory—Revised* (SPSI-R) evaluates whether the program is improving the way participants solve problems (problem-solving skills and approaches). The Texas Christian University Criminal Thinking Scales (TCUCTS) evaluates whether the program is changing what participants think.

There is an *Exit Survey* that we will use to ask participants' opinions about the program, and for those who left the program before completing it, why they did so (e.g., needed to attend another program or job assignment, etc.).

### **2. How will I get the surveys?**

Contact [gayle.bickle@odrc.state.oh.us](mailto:gayle.bickle@odrc.state.oh.us) (or call 614-752-1267 if that's easier) at least two weeks prior to the class with the start date and information about: 1) how many will be in the class & 2) who is facilitating the class. Also, let me know how long the class will be running so I know when to expect the follow-up surveys.

### **3. When should the surveys be administered?**

The two program evaluation surveys (*SPSI-R* & *TCUCTS*) should be administered at the beginning of the second session (pre-surveys) and then re-administered at the end of the program (post-surveys), preferably the last session. NOTE: If they took the surveys last summer (2010) as part of the comparison group, they should not take the SPSI-R & the TCUCTS again now. The two surveys should take about ½ to ¾ hours to administer, although some people take longer.

For those who complete the program, the *Exit Survey* should be given to them at the end of the last class, along with the pre-addressed interoffice envelopes I will provide. Those who do not complete the class should also get an *Exit Survey*—just e-mail me the ID & names of those who are terminated from the program, and I'll mail it out to them in interoffice mail (if you know why the person was terminated, please let me know). We want all participants to complete this survey, even if they did take the other two surveys last summer as part of the comparison group.

### **4. How will I collect and return the surveys?**

Collect them and place them in the pre-addressed envelope while in the class, and send them back to me with the Survey Checklist form, which indicates that it was your class, who didn't participate, and whether they are the "pre"-surveys or the "post"-surveys. When collecting the surveys, make sure they have their inmate ID on them (the SPSI-R also asks for DOB and date of survey administration). IMPORTANT: When returning the surveys completed at the end of class, please include any surveys that were not used (we had to pay for the use of the SPSI-R's, and will need all of them for the evaluation).

### **5. What will I say to them when I hand out the surveys?**

See the attached instructions (next page) for the program evaluation survey "script." The *Exit Survey* will include a cover letter from me with instructions. You can tell them the *Exit Survey* is a survey asking them what they think about the program, and they can complete it and return it in the pre-addressed envelopes.

## Survey Script

*{To be delivered to the whole group. If anyone arrives late, tell them the same thing, but individually. They will need pencils or pens to complete the surveys}.*

Researchers who are working for the DRC are evaluating the Thinking for a Change program.

They need to know whether the program is working, so they are surveying you about your beliefs and opinions before and at the end of the program. Your answers will be completely confidential, no one other than researchers (not even me) will look at the answers on your survey, and findings will be presented in groups. We do ask for your inmate ID on the form so we can match this survey to the survey at the end of the program. We really hope you will complete the surveys, but you do not have to. If you took the survey this past summer (2010) as part of the comparison group, please do not take it now.

This first survey, the TCUCTS, asks you to choose how much you agree or disagree with statements about your beliefs. It is real important that you be as honest as possible when you circle your response, and that you think about your beliefs as they are NOW. If you need to change an answer, just X out the one you want to change and circle the one you really want.

When you are done with the survey, please put your pencil down and I will come by and pick it up. Are there any questions before we begin? Please be sure to answer every question.

*{pass out the first survey. Once the first survey is completed, collect the first survey, check and make sure the ID is on it, and put it in the envelope & give them the next one. Some may take longer than others on the first one, so the second can be passed out individually with instructions given individually, if you want.}*

The second survey (SPSI-R) asks you to choose the response that shows how much the statement is true of you. A 0 means it is not at all true of you, 1 is slightly true, 2 is moderately true, 3 is very true, and 4 is extremely true of you. Circle the answer that is most true of you. If you need to change an answer, just X out the one you don't want and circle the one you really want.

Please put your inmate ID number in the box where it says "client ID." It is real important that you answer as honestly as possible about how true the statements are of you NOW. Please be sure to answer every question.

*{When they are done, pick up the surveys (check that the ID is on them) and put them in the pre-addressed envelope, fill out the Survey Checklist sheet (enclosed in packet), and send them to me. Remember: return unused surveys after the second survey is administered. THANK YOU! }.*