

Electronic Theses and Dissertations, 2004-2019

2012

Recidivism Measurement And Sanction Effectiveness In Youth Diversion Programs

Thomas T. Maroney
University of Central Florida

 Part of the [Public Affairs, Public Policy and Public Administration Commons](#)
Find similar works at: <https://stars.library.ucf.edu/etd>
University of Central Florida Libraries <http://library.ucf.edu>

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Maroney, Thomas T., "Recidivism Measurement And Sanction Effectiveness In Youth Diversion Programs" (2012). *Electronic Theses and Dissertations, 2004-2019*. 2476.
<https://stars.library.ucf.edu/etd/2476>

RECIDIVISM MEASUREMENT AND SANCTION EFFECTIVENESS IN YOUTH
DIVERSION PROGRAMS

by

THOMAS T. MARONEY

B.S. State University of New York, College at Fredonia, 1995

M.B.A. Canisius College, 1999

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctorate of Philosophy
in the Department of Public Affairs
in the College of Health and Public Affairs
at the University of Central Florida
Orlando, Florida

Summer Term
2012

Major Professor: Thomas T.H. Wan

© 2012 THOMAS T. MARONEY

ABSTRACT

With the rapid growth of juvenile offender diversion programs, which use many non-traditional sanctions, the effectiveness of sanction combinations in juvenile diversion programs and in each individual program needs to be evaluated. Those making sanctioning decisions currently do so based on intuition rather than using an evidence- or theory-based approach. Considerable research has examined the relationship between offender risk factors and recidivism (who is more likely to reoffend?) and between offender risk factors and sanctions (who is more likely to receive what sanctions?), but little is known about the relationship between sanctions and recidivism (which sanctions best reduce recidivism and for whom?). Furthermore, recidivism studies vary drastically in how they measure or quantify recidivism. This variability of approach makes comparing studies difficult and provides a less-than-complete picture of recidivism in general.

The present study used data from one specific youth diversion program to test certain hypotheses of sanctioning by developing and testing a model for assigning sanction combinations to certain offenders on the basis of their individual characteristics. The study first developed measurement models for Offender Risk Propensity, Multiplicity of Sanctions, and Recidivism using structural equation modeling (SEM). Then predictive models were developed to test specific relationships. Understanding the effectiveness of certain sanction packages on certain offenders can form the basis for effective sanctioning in youth diversion programs.

This study sought to answer three research questions: What is the best way to measure recidivism? Does completion of a restorative justice program reduce recidivism? Which sanctions, if any, reduce recidivism for specific offender types? To answer the first question: a

multi-indicator latent construct of recidivism did a very good job of measuring variation in recidivism. Multiple indicators analyzed simultaneously produced a robust tool that can be used in other recidivism studies and help to reduce comparability issues between studies.

The recidivism construct, when tested as a function of completion of the restorative justice program, was seen to produce a significant model having an overall good fit with the data. Thus to answer the second research question: offenders' completion status for the restorative justice program was shown to be a significant predictor of the latent construct of recidivism at the 0.05 level (two-tailed), with those who failed to complete (or chose not to participate) having higher recidivism than did those who completed the program. To answer the third research question: the assignment of specific sanctions (both those suggested by research and theory and those traditionally assigned by this and similar programs) on the entire data set (and on various subsets) of this study have no statistically significant impact on recidivism at the 0.05 level (two-tailed).

The findings suggest many policy implications. Consistency is all but nonexistent in recidivism measurements in the academic literature and in program review studies. A multi-indicator latent construct of recidivism, such as the one proposed and proven effective in this study, provides a more complete picture than simply conceptualizing recidivism by one dummy variable. This recidivism model can be used as the endogenous variable to evaluate programs and their practices and could reduce the problem of study comparability. This could lead to a better understanding of program characteristics and their impact on offender success.

This study also found that completion of the Neighborhood Restorative Justice Program was a significant predictor of recidivism, yet none of the eleven most commonly assigned

sanctions were seen to have a significant impact on recidivism for any subgroup. Proponents of restorative justice argue that it is the programs' characteristics and not their specific activities that make the programs successful. Reintegrative Shaming Theory and Labeling Theory support this claim and suggest the best approach to address youth criminal behavior is to admonish the act and not the actor, have the offender and community agree on a plan to make the community whole after that criminal act, and prevent repeated interaction with the formal criminal justice system which encourages the youth to see themselves as a deviant and engage in further deviant behavior. These characteristics should be further examined and widely employed if confirmed.

Keywords: Restorative justice, recidivism, sanctioning, structural equation modeling

This work is dedicated to the man I most admire, Thomas P. Maroney, Sr.,
my grandfather.

ACKNOWLEDGMENTS

This work would not have been possible without the guidance of a fantastic committee, especially the chair Thomas T. H. Wan, the data provided by Bryan McGuiness of the Ninth Judicial Circuit Court's Restorative Justice Program, and the support of friends and family.

Thank you all.

TABLE OF CONTENTS

LIST OF FIGURES	xii
LIST OF TABLES	xiv
LIST OF ACRONYMS/ABREVIATIONS.....	xvii
CHAPTER 1: STUDY PROBLEM	1
Introduction.....	1
Research Problem	5
Nomenclature.....	6
Significance of Research Problem.....	7
Weaknesses in Current Research.....	9
Specific Research Questions.....	11
CHAPTER 2: LITERATURE REVIEW	12
Theoretical Focus and Informed Perspective.....	12
Contingency Theory.....	13
System Theory	13
Punishment Philosophies and Related Assumptions	14
Application of Literature and Theory on the Current Study.....	15
Conceptualizing Recidivism.....	16
States’ Definitions of Recidivism	17
Researcher’s Definitions of Recidivism	18
Reoffending.....	19
Time Period Examined	22
Data Sources	23
Reoffending Severity and Frequency.....	23
Case Outcome as a Measure of Recidivism.....	24
Reporting Recidivism	25
Application of the Recidivism Literature and Theory to the Current Study.....	27
Restorative Justice Programs	27
Restorative Justice Defined.....	28
Restorative Justice Characteristics.....	30
Theories Related to Restorative Justice	31
Labeling Theory.....	31
Differential Association	32
Reintegrative Shaming Theory	32
Restorative Justice and Recidivism	33
Australian/New Zealand Studies.....	35
U.S. Studies.....	37
Meta-Analysis	38
Selection Bias.....	39
Facilitator Background.....	40
Application of Restorative Justice Literature and Theory on the Current Study.....	41
Offender Risk Propensity.....	43
Prior Arrest History.....	45

Gender	46
School Performance	48
Age	50
Race.....	50
Crime Type	52
Drug Offenders	52
Violent Offenders.....	55
Status Offenders.....	55
Property Offenders.....	56
Other Characteristics.....	56
Application of Offender Risk Propensity Literature and Theory to the Current Study	57
Multiplicity of Sanctions	59
Boot Camps.....	60
Community Service	61
Curfews	62
Letters of Apology	65
Restitution	65
Drug Testing	66
Anger Management.....	67
No-Contact Orders	68
Intense Supervision.....	69
Aftercare.....	69
Application of Literature and Theory on Sanctions to the Current Study	71
Conclusions.....	72
CHAPTER 3: METHODOLOGY	74
Specific Hypotheses Tested	74
Data Source.....	76
Data Security.....	77
IRB Determinations	78
Unit of Analysis	79
Hypothesis Testing Methods	79
Structural Equation Modeling (SEM).....	80
Measurement Model of Recidivism.....	81
Measurement Model of Offender Risk Propensity	88
Measurement Model of Multiplicity of Sanctions	91
Control Variables	92
Predictive Models	94
Explanation of Statistical Design for Testing Hypotheses	94
Model Modification	95
Chi-Square Test	95
Mann-Whitney U Test	97
Hypothesis Testing	98
Hypothesis 1: Recidivism Measurement Model	98
Hypothesis 2: The Data Fit Common Conventions About Recidivism	99

Hypothesis 3: Completion of Restorative Justice Program and Recidivism	100
Hypothesis 4: Drug Testing and Recidivism	100
Hypothesis 5: Curfews and Recidivism	102
Hypothesis 6: No-Contact Orders and Recidivism	102
Hypothesis 7: Letters of Apology and Essays on Recidivism	103
Hypothesis 8: Over-All Main, Moderating, and Mediating Models	104
Measurement Reliability and Validity of Study Variables:	104
Validity	104
Face Validity	105
Content Validity	105
Predictive Validity	106
Reliability Measures	106
Power	107
Addressing Missing Data	108
Descriptive Statistics	110
Descriptive Statistics Related to Hypothesis 1	110
Descriptive Statistics Related to Hypothesis 2	113
Gender and Recidivism	113
Age Group and Recidivism	114
Suspension and Recidivism	114
Descriptive Statistics Related to Hypothesis 3	116
Descriptive Statistics Related to Hypothesis 4	117
Descriptive Statistics Related to Hypothesis 5	118
Descriptive Statistics Related to Hypothesis 6	122
Descriptive Statistics Related to Hypothesis 7	123
Descriptive Statistics Related to Hypothesis 8	125
CHAPTER 4: RESULTS	129
Testing Hypothesis 1	129
Testing Hypothesis 2	131
Testing Hypothesis 3	135
Testing Hypothesis 4	137
Testing Hypothesis 5	138
Testing Hypothesis 6	140
Testing Hypothesis 7	143
Testing Hypothesis 8	147
Exploratory Research	151
Summary of Findings	153
CHAPTER 5: DISCUSSION	155
Limitations	155
Research Design	155
Selection Bias	156
Removal Reason Unknown	156
Generalizability	157
Factors Not Considered	158

Limited Use of Control Variables	159
Sample Size.....	159
Data Censorship	159
Sanction Assignment vs. Sanction Completion	160
Evaluation: Ethical, Punishment Philosophy, and Economic.....	161
Policy Relevance and Theoretical Contributions.....	162
CHAPTER 6: CONCLUSIONS	165
APPENDIX A: TABLES.....	167
APPENDIX B: FIGURES	226
APPENDIX C: INTERNAL REVIEW BOARD FINDINGS.....	251
REFERENCES	253

LIST OF FIGURES

Figure 1: Recidivism Based on Offender Risk Propensity	227
Figure 2: Multiplicity of Sanctions Based on Offender Risk Propensity	227
Figure 3: Simplified Main Effect Model	227
Figure 4: Simplified Mediating Effect Model	228
Figure 5: Simplified Moderating Effect Model	228
Figure 6: Offender Characteristics and Recidivism.....	229
Figure 7: Recidivism as a Function of Sanctions Based on Offender Characteristics.....	229
Figure 8: Measurement Model of the Latent Construct of Recidivism	230
Figure 9: Predictive Model of Recidivism Based on NRJP Completion Status	230
Figure 10: Measurement Model of Offender Risk Propensity.....	231
Figure 11: Predictive Model of Recidivism Based on Offender Risk Propensity	231
Figure 12: Measurement Model of Multiplicity of Sanctions	232
Figure 13: Expanded Main Effect Predictive Model	233
Figure 14: Expanded Mediating Effect Predictive Model	234
Figure 15: Expanded Moderating Effect Predictive Model.....	235
Figure 16: Predictive Model of Recidivism Based on Gender, Age, and Suspension History...	236
Figure 17: Predictive Model of Recidivism Based on Drug Testing Assignment.....	237
Figure 18: Predictive Model of Recidivism Based on Curfew Assignment.....	237
Figure 19: Predictive Model of Recidivism Based on the Assignment of a No Contact Order .	238
Figure 20: Predictive Model of Recidivism Based on Essay and Apology Letter Assignment .	239
Figure 21: Predictive Model of Recidivism Based on the Assignment of Eight Commonly Assigned Sanctions	240
Figure 22: Pie Chart of Severity of New Charge for Entire Sample	241
Figure 23: Results of Testing the Measurement Model of Recidivism With All Six Indicators	241
Figure 24: Results of Testing the Revised Measurement Model of Recidivism With Five Indicators.....	242
Figure 25: Predictive Model of Recidivism Based on Age Group, Gender, and Suspension History.....	242
Figure 26: Revised Predictive Model of Recidivism Based on Age Group and Suspension History.....	243
Figure 27: Predictive Model of Recidivism Based on Completion of the Restorative Justice Program.....	243
Figure 28: Predictive Model of Recidivism Based on the Assignment of Drug Testing	244
Figure 29: Predictive Model of Recidivism Based on the Assigned Curfew Time.....	245
Figure 30: Predictive Model of Recidivism Based on the Assigned Curfew Time Controlled for Age Group.....	245
Figure 31: Predictive Model of Recidivism Based on the Assignment of No-Contact Order....	246
Figure 32: Predictive Model of Recidivism Based on the Assignment of Apology Letters and Essays.....	247
Figure 33: Results of Testing the Revised Main Effect Predictive Model of Recidivism Based on Offender Risk Propensity and Multiplicity of Sanctions.....	248

Figure 34: Results of Testing the Revised Mediating Effect Predictive Model of Recidivism
Based on Offender Risk Propensity and Multiplicity of Sanctions 249

Figure 35: Results of Testing the Revised Moderating Effect Predictive Model of Recidivism
Based on Offender Risk Propensity, Multiplicity of Sanctions, and the Interaction Construct.
..... 250

LIST OF TABLES

Table 1 Time Periods Used in Juvenile Recidivism Studies	168
Table 2 Tests Appropriate According to Variable Types	169
Table 3 Conceptualization, Scale, Type, and Coding for Study Variables.....	170
Table 4 Severity of New Charge Frequencies and Percentages by Category for the Data Set and for Those Who Reoffended.....	172
Table 5 Crime Level Change Observed: Frequencies and Percentages for Entire Data Set	173
Table 6 Number of New Charges, Observed Frequencies, and Percentages for Entire Data Set	174
Table 7 Most Severe Sanctions, Observed Frequencies, and Percentages for Entire Data Set ..	175
Table 8 Pearson Correlations for Indicators of the Recidivism Construct	175
Table 9 Path Analysis of Recidivism Indicators for Measurement Model Using the Entire Data Set	176
Table 10 Path Analysis of Recidivism Indicators for Revised Measurement Model Using the Entire Data Set	176
Table 11 Goodness of Fit Statistics for Recidivism Measurement Model	177
Table 12 Severity of New Charge Categories and Frequencies and Percentages by Gender, Age Group, and Suspension History	178
Table 13 Most Severe Sanction: Frequencies and Percentages by Gender, Age Group, and Suspension History	179
Table 14 Crime Level Change: Frequencies and Percentages by Gender, Age Group, and Suspension History	180
Table 15 Number of New Charges: Frequencies and Percentages by Gender, Age Group, and Suspension History	181
Table 16 Days Charge Free Statistics: Based on Gender, Age Group, and Suspension History	182
Table 17 Pearson Correlations Between Gender, Age Group, and Suspension History	182
Table 18 Path Analysis for Predictive Model of Recidivism Based on Age Group, Gender, and Suspension History	183
Table 19 Goodness of Fit Statistics for Predictive Model of Recidivism Based on Age Group, Gender, and Suspension History	183
Table 20 Chi-Square and Mann-Whitney U Tests for Distributions of Suspension History, Age Group, and Gender and Among the Indicators of Recidivism.....	184
Table 21 Chi-square and Mann-Whitney U Tests for Distributions of Suspension History, Age Group, and Gender and Among Condensed Indicators of Recidivism.....	185
Table 22 Severity of New Charge Categories: Frequencies and Percentages Based on Competition Status in the NRJP	186
Table 23 Most Severe Sanction: Frequencies and Percentages Based on Competition Status in the NRJP	186
Table 24 Crime Level Change: Frequencies and Percentages Based on Competition Status in the NRJP	187
Table 25 Number of New Charges: Frequencies and Percentages Based on Competition Status in the NRJP	188
Table 26 Days Charge Free: Statistics Based on Competition Status in the NRJP	188

Table 27 Path Analysis for Predictive Model of Recidivism Based on Completion of Restorative Justice Program.....	189
Table 28 Goodness of Fit Statistics for Predictors of Recidivism Based on Completion Status in the NRJP	190
Table 29 Chi-Square and Mann-Whitney Tests for Distributions of JRJP Completion Among the Indicators of Recidivism.....	191
Table 30 Severity of New Charge: Frequencies and Percentages Based on the Assignment of Drug Testing for Drug Offenders	192
Table 31 Most Severe Sanction: Frequencies and Percentages Based on Assignment of Drug Testing to Drug Offenders	193
Table 32 Crime Level Change: Frequencies and Percentages Based on Assignment of Drug Testing to Drug Offenders	194
Table 33 Number of New Charges: Frequencies and Percentages Based on Assignment of Drug Testing to Drug Offenders	194
Table 34 Days Charge Free: Statistics Based on Assignment of Drug Testing to Drug Offenders	195
Table 35 Chi-Square and Mann-Whitney Tests for the Distributions of Drug Testing Assignment to Drug Offenders Among Indicators of Recidivism.....	196
Table 36 Severity of New Charge: Frequencies and Percentages Based on the Assigned Curfew Time for Data Set and for Older Male Offenders	197
Table 37 Most Severe Sanction: Frequencies and Percentages Based on the Assigned Curfew Time for Data Set and for Older Male Offenders	198
Table 38 Crime Level Change: Frequencies and Percentages Based on the Assigned Curfew Time for Data Set and for Older Male Offenders	199
Table 39 Frequencies and Percentages Based on the Assigned Curfew Time for Data Set and for Older Male Offenders	200
Table 40 Days Charge Free: Statistics Based on the Assigned Curfew Time for Data Set and for Older Male Offenders	201
Table 41 Path Analysis for Predictive Model of Recidivism Based on the Assigned Curfew Time for Entire Data Set	201
Table 42 Goodness of Fit Statistics for the Predictive Model of Recidivism Based on the Assigned Curfew Time for Entire Data Set and for Older Male Offenders	202
Table 43 Severity of New Charge: Frequencies and Percentages Based on the Assignment of No-Contact Order for Entire Data Set.....	202
Table 44 Most Severe Sanction: Frequencies and Percentages Based on the Assignment of No-Contact Order for Entire Data Set.....	203
Table 45 Crime Level Change: Frequencies and Percentages Based on the Assignment of No-Contact Order for Entire Data Set.....	203
Table 46 Number of New Charges: Frequencies and Percentages Based on the Assignment of No-Contact Order for Entire Data Set	204
Table 47 Days Charge Free: Statistics Based on the Assignment of No-Contact Order for Entire Data Set.....	204
Table 48 Goodness of Fit Statistics for Predictive Model of Recidivism Based on No-Contact Order for Entire Data Set and Gender, Age, and Drug Offender Subgroups	205

Table 49 Chi-Square and Mann-Whitney Tests for Distributions of Indicators of Recidivism Based on the Assignment of No-Contact Order for Entire Data Set	206
Table 50 Frequency and Percentage of Assignment of Essays and Apology Letters for the Entire Data Set.....	207
Table 51 Severity of New Charge: Frequencies and Percentages Based on Assignment of Apology Letters and Essays for the Entire Data Set.....	208
Table 52 Most Severe Sanction: Frequencies and Percentages Based on Assignment of Apology Letters and Essays for the Entire Data Set.....	209
Table 53 Crime Level Change: Frequencies and Percentages Based on Assignment of Apology Letters and Essays for the Entire Data Set.....	210
Table 54 Number of New Charges: Frequencies and Percentages Based on Assignment of Apology Letters and Essays for the Entire Data Set.....	211
Table 55 Days Charge Free Statistics Based on Assignment of Apology Letters and Essays for the Entire Data Set	212
Table 56 Goodness of Fit Statistics for Predictive Model of Recidivism Based on Assignment Apology Letters and Essays for the Entire Data Set, All Females, and Older Female.....	213
Table 57 Chi-Square and Mann-Whitney Tests for Distributions of Indicators of Recidivism Based on Assignment Apology Letters and Essays for the Entire Data Set.....	214
Table 58 Chi-Square and Mann-Whitney Tests for Distributions of Indicators of Recidivism Based on Assignment Apology Letters and Essays for Females at Least 15 Years of Age .	215
Table 59 Pearson Correlations Between Indicators of Offender Risk Propensity.....	216
Table 60 Pearson Correlations Between Indicators of Multiplicity of Sanctions.....	216
Table 61 Pearson Correlation Between Indicators of the Interaction Construct, Offender Risk Propensity, and Multiplicity of Sanctions.....	217
Table 62 Goodness of Fit Statistics for Measurement and Predictive Models	221
Table 63 Path Analysis for Offender Risk Propensity Measurement Model.....	222
Table 64 Path Analysis for Multiplicity of Sanctions Measurement Model	222
Table 65 Goodness of Fit Statistics for Predictive Model of Recidivism Based on Assignment of Apology Letters and Essays for Entire Data Set.....	223
Table 66 Chi-Square Tests for the Distribution of Specific Sanction Assignment Across the Categories of Reoffended for the Entire Data Set	223
Table 67 Chi-Square Tests for the Distribution of Specific Sanction Assignment Across the Categories of Reoffended for Older Offenders.....	224
Table 68 Chi-Square Tests for the Distribution of Specific Sanction Assignment Across the Categories of Reoffended for Older Female Offenders	224
Table 69 Chi-Square Tests for the Distribution of Specific Sanction Assignment Across the Categories of Reoffended for Older Male Offenders	225
Table 70 Chi-Square Tests for the Distribution of Specific Sanction Assignment Across the Categories of Reoffended for Drug Offenders.....	225

LIST OF ACRONYMS/ABREVIATIONS

AGFI	Adjusted goodness of fit index
CFA	Confirmatory Factor Analysis
CFI	Comparative fit index
CITI	Collaborative Institutional Training Initiative
DUI	Driving under the influence
DWI	Driving while intoxicated
FBI	Federal Bureau of Investigation
FDOC	Florida Department of Corrections
FLDJJ	Florida Department of Juvenile Justice
Freq	Frequency
GFI	Goodness of fit index
JAM	Juvenile Arrest and Monitor
NCIC	National Criminal Information Center
NRJP	Neighborhood Restorative Justice Program
IRB	Internal Review Board
NRJP	Neighborhood restorative justice program
RJ	Restorative Justice
RMSEA	Root mean square error of approximation
SEM	Structural Equation Modeling
USB	Universal Serial Bus
VOM	Victim offender mediation

CHAPTER 1: STUDY PROBLEM

Introduction

An increasing number of juveniles who are charged with crimes are not sent through the juvenile court system. Many offenders, especially those who have committed minor offenses or are first-time offenders, are diverted to a wide variety of pre-trial court diversion programs. In the Ninth Judicial Circuit of Florida, these programs have included Teen Court, Drug Court, the Neighborhood Restorative Justice Program (NRJP), and the Juvenile Civil Citation Program, but programs of varying specialty and characteristics exist across the United States. A common characteristic of these programs is non-traditional punishments for youthful offenders.

Sanctions in such programs can include any combination of community service, boot camps, apology letters, essays, curfews, random drug testing, substance abuse or mental health treatment, school progress reports, and anything deemed appropriate for the offender and offense. Many programs use civilian volunteers (sometimes juveniles themselves) to set sanctions for youthful offenders who have committed criminal acts. In an analysis of 19 victim offender mediation programs, Nugent, Williams, and Umbreit (2004) found about 80% used trained volunteers. The sanctioning bodies assign punishment packages based on volunteers' experiences and hunches and not on proven research as to the effectiveness of sanction packages for particular types of offenders. It is problematic that volunteers are making decisions on sanctions intuitively, without the benefit of research to guide them.

Diversion programs have become an invaluable tool for the courts to dispose of minor cases or specific needs cases (e.g., drug offenders) without burdening the formal court system.

According to the Florida Department of Juvenile Justice 2008/2009 Delinquency Report, of the 85,527 juvenile cases referred to the court in that year statewide, 37,783 (44%) were disposed of by using diversion programs, compared to 40,356 (47%) disposed of by the court. The remaining 9% were disposed of by mental health referrals or were sent to adult court. In Orange County, FL, of the 6,156 youth cases referred to the courts during the same time period, 2,353 (38%) were diverted from the courts and 2,784 (45%) were disposed of by the courts. The remaining 17% were disposed of by transfer to mental health services or to adult court (Florida Department of Juvenile Justice [FLDJJ], 2010). These numbers speak to the significance of this issue, which will be examined in detail in this study.

Criminal Justice literature abounds in studies of recidivism based on offender risk characteristics (see Figure 1 in Appendix B). These studies seek to identify the groups at the greatest risk of reoffending, however recidivism is defined. Conventional wisdom based on these studies suggests young, male, drug-based offenders from poor family environments who have records of deviancy are most likely to reoffend.

The literature also abounds with research on the relationship between Offender Risk Propensity and Multiplicity of Sanctions including, but not limited to, sanction severity (see Figure 2 in Appendix B). This body of research seeks to determine the relationship between offender demographic characteristics and the sanctions they are likely to receive or have received. The conclusions drawn by these studies suggest that older, minority, and violent male offenders with repeated violations receive the most and more severe sanctions.

Little has been done to study the effectiveness of certain punishment combinations on specific offender typologies, especially offender groups proven to be at a higher risk of

reoffending. To examine that question, this study tests three conceptual models: Main Effect, Mediating Effect, and Moderating Effect (see Figures 3–5 in Appendix B). These types of effects are discussed next.

“In the classic terms, a moderator is a qualitative (e.g., sex, race, class) or quantitative (e.g., level of reward) variable that affects the direction and or strength of the variable between an independent or predictor variable and the dependent or criterion variable” (Baron & Kenny, 1986, p. 1174). The moderator model has three paths to the outcome variable: a) from a predictor (Offender Risk Propensity) to Recidivism, b) from a moderator (Multiplicity of Sanctions) to Recidivism, and c) the predictor times the moderator to Recidivism. This model is different from how a mediator variable functions.

In general, a given variable may be said to function as a mediator to the extent that it accounts for the relationship between the predictor and the criterion. . . . Where as moderator variables specify when certain effects will hold, mediators speak to how or why such effects occur. (Baron & Kenny, 1986, p. 1176)

A mediator, as the name implies, functions between the independent variable and the outcome variable. Like the moderator model, the mediator model had three paths: a) from the independent variable (Offender Risk Propensity) to the outcome variable (Recidivism), b) from the independent variable (Offender Risk Propensity) to the mediator (Multiplicity of Sanctions), and c) from the mediator (Multiplicity of Sanctions) to the outcome variable (Recidivism).

In each model, Recidivism is not an indicator, but rather a latent construct. Rather than conceptualizing recidivism by one variable at a time (e.g., did the offender reoffend?), this research developed a measurement model for recidivism that simultaneously encompasses multiple methods of measuring recidivism. The resulting measurement model for Recidivism is explained in detail later. Structural equation modeling makes such a model possible.

The relationship between Offender Risk Propensity and Multiplicity of Sanctions and the relationship between Offender Risk Propensity and Recidivism are strongly supported in the literature. The model developed here has the benefit of testing the relationship between Offender Risk Propensity and Recidivism while controlling for Multiplicity of Sanctions, while at the same time testing the relationship between Multiplicity of Sanctions while controlling for Offender Risk Propensity.

In the mediating effect model (in Figure 4 in Appendix B), the Multiplicity of Sanctions is tested as a possible mediator between the Offender Risk Propensity and Recidivism. It has been well established in the literature, as discussed in detail in the next chapter, that Offender Risk Propensity, the initial variable, is correlated with recidivism, the outcome. This model suggests Multiplicity of Sanctions influences Recidivism as an intervening variable, based on the Offender Risk Propensity (Baron & Kenny, 1986). Figure 5 in Appendix B illustrates the moderating effect and suggests a combination of Offender Risk Propensity and Multiplicity of Sanctions impacts Recidivism.

This chapter identifies the study problem, explaining its significance and the theoretical focus. Second, the research framework, causal paths, theoretically informed perspective, and research questions are presented. Next the conceptualization of variables, constructs, hypotheses and the analytical tools used (as well as how the tools will be used and interpreted) will be elaborated. This will include the development of measurement models for Recidivism, Offender Risk Propensity, and Multiplicity of Sanctions and a multi-layer structural equation model (SEM) where the relationships between these latent constructs, as measured by their indicators, will be tested. Then research design issues (like unit of analysis, data sources, and sample size)

and anticipated results based on the literature will be discussed. Lastly, limitations of any findings and policy implications will be enumerated. Understanding effective sanctions could improve the juvenile justice system.

Research Problem

Keeping juveniles from entering or becoming more deeply involved in the juvenile justice system should be a goal of educators, parents, law enforcement, the courts, correction agencies, and the community. Crimes committed by the youth damage the community and jeopardize the youths' future and even their personal safety. Resources must be best allocated to address this public policy problem, and doing so requires understanding how to measure the effectiveness of diversion programs and the sanctions they assign.

The public policy issue addressed by this study is the lack of evidence-based research to guide pre-trial juvenile court diversion programs in the sanctioning of offenders. There is limited understanding and application of sanctioning effectiveness based on characteristics of the offender. Public policies should require the evaluation of sanctioning effectiveness to identify the sanctions or combinations of sanctions that are most effective in reducing recidivism for certain types of offenders. It is equally important that these results be communicated to and implemented by those who make sanctioning decisions.

In this context, an evidenced-based approach means a scientific and systematic examination of well-defined (valid and reliable) measures of the indicators of latent constructs (e.g., Multiplicity of Sanctions and Recidivism) as they can be shown to affect each other. Such an approach is theoretically informed, builds on available literature, tests falsifiable hypotheses,

and can be replicated using conventional means. The verifiable information obtained builds upon a body of knowledge and seeks to improve the human condition.

Nomenclature

“The beginning of wisdom is the definition of terms” (Socrates, 470 BC–399 BC). This section defines key terms used in this study. First, important but often confused distinctions must be made: those among efficiency, effectiveness, efficacy, and cost effectiveness, which are sometimes used interchangeably in the literature, and thus obscure important issues.

The term effectiveness “indicates the degree to which the objective has been achieved” (Carter, 1989, p. 98), “[W]here effectiveness focuses exclusively on results, the second criterion—efficiency—balances results and costs. The most efficient may not be the most effective. Rather it is the one that achieves the optimum balance of benefits and costs” (Salamon, 2002, p. 23). Efficacy can refer to the benefits gained in an ideal experiment using a randomized trial procedure. Cost effectiveness indicates a less experimental situation in practical settings without the benefit of randomization and where the costs are considered.

This study used historical data, the make-up of the sanction packages was not randomly assigned, and no control group was used. Further, because the cost (direct, indirect, or opportunity) of assigning and implementing the sanction package was not considered, the term effectiveness is used here throughout to mean simply the ability of a sanction package to reduce recidivism.

The term recidivism (“r”) refers here to a generic rate of reoffending or is a term as defined in a specific study. The term Recidivism (“R”) refers to the latent construct that is a function of multiple indicators contributing simultaneously to the understanding of that

construct. Latent constructs, like Offender Risk Propensity, are distinguished by capital initial letters. These constructs, their measurement, and testing are discussed in detail in later chapters.

The last nomenclature distinction to be made refers to first offenders. In this study, the term “first offender” is used to denote an individual who has been charged with a first criminal act and not necessarily one who has merely committed a first criminal act.

Significance of Research Problem

“Stated most simply, public policy is the sum of government activities, whether pursued directly or through agents, as those activities have an influence on the lives of citizens” (Peters, 2004, p. 4). The significance of a public policy issue is considered as a function of scope, prevalence, trends, costs, seriousness, incidence, and prevalence rates (Bardach, 2000). This section examines the significance of the public policy problem of sanction effectiveness and Recidivism measurement in juvenile diversion programs. These issues are significant for three main reasons: the costs of juvenile crimes on society, the frequency of youth arrests, and the rapid growth and use of youth diversion programs in the United States. Each reason is examined in this section.

Policy and research issues often revolve around costs, and juvenile crime is expensive. The Florida Department of Juvenile Justice (FLDJJ) spent \$633.6 million in FY 2002–2003 and 642.8 and 619.2 in 2008–2009 and 2009–2010, respectively (FLDJJ, 2009, 2010). These costs did not include the costs of law enforcement or costs for the victims of these crimes, which would increase the total costs dramatically. The Florida Department of Juvenile Justice claimed that a one percent reduction in juvenile crime saves an estimated \$10.3 million in criminal justice and \$5.3 million in costs to victims (FLDJJ, 2003). Such high costs cannot be ignored.

In addition to considering the direct and indirect costs, policymakers must consider opportunity costs, the cost of funding one program or approach at the expense of another. Restorative justice has become a rallying flag for such divergent activities as “justifying more dollars for diverting first-time offenders from the court” (Umbreit, Coates, & Vos, 2007, p. 24). Corrections administrators feel compelled to use community-based diversion programs, not because they are more effective at reducing recidivism but because they are cheaper (Martin, 2003). These programs and their activities should be examined in comparison to other programs and approaches that can serve the same ends.

Policy and research issues also arise as a function of the number of people affected, and again, crime affects many people. According to the United States Department of Justice, in 2008, Florida was sixth in the nation at the rate at which juveniles were arrested for property-related crimes (Puzzanchera, 2009). The author does warn that “state variations in juvenile arrest rates may reflect differences in juvenile law-violating behavior, police behavior, and/or community standards; therefore, comparisons should be made with caution” (p.11). According to the Florida Department of Law Enforcement [FDLE] in 2002, 123,270 juveniles were arrested in Florida. In Orange County, Florida, more than two-hundred juvenile cases are referred to the courts in an average week (Orange County Public Schools SAFE Program, 2004). Youth arrests are too common to ignore.

Lastly, public policy and research issues arise when issues not previously publicly discussed come to public light. In 2001, there were just over 300 restorative justice programs (NRJPs) in the United States (Nugent et al., 2004; Prison Fellowship International, 2001). In 2002, there were more than 875 such programs nationwide, an increase of almost 300% in one

year (Dick, Pence, Jones, & Geetresen, 2004). There are currently 33 NRJPs in Florida, which receive about \$550,000 in grant funds; 13 counties offer civilian volunteers trained by the Florida Department of Juvenile Justice (FLDJJ, 2003). In Europe, there are currently more than 1,000 such programs (Nugent et al., 2004). Further, the number of teen courts, peer courts, drug courts, and a variety of other diversion programs that follow a similar process is growing rapidly. Law enforcement agencies are increasingly involved with the operation of these programs (FLDJJ, 2003). With so many first-time juvenile offenders sent to court diversion programs, a firm understanding of the effectiveness of the punishment combinations they employ is crucial. The literature, however, has yet to address that concern, as is discussed in the next section.

Weaknesses in Current Research

The current research in this field has been limited to three distinct areas: offender recidivism, sanction assignment, and program reviews. Traditional recidivism studies have examined characteristics of the offender, the crime, and the environment to identify groups at the greatest risk of reoffending (see Figure 6 in Appendix B). These studies have focused on such offender characteristics as antisocial behavior, gender, criminal history, family environment, mental health, substance abuse, intelligence, school performance, social skills, and age to determine their association with recidivism (Mulder, Brand, Bullens, & Van, 2010).

A second body of literature has focused on offender characteristics and their association with the sanctions' characteristics, including severity. These studies look at arrest, prosecution, and punishment of offenders and the relationships between these events and offender and crime characteristics (Applegate, Turner, Sanborn, Latessa, & Moon, 2000; Bishop & Fraizer, 1996; Worden & Myers, 1999).

Studies have also been conducted on the effectiveness of certain types of punishments. For juveniles, these studies have traditionally focused on the role of boot camps, curfews, and community service in reducing recidivism (Bouffard & Muftic, 2007; Martin, 2003; Steiner & Giacomazzi, 2007). Applegate and colleagues (2000) reported, for example, that

judges focus primarily on offense characteristics, and are influenced only marginally by the offender's social characteristics. These findings are more consistent with the view that juvenile courts are becoming "criminalized" than with the view that individualized treatment is the goal. (p. 309)

Lastly, studies have examined such overall programs as juvenile restorative justice for their effectiveness at reducing recidivism (Rodriguez, 2007). The problem with program reviews as a method of determining sanction effectiveness is that programs could have used any combination of sanctions, so conclusions about sanctioning cannot be drawn from those studies. These studies and their findings are discussed in more detail in the literature review chapter.

The drawback in these studies that makes them inapplicable in practice for those actually assigning sanctions to juvenile offenders is that offender recidivism and sanction effectiveness have been examined separately. Hypothetically, a study may show that letters of apology do not serve as effective sanctions overall, but they may be effective for white female offenders aged 15–18 charged with a minor property-related offense such as misdemeanor retail theft. Drug testing may not be an effective sanction overall, but one would believe intuitively that it would be an effective sanction for drug offenders. The causal path examined by this study is the effectiveness of certain sanctions or sanction combinations on certain offender typologies (see Figure 7 in Appendix B).

Specific Research Questions

The research questions, simply stated, are as follows: How should we measure recidivism; does completion of juvenile restorative justice diversion reduce recidivism; and does sanctioning matter, and if so, what is the best combination of sanctions to reduce reoffending for certain offender typologies?

CHAPTER 2: LITERATURE REVIEW

Before research is started, a review of the available knowledge, theories, and findings should be completed. This chapter presents the theoretically informed perspective for the study, to be accomplished in many steps, each building on the last. First, the specific theories to be used here are explained and applied to the current study. Second, punishment philosophies, in particular those applicable to this study, are presented. Next, the construct of Recidivism is discussed, including its current conceptualizations and measurements as well as the problems these techniques present. After the concept of restorative justice is introduced, defined, and explained, the conceptualizations of recidivism are applied to studies of restorative justice from outside as well as inside the United States, including meta-analyses. Lastly, the literature related to Offender Risk Propensity and Multiplicity of Sanctions is explored with a focus on recidivism. This chapter provides the framework for the hypotheses and methodologies, which are discussed in the next chapter.

Theoretical Focus and Informed Perspective

A research study should have a theoretically informed framework that builds on the published science and literature. A theory is a “set of interrelated constructs, definition and propositions that present a systematic view of phenomena by specifying the relationship between variables with the purpose of explaining and predicting a phenomena” [sic] (Kerlinger, 1986, p. 9). Theory sets the foundation for the way a problem or process is conceptualized. The theoretical focus of this study, or the guiding principle under examination, is the presumption that all punishment combinations are not equally beneficial in reducing recidivism (the rate at

which juveniles reoffend) for all offender typologies, and that those who assign sanctions should seek to understand and apply what has been proven to be most effective. Practitioners who fail to use theory, or even an evidence-based approach “continue to be guided by commonsense, intuitive, or idiosyncratic notions for reducing recidivism” (Dick et al., 2004, p. 1450).

Contingency Theory

The discussion of theory begins with Contingency theory. Contingency theory suggests that the best sanction package for a youth is contingent on a variety of factors and not all sanctions would be equally effective under all situations (Lawrence & Lorsch, 1967). Levine (2009), in applying contingency theory to violent offending and reoffending, paraphrased Steadman et al. (2000):

The contingency framework provides a broad basis on which to consider key risk factors widely thought to relate to violence. The framework also acknowledges that the predictors of violence [and for that matter reoffending in general] may vary for different people in different situations. Also the framework explicitly considers that complex contingencies between risk factors occur. (pp. 167–168)

Contingency theory, although applicable, is not the most accepted theory in academic circles.

System Theory

The theory most directly applicable to this study is Systems theory. This theory as described by Kraska (2004) suggests an open/rational system approach, adaptation and awareness of external forces (and their impact), streamlining operations, and a focus on rational decision making, effectiveness, and reduced discretion. This theoretically informed framework can be further strengthened by an understanding of punishment philosophies, which are discussed next.

Punishment Philosophies and Related Assumptions

Before any theoretically informed framework can be applied to sanctioning, a basic understanding of punishment philosophies is required. The punishment philosophy sets the stage for the application of public policy, for research actions, and ultimately for the choice of sanctions. According to Von Hirsch (1976) as cited in Worden and Myers (1999), there are four main punishment philosophies: rehabilitation, retribution, deterrence, and incapacitation. This section discusses those four and a fifth, restorative.

Rehabilitation is a punishment philosophy that attempts to make a person less likely to reoffend by changing his or her values, habits, and skills through training, counseling, drug treatment, and education. Retribution is a philosophy that suggests the offender owes a debt to society and should receive punishments based only on the crime committed. Proponents of the deterrence philosophy believe rational people are motivated not to commit crimes out of fear of certain (or at least likely) punishment. This philosophy has two aims, specific (to prevent a specific person from offending) and general (to set an example to society). The incapacitation philosophy professes that the best way to deal with criminals is to separate them from society, their potential victims (Lab, Williams, Holcomb, King, & Buerger, 2003, pp. 5–6). The restorative philosophy, which is explored later in this paper, envisions “a process whereby parties with a stake in a specific offense collectively resolve how to deal with the aftermath of the offense and its implications for the future” (Marshall, 1999, p. 5). In many cases components of these theories are used in conjunction.

For the purpose of this research, two important assumptions are made. First, it is assumed that for juvenile offenders (and especially those sent to pre-trial diversion programs) the

underlying punishment theories are rehabilitation and specific deterrence. The second assumption is that the offenses disposed of by juvenile diversion programs are relatively minor offenses, such as status offenses (e.g., possession of alcohol by a minor) and misdemeanor crimes such as criminal mischief (graffiti) and shoplifting.

These assumptions are commonly understood in the juvenile justice literature. Lundman, Sykes, and Clark (1978) agreed with Black and Reiss (1970) that a majority of police-juvenile encounters involve very minor (if any) criminal behavior. Worden and Myers (1999) were in agreement, with the conclusion that most juvenile police encounters involved non-serious issues such as loitering or rowdiness, and “less than one tenth of the encounters concerned violent crimes” (p. 15). These distinctions are significant to the understanding of juvenile crime, its sanctions, and the theories that apply, particularly as they are applied in restorative justice programs. Three theories specific to restorative justice (Labeling, Differential Association, and Reintegrative Shaming) are discussed later in this chapter. The point of sanctioning, most would agree, is to prevent the youth from reoffending. The concept of recidivism is discussed after the brief summary of this theoretical focus section that follows here.

Application of Literature and Theory on the Current Study

The theories presented above start to develop the theoretically informed perspective for the current study. Contingency and Systems Theories were introduced, as well as the punishment philosophies of rehabilitation, retribution, deterrence, incapacitation, and restorative. This section provides a summary of their applications to the current study.

The Open System Theory suggests that organizations and individuals are affected by their environments (Scott, 2003). Youths who interact with the criminal justice system are affected by

the encounter. Sanctions are part of that encounter. All sanctions are not expected to be equally beneficial to all individuals in all settings; this is the application of Contingency Theory (Lawrence & Lorsch, 1967). The Open/Rational System Theory urges adaptation and awareness of external forces (and their impact), streamlining operations, and a focus on rational decision making, effectiveness, and reduced discretion (Kraska, 2004). Specific theories related to sanctioning and restorative justice (Reintegrative Shaming Theory, Labeling Theory, Differential Association Theory, and Social Capital) are explained and applied later in this chapter.

Sanctioning philosophies of rehabilitation, retribution, deterrence, incapacitation and a restorative approach have been discussed above (Von Hirsch, 1976; Zehr, 2002). For the purpose of this study the punishment philosophy combines rehabilitation, specific deterrence, and restorative. Restorative as a punishment philosophy is discussed in more detail in the restorative justice programs section below.

Youths are changed by their environment as they go through the criminal justice system, whether the formal system or a diversion program. All sanctions are not expected to be equally beneficial at reducing recidivism. The literature on recidivism is discussed next.

Conceptualizing Recidivism

Recidivism, like any latent construct, can be measured in countless ways. Different definitions of a reoffense (police contact, parole or technical violation, arrest, prosecution, court contact, and conviction), different time periods (0–20 years), different data sources (self-reported, city, county, state, and national databases), and different reporting methods (offenses per 100,000, percentage of reoffending, and time to live) all make valid comparability of studies and program reviews virtually impossible. This section discusses recidivism in terms of

definitions, time periods, data sources, and severity. Understanding a study's methods of calculating recidivism rates can be more important than its results.

States' Definitions of Recidivism

How is recidivism defined? There is little agreement in the academic literature, and even states' reviews of their own programs present comparability problems. The twelve-month reviews of recidivism for juveniles released from incarceration in different states vary wildly because of the differences in how recidivism is conceptualized. Even studies from the same state have different criteria. Florida, for example, has studies that report recidivism four different ways.

Public officials and policymakers might be influenced, for example, to revamp New York's system (which reports a very poor 55% recidivism rate) to be more like that of Arkansas (which boasts an amazing 12% recidivism rate) (Harris, Lockwood, & Mengers, 2009). In fact Harris and colleagues reviewed 46 recidivism studies and could not find two that were comparable. They all differed in terms of definitions of recidivism, data sources, and time periods reviewed.

Florida, New York, and Virginia, which have studies which consider recidivism as only re-arrest in either the juvenile or adult systems, report recidivism rates as high as 55%. Colorado and Maryland conceptualize recidivism as referral to adult or juvenile courts, so not just arrest, but some level of prosecution is required. They report 45% recidivism. Arkansas, Florida, Georgia, Kentucky, Maryland, North Dakota, Oklahoma, and Virginia have program reviews that use reconviction/adjudication in adult or juvenile court systems as an indicator of recidivism and report 33% recidivism. Florida, Maryland, and Virginia (the states with the widest variety of

recidivism conceptualizations) also reported 24% recidivism when it was defined as not only a criminal act but a criminal act that resulted in referral to adult or juvenile court, prosecution, and adjudication and re-incarceration. Lastly, Arizona, Montana, and New Mexico have state programs that report an amazingly low 12% recidivism. Those studies defined recidivism as a new criminal act resulting in prosecution, adjudication/conviction, and re-incarceration only in the juvenile court system: participants who committed any criminal violation as adults, regardless of the outcome, were not considered to have reoffended (Harris et al., 2009). One might expect that evaluators of state programs would use standardized recidivism definitions. This assumption would be wrong. Researchers' measures are discussed next.

Researcher's Definitions of Recidivism

In research, recidivism is defined even less consistently than in state program reviews. Hayes and Daly (2003), for example, who studied reoffending in restorative justice and conferencing programs, defined recidivism as:

any new official incident (which might have involved multiple charges or counts) to which the police responded to with arrest or apprehension after the date of the ... conference. These incidents were dealt with by formal caution, conference, or court. ... All violent offenses, property offenses, and driving offenses (normally drunk driving offenses) were included, and breaches of good-behavior bonds which were related to previous sentences (or undertaking), were excluded because they did not reflect a fresh incident. ... The postconference [sic] window of time was thus 8 to 12 months. (p. 741)

Although they were very specific as to what constituted a reoffense, the authors' data source was not specified. Data sources are discussed in detail later in this chapter.

Nugent and colleagues (2004) in their meta-analysis examined 15 studies of 19 different victim offender programs. They found three definitions of recidivism. The first was a charge resulting in an adjudication of guilty (Nugent & Paddock, 1996; Umbreit & Coates, 1993). The

second was a charge not resulting in dismissal or exoneration (Schneider, 1990). Most studies defined recidivism as any official contact with law enforcement, such as re-arrest, but in some states official contact could include traffic citations, field interview reports, and trespass warnings.

The call for standardized conceptualizations of recidivism is not new. In 1976, the National Advisory Committee on Criminal Justice Standards and Goals under the United States Department of Justice released this statement: “A major problem in research on criminal justice is the absence of standardized definitions. . . . The confusion over definitions has not only impeded communication among researchers and practitioners, but also hindered comparisons and replications of research studies” (Harris et al., 2009, p. 6). The present study proposes a measurement model that considers many of the conceptualizations of recidivism all at the same time: what act constitutes a reoffense, the timing of the act (or time to live), the severity of the offense, and the outcome of the new offense. These components are discussed next.

Reoffending

What act constitutes a reoffense? Recidivism measures all begin with some action, which indicates further deviant behavior, and again there are wide differences in terms of what that means. This section looks at the different definitions used by researchers.

Researchers have used definitions of recidivism as minor as a re-sanctioning for violations of their particular program (Guerra & Slaby, 1990; Leeman, Gibbs, & Fuller, 1993; Martin, 2003; Peters, Thomas, & Zamberlan, 1997; Wiebush, 1993). Thus recidivism could be as simple as missing curfew or a failed drug test while in a program. Obviously, with definitions this strict, programs will report higher recidivism rates, all other things held constant.

An apparent majority of the studies have considered re-arrest to be either the sole measure of recidivism or some component of recidivism (Borduin et al., 1995; Fagan, 1990; Fredrick & Roy, 2003; Harris et al., 2009; Herzfeld et al., 2008; just to name a few). It should be noted that these studies had different data sources. Using only county or state records is likely to produce lower recidivism rates than does using national records. Although the effect is likely to be minimized in juvenile populations because they are less migratory than adult populations, it still should be considered.

The next conceptualization of recidivism is as an act resulting in a court referral (Bank, Marlowe, Reid, Patterson, & Weinrott, 1991; Borduin, Henggeler, Blaske, & Stein, 1990; Leeman et al., 1993; Smith & Monastersky, 1986; to name a few). For these authors, some form of prosecutorial response is required if an action is to be considered a reoffense.

Some studies have further restricted the construct of recidivism to include only felony court filings (Roskey, Pasini-Hill, Lowden, Harrison, & English, 2004; Tarte, Mackin, Cox, & Furrer, 2007). Levine (2009), for example, defined recidivism as regarding only specific serious charges: “murder, robbery, assault, burglary, sex offenses, weapons, and other offenses such as kidnapping” (p. 169). The more specific and obscure the definition of recidivism is, the less comparability the studies using it will have with others.

A more restrictive conceptualization of recidivism is re-conviction (Fredrick & Roy, 2003; Hagan & Cho, 1996; Nugent & Paddock, 1996; Peters et al., 1997; Schneider, 1990; Umbreit & Coates, 1993; & Wiebush, 1993). These studies often use multiple offense types when measuring recidivism. Martin (2003), for example, in a study of community-based

sanctions in Oregon, conceptualized recidivism as “reconviction of a felony, re-sanctioning for a violation and re-arrest for any violation of law” (p. 26).

Lastly, some studies have used commitment to define recidivism. Thus, to count as a new offense the action must result in arrest, prosecution/referral, adjudication, and re-incarceration. Of those studies, some have even specified juvenile or adult commitment (Harris et al., 2009). Studies with this definition, all other things held constant, are likely to report the lowest recidivism rates.

Harris and colleagues (2009), in a review of 53 recidivism studies, noted the actions used to define recidivism: technical violation (17%), petition (60.4%), physical arrest (22.6%), petition or physical arrest (69.8%), adjudication (30.2%), juvenile commitment (3.8%), and adult commitment (1.9%). Over 43% of the studies reviewed used multiple measures at the same time to define an act that constituted reoffending. Whatever the researchers choose for the specific indicator of reoffending, once it is set the researcher must choose the time period to be studied and a method of reporting the findings. Time periods for study are discussed next.

Simply considering recidivism a function of a dummy variable such as whether or not the offender was charged with an additional crime (yes/no) during a specific time period

is wasteful of information because it ignores the variation on either side of the cutoff point. One might suspect, for example, that a person arrested immediately after release has a higher propensity towards recidivism than one arrested 11 months later. (Allison, 1982, p. 64)

Some researchers, such as Levine (2009), discussed later, have attempted to address this issue by making the dependent variable in their analysis time to live. This approach encounters new problems, known as censorship. Censorship, which relates to the unknown characteristic of

the dependent variable if no re-arrest occurred during the examination period, is discussed in detail later in this study.

Time Period Examined

Even when they agree about the act that constitutes a reoffense (and the section above shows they do not), the studies in the literature have an enormous range and lack of agreement in terms of the time periods examined. Although the commonly used time periods appear to be 0.5, 1 and two years, the range is 0–20 years. Some review periods, like 2.04 years (Bottcher & Ezell, 2005) and 1.75 to 4.08 years (Borduin et al., 1990) immediately raise doubts about their comparability with studies with more commonly used time periods. In a review of 22 recidivism studies, 25 different time periods were used (see Table 1 in Appendix A).

In their review of 53 recidivism studies, Harris et al. (2009) noted the time periods examined. They reported their percentages as less than one year (20.8%), 1 year (34%), 1.5 years (7.5%), 2 years (24.5%), three years (5.7%), more than three years (18.9%); moreover 28.3% of these studies reported multiple follow-up periods. Harris et al. reported the average maximum follow-up time period as 2.6 years. Although studies show the recidivism occurs sooner after correctional services (Levine, 2009), all things held constant, the longer the time period reviewed, the greater the recidivism rate is likely to be. Looking at time to reoffending, Levine (2009) examined a 20-year period following incarceration. During that time period, he found a recidivism rate of 66%. This is especially high, particularly because he defined reoffending as only specific serious acts. The next variable in recidivism studies to be considered is the source of the data used.

Data Sources

Even if researchers agreed completely on the definitions of re-offense and on the time period to examine, issues of comparability can arise because of the data sources for recidivism. Sources for reoffending data include a wide range, from self-reported to national databases.

Some studies have used records from their respective departments of parole/probation (Guerra & Slaby, 1990; Leeman et al., 1993; Peters et al., 1997; & Wiebush, 1993). Most studies, particularly of juvenile recidivism, and most program reviews have used arrest records from state databases (Fagan, 1990; Fredrick & Roy, 2003; Herzfield et al., 2008; Hagan & Cho, 1996; Nugent & Paddock 1996; Peters et al., 1997; Schneider, 1990; Umbreit & Coates, 1993). Some studies have used national databases, such as the National Criminal Information Center (NCIC) in the United States (Borduin et al., 1995; Wiebush, 1993) or national databases of Australia (Hayes & Daly, 2003). Other studies have used county court records as data sources (Bank et al, 1991; Borduin et al., 1990; Fredrick & Roy, 2003; Leeman et al., 1993; Roskey et al., 2004; Smith & Monastersky, 1986; Tarte et al., 2007). All things held constant, the wider the data source, the higher the reported recidivism rates to be expected.

The next two sections focus on measures not regularly used in recidivism research: frequency and severity of reoffending, and case outcome.

Reoffending Severity and Frequency

Preventing offenders from reoffending is assumed to be the goal of correctional programs. So, by extension, preventing offenders from reoffending with an offense more severe than their original one and preventing them from reoffending with any frequency should also be a goal. Studies on recidivism have failed to capture measures that reflect those goals. Most

recidivism studies report on how many individuals reoffended as a function of yes/no. Thus, an individual who reoffended once is considered to be the same as an individual who reoffended thirty times. These measures fail to capture a true picture of recidivism.

The studies also do not compare the original charge's severity to the severity of any new violation. For example, Youth A's original charge was assault (a second-degree misdemeanor in Florida); he or she completed a correctional program and immediately reoffended with a capital offense (murder for example). That set of circumstances would be considered simply a yes for recidivism. Youth B, whose original crime was burglary (a third-degree felony in Florida), completed a correctional program and reoffended by committing assault 15 years later. Youth A reoffended right away with a crime far more severe than the original charge. Youth B reoffended 15 years later with a crime less severe than the original crime charged. As long as the new criminal acts were both in the defined time period and the acts were considered a re-offense, these two hypothetical individuals would both simply score a "yes" for reoffended. This measure fails to provide a complete conceptualization of recidivism.

Case Outcome as a Measure of Recidivism

Some studies have conceptualized recidivism as incarceration (Harris et al., 2009), but no reviewed study looked at the case outcome as a continuum of recidivism. If arrest was the definition of offense, it does not matter to researchers how that case was disposed of (exoneration, nolle prosequi (dropped by the state), judicial warning, adjudication of guilty, probation, or incarceration). Case outcome can bring into consideration the strength of the evidence, seriousness of the violation, and past criminal record of the violator. Some also might

argue that case outcome is a function of demographics such as race. These issues are beyond the scope of this study.

Reporting Recidivism

Once all of the factors above are defined and their relationship studied, researchers report their findings. The literature includes three general ways of reporting recidivism. This section discusses these methods. They include recidivism rates, time to live, and offenses per number of offenders per a set time period.

Most studies report recidivism as a percentage of previous offenders who take new criminal actions, however defined. Levine (2009), for example, examined a population of 4,146 young males committed to the California Youth Authority in 1964 and 1965. A random sample ($n = 511$) was drawn, which was reduced to 413 after cases with missing data were eliminated. An official search for arrest and conviction records over a 20-year period revealed that 273 (66%) had reoffended seriously. Seriously was defined as these specific charges: “murder, robbery, assault, burglary, sex offenses, weapons, and other offenses such as kidnapping” (p. 169). The percentages of individuals with new actions that constituted a reoffense were reported as the measure of recidivism. Reporting either the recidivism rate or the success rate (1 - recidivism rate) is by far the most common reporting method.

A second recidivism reporting method is time to live, some measure of the time between program completion and a new criminal act. Levine (2009), for example, did a tree analysis of groups based on the timing of their recidivism. The median time to recidivism was as low as 0.5 years for non-Caucasian males who had a diagnosed mental illness and IQ less than 96.3. The

longest median survival time was ten years for Caucasians with fewer than two prior arrests (0 or 1), who had married parents, and who had been arrested after the age of 18.

The third method of reporting recidivism observed in studies was the number of offenses per a certain group of offenders per a certain time period. Sherman, Strang, and Woods (2000), for example, studied an Australian restorative justice youth diversion program and reported that, when compared to those whose case was disposed by the court, those originally charged with violent crimes who participated in the diversion program had a “big drop in offending rates by violent offenders (by 38 crimes per 100 per year) [and a] very small increase in offending by drunk drivers (by 6 crimes per 100 offenders per year)” (p. 3). Although this measure is more confusing than a simple percentage rate, it does account for multiple offenses committed by the same offender and in certain cases can provide a fuller conceptualization of recidivism.

Recidivism, like any latent construct, can be conceptualized in many different ways. Different specifications of what act constitutes a re-offense, different time periods examined, considering the timing and severity of violations, different data sources, and different reporting methods all contribute to variations in the figures reported. At best, the outcomes of these varying decisions make comparing studies and treatment methods difficult. At worst, the variations can be chosen intentionally to make a particular program look better so more funding may be secured (and perhaps even diverted from more effective programs). Either way, standardization is absent and badly needed in recidivism literature. Now that recidivism has been examined, it can be applied specifically to studies of restorative justice programs and Offender Risk Propensity, in the next two sections.

Application of the Recidivism Literature and Theory to the Current Study

A review of the recidivism literature reveals a few blatant problems. First and foremost, studies use different recidivism definitions. Different acts that constitute a reoffense, different time periods, different data sources, and different reporting methods all present researchers and policymakers with comparability problems between studies and between interventions. Second, most studies only use one indicator of recidivism (such as re-arrest). Few studies use multiple indicators, but almost none use multiple indicators at the same time. That is also the case for offender characteristics and for sanctions, issues addressed later in this study. Unless several indicators of recidivism are used at the same time, the picture of recidivism is incomplete.

This current study offered and tested a measurement model of the latent construct of Recidivism (see Figure 8 in Appendix B). Indicators (e.g., the time from services to reoffending, offense severity, change in severity level, case disposition, number of new charges, and re-offense (yes/no)) were tested to determine their contribution to the measurement of the latent construct of recidivism. The belief was that these indicators used simultaneously would present a better conceptualization of recidivism than would simply the use of one indicator (such as re-arrest, time to live, or new crime severity) one at a time. This measurement model of recidivism was expected to more clearly reveal the relationship between the latent constructs of Recidivism and Offender Risk Propensity and Sanctions Characteristics of youths in restorative justice programs. These programs are discussed next.

Restorative Justice Programs

Before examining recidivism in restorative justice diversion programs, the restorative justice concept should be understood. This section first defines the concept as explained by

researchers and by practitioners using it. Then the components and requirements of such programs are discussed. Next, the theories thought to explain the impact of these programs are introduced: Restorative Shaming (also known as Reintegrative Shaming), Labeling, and Differential Association. The remainder of this section delves into recidivism research to examine national, international, and meta-analysis studies, their methodologies, and findings.

Restorative justice programs are for the most part diversionary. The concept of a diversion program is a bit of a contradiction.

Many if not all of those in the ranks of the police, the judiciary, probation departments, and departments of correction are there because they wanted to help those who had been hurt by crime and bring about change in individual behavior. (Umbreit et al., 2007, p. 24)

And it is those members who administer and staff diversion programs. Thus, youth who go through diversion programs may still interact with police officers, prosecutors, court officials, judges, and even correctional and probation officers. The path their cases take through the criminal justice system, however, is often very different from the traditional process, especially in restorative justice programs, which are discussed next.

Restorative Justice Defined

“While still not the mainstream in any nation, restorative justice has clearly moved beyond the margins of social change in many locations and is beginning to enter the mainstream criminal justice policy” (Umbreit et al., 2007, p. 23). Restorative justice programs have become popular all over the globe. Australia, New Zealand, United States, Canada, England, Norway, Scotland, and Japan all have such programs (Latimer, Dowden, & Muise, 2005). There has been a great deal of confusion and debate about exactly what constitutes restorative justice (Doolin, 2007). The term has been used as synonymous with peacemaking criminology, transformative

justice, community justice, relational justice, and victim offender mediation (Bazemore & Walgrave, 1999). Programs using restorative justice principles have been called sentencing circles, neighborhood accountability boards, and juvenile restorative justice programs (Umbreit, Coates, & Vos, 2001). This section attempts to define restorative justice through a review of the leading authors in the field and identify the elements of the definition upon which they all agree.

In the literature, Marshall's (1999) definition appears most frequently. He defined restorative justice as "a process whereby parties with a stake in a specific offense collectively resolve how to deal with the aftermath of the offense and its implications for the future" (p. 5). Some debate exists on exactly how to deal with the aftermath and who exactly are the stakeholders (Doolin, 2007). The authors do agree on the overall goals of these programs: 1) to repair the harm caused by delinquent acts, 2) to involve and empower victims, and 3) to use community members to reintegrate offenders into the community (Bazemore & Walgrave, 1999; Braithwaite, 1989; De Beus & Rodriguez, 2007; Doolin, 2007; Hayes & Daly, 2003; Rodriguez, 2007; Umbreit et al., 2007; Wheeldon, 2009; and Zehr, 2002). All of these authors accept the doctrine that "crime is a violation, violation creates an obligation, and reparation fulfills obligation" (De Beus & Rodriguez, 2007, p. 2).

Restorative justice programs invite all parties affected by the crime to discuss its effects on them with the offender. Then the offender and the group agree on a contract, which includes activities the offender must do to restore the community to its pre-crime state. The term *invited* is used because a majority of victims, especially victim businesses, generally do not participate in these programs. The programs are almost always unique in their operations, structures, and the offenders they serve. These characteristics are explored next.

Restorative Justice Characteristics

Just as the definitions of restorative justice vary, so do the program characteristics. There are some commonalities in characteristics. Almost all restorative justice programs require the offender to admit to his or her offense and take responsibility for it (Doolin, 2007). Some require the victim to agree to the program. The programs generally seek to get the offenders to become aware of the harm they have caused and make efforts to repair that harm (Zehr, 2002).

These programs are almost always voluntary. Juvenile offenders and their parents are offered the opportunity to participate if they agree to the guidelines of the specific program, if one is available. If they decline to participate, refuse to admit to the original offense, or are not interested in repairing the harm caused by their actions, their cases are sent back to the court for standard prosecution. Other requirements may also exist.

The crime must meet the requirements of the particular restorative justice program. Most programs usually accept only minor offenses (Levrant, Cullen, Fulton, & Wozniak, 1999), but some New Zealand programs do accept felony cases (Hayes & Daly, 2003). Some programs accept only offenders who live in or who have committed the crime in a certain neighborhood or jurisdiction. The Ninth Judicial Circuit's NRJP program (the data source for this study), according to its charter, accepts only offenders with first-time non-violent misdemeanors who reside in specific areas of Orange County, Florida. But minor violent crimes, some felonies, and a few for which the offenders reside outside the service area have been accepted into the program.

There have been faith-based restorative justice programs, such as the Faith-Based Neighborhood Accountability Board Project, which was run through the African Methodist

Episcopal Churches in Orange, Alachua, Pinellas, Dade, and Duval counties in Florida. This non-denominational program no longer exists today, according to Remera Jones, Director of the Bethel by the Lake Youth Community (Remera Jones, personal communication, October 20, 2004).

The time period for completion of such youth restorative justice programs is generally about 90 days or three months. Once the contract is agreed upon, the youth has 90 days to complete the assigned sanctions, treatments, and other stipulations such as drug testing and curfews (if assigned). The time period allows for intense supervision, which is discussed below in the Multiplicity of Sanctions section.

Theories Related to Restorative Justice

Two classical sociological theories are specific to restorative justice programs: Labeling and Differential Association. A newer theory is Reintegrative Shaming. This section presents these theories, starting with the Labeling Theory.

Labeling Theory

The labeling theory proposes that certain aspects of the criminal justice system produce more delinquency than they prevent. As youths continue to encounter agents of the criminal justice system, “they begin to see themselves as delinquents [and] develop and [that serves to] reinforce delinquent self-concepts ...” (Dick et al., 2004, p. 1451). “The stigma attached to contact with justice agencies may act as a self-fulfilling prophesy in which the individual’s behavior conforms to the label ...” (p. 1452). An application of labeling theory to diversion programs suggests that any success of a program could be attributed to a less formal, less

confrontational, and less intrusive program. Such program characteristics may avoid the negative label and thus future delinquency based on that label. Further, the practice in restorative justice of denouncing the act but not the actor may help to avoid internal labeling.

Differential Association

The next theory applicable to restorative justice programs is differential association. The differential association theory professes that “if an individual is exposed to an excess of definitions favorable to law-violating behavior, then there is an increased likelihood that said individual will engage in delinquent activity” and vice versa (Dick et al., 2004, p. 1453). Following this theory, diversion programs seek to frequently and intensely express negative reactions to the delinquent act (but not the offender) to prevent the youth from internalizing antisocial definitions.

Reintegrative Shaming Theory

Proponents of restorative justice would argue that it is not the sanctions but the program’s process and philosophy that reduce recidivism. This section discusses the Reintegrative Shaming Theory. Altogether the three theories provide the theoretical background to explain the impact of these programs.

Reintegrative Shaming Theory, one of the first theoretical frameworks for restorative justice, was developed by Braithwaite (1989). Braithwaite observed that sanctions imposed by members of the offender’s community were more effective than those set by members of the criminal justice system, who were seen as anonymous. According to Hipple and McGarrell (2008):

Reintegrative shaming involves a conscious effort to shame the action of the offender but not the offender as a person. Disapproval is expressed with regards to the action but the person who committed the act is not labeled as deviant or evil. A level of respect is maintained towards the offender. The offender is then given a chance to rejoin the community of law-abiding citizens. Individuals who are closest to an offender are best able to instill a sense of Reintegrative shaming. Essentially, shaming of the offender occurs while maintaining the bonds of respect. (p. 557)

The youth's actions but not the youths are condemned, while allowing the offender to make up for the action. Then, after restoration, the youth is welcomed back into the community. Whether the sanctions assigned or the process of the system deters future violations, the question arises: do these programs work?

Restorative Justice and Recidivism

Perhaps the most important measure of the success of any criminal justice intervention is recidivism, however it is defined. This section examines the relationship between participation in restorative justice programs and recidivism. Restorative justice programs have received mixed reviews. Hayes and Daly (2003) stated: "The literature on restorative justice and reoffending consists largely of comparative analyses of traditional and restorative interventions and suggests small but significant differences or no differences in reoffending" (p. 725). "Restorative justice studies have not always found restorative justice programs reduce recidivism" (De Beus & Rodriguez, 2007). Having noted that, it can be further noted that very few studies have found participation in restorative justice programs to have increased recidivism, especially in any statistically significant way (Nugent et al., 2004)—unlike studies of boot camps (Blair, 2000). Sherman and colleagues (2000) reported that for offenders originally charged with driving while intoxicated recidivism rates for program participants as compared to those sent to court were

actually higher “by 6 crimes per 100 offenders per year” (p. 3). This section examines the literature on restorative justice programs and recidivism.

Howard Zehr directed the first restorative justice program in the United States and is considered the grandfather of the concept. On the subject of recidivism, Zehr (2002) stated:

There are good reasons to believe that, in fact, such programs will reduce offending. Indeed, the research thus far—centering mainly on juvenile offenders—is quite encouraging on this issue. Nevertheless, reducing recidivism is not the reason for operating restorative justice programs. Reducing recidivism is a byproduct, but restorative justice is done first of all because it is the right thing to do. (p. 9-10)

This study does not consider victim satisfaction and involvement, offender and community involvement in the process, or the reduced amount of conflict in the process. All of these are the goals of restorative justice and are supported by the literature (Hayes, 2005; Sherman et al., 2000; Umbreit et al., 2007; and Zehr, 2002). This study focuses exclusively on recidivism as a measure of program success.

The East Point, Georgia, Police Department implemented a diversion program in 1996 and in 2001 reported that not one of its more than 150 youth participants had reoffended. They further reported an increase in school performance and respect for law enforcement officers. Moreover, most of the participants “bonded with officers ... [and chose to] stay involved after the four weeks” (Georgia ‘juvenile diversions’, 2001, p. 1). This program reported extraordinary results. Some would say that it succeeds because it builds social capital, or the “connections among individuals—social networks and the norms of reciprocity and trustworthiness that arise from them” (Putnam, 2000, p. 19).

The East Point study, admittedly published in a practitioner publication, not an academic journal, illustrates most of the criticisms of the recidivism literature. The reports are anecdotal;

recidivism is not always defined (and when it is the definitions and time periods differ); there is no (or poor) experimental design and no control group; participants are cherry-picked; there is selection bias; and generalizability is extremely questionable.

The City of East Point enacted a city ordinance that allows police officers to arrest truant youths, those not in schools during the day. After three violations, the parents are fined. In order to drop the charge against the youth, the child must participate in a four-week, police-run program. There, youths spend the weekend with officers learning about local history and the Constitution, undergo counseling, and participate in afterschool activities, cultural events, and athletic sports alongside senior police officers. This police-run program also accepts first-time, non-violent misdemeanor offenders.

In 2002, the City of East Point had an estimated population of 82,243 with 30,577 households (2.7 persons per household). The median household income \$43,071, and 45.1% of the households had household income over \$50,000. The workforce was 51% white collar (East Point, 2011). In addition to the study's not defining recidivism, not having a control group, and cherry-picking the participants, the question of its generalizability arises. This section examines other studies on restorative justice and recidivism and presents findings from foreign and domestic studies, starting with foreign studies.

Australian/New Zealand Studies

Considerable research has been conducted on restorative justice and conferencing in Australia and New Zealand, since both these countries have been utilizing restorative justice longer (Hayes & Daly, 2003, p. 726). Sherman and colleagues (2000) compared program participants to a control group who went through the formal court process (n = 1,300). The

program was the Canberra Reintegrative Shaming Experiment (RISE). They had intended to use self-reported criminal history questionnaires, but they reported recidivism based only on official records provided by the Australian Federal Police. They randomly assigned youths to either a restorative justice program or the control group, which went through the normal court process. Their time period was two years. They reported that violent offenders who participated in the program reoffended less (38 fewer offenses per year per 100 offenders). They found no statistically significant difference in property offenders, and they reported a slight increase in recidivism for offenders charged with drunk driving, although “the effect size of that difference was a mere 0.1, which is modest indeed” (p. 13).

Australian research also examined what in these programs contributed to the success or failure of participants. In a study funded by the Australian Research Council, Hayes and Daly (2003) attempted to identify bivariate relationships between selected variables and the rate of reoffending in restorative justice and conferencing programs. Certain factors were clearly identified as predictors of recidivism. The offender treating the police officer with disrespect and the coordinator not permitting all parties to have a say in the process were both seen in 100% of those who later reoffended. Offenders’ belief that the process was a waste of time was seen in 78% of those who reoffended. They found that those who had three or more addresses on file with the police accounted for 72% of those who had reoffended. Offenders who did not know their victims personally and offenders who victimized organizations accounted for 50% and 56%, respectively, of those who reoffended.

The characteristics found in those who reoffended less included treating police officers with respect, outcome decided by genuine consensus, ending the conference on a high note,

offender not being defiant and accepting responsibility for his or her actions, offender providing a clear and honest account of the incident, and assurances by those who personally knew the victims and offenders that the offense would not happen again. Surprisingly, the offender crying during conferencing was actually seen more often in those who subsequently reoffended (Hayes & Daly, 2003). Factors which appear to be correlated to recidivism will be discussed in more detail in the Offender Risk Propensity section. Next, restorative justice studies from the United States will be examined.

U.S. Studies

Many U.S. studies have concluded that restorative justice program participants had lower recidivism compared to control groups who went through the court system. Studies of programs in Pennsylvania, Arizona, and Tennessee all concluded that restorative justice program participants had lower recidivism than control groups sent through youth courts.

The Tennessee Juvenile Victim Offender Mediation (VOM) Program had significantly lower recidivism than a comparison group (Umbreit et al., 2001). This was echoed by Rodriguez (2005) in an empirical study using official court records of youth who participated in a restorative justice program in urban Arizona. McCold and Wachtel (1998) studied a Pennsylvania restorative justice program and reported that choosing to participate in a restorative justice program was found to have a significant impact on reducing recidivism. This brings up the issue of selection bias, which is addressed later. These are only three studies of specific programs, but the results of two meta-analysis studies, which covered 54 separate programs, echoed their results. The meta-analysis studies are discussed next.

Meta-Analysis

According to Glass, McGaw, and Smith (1981) as cited in Latimer and colleagues (2005), “A meta-analysis can be understood as a statistical analysis of a collection of studies that aggregate the magnitude of a relationship between two or more variables” (p. 130). Meta-analysis reviews are generally regarded as a method of research synthesis that is superior to traditional narrative reviews, being “more systematic, more explicit, more exhaustive and more quantitative (Rosenthal, 1991, p. 17, as cited in Latimer et al., 2005, p. 130). In laymen’s terms a meta-analysis is a study of studies or a synopsis of many studies.

In one such meta-analysis, Nugent and colleagues (2004) examined 15 studies of 19 different victim offender programs, which provided services to over 9,000 juvenile offenders. Their study focused only on juvenile offenders and concluded “the odds of VOM participants [reoffending] were only about 0.7 as great as the odds of nonparticipants reoffending” (p. 408). In addition to different definitions of recidivism, these studies had different degrees of restorative justice. The Schneider (1990) study compared restitution (only part of which might be considered restorative justice). She reported that for every 100 youths who completed restitution programs, there was a decrease in 18 offenses per year. None of the studies examined in this meta-analysis “provided any evidence that [victim offender mediation] participation was associated with an increase in delinquent behavior” (Nugent et al., 2004, p. 415).

Latimer and colleagues (2005), in another meta-analysis of 22 studies on 35 individual programs, concluded that “restorative justice programs, on average, yield reductions in recidivism compared to non-restorative approaches to criminal behavior” (p. 137). During

follow-up, as compared to a control group those who participated in restorative justice programs were significantly less likely to reoffend (Latimer et al., 2005).

In the 59 studies on restorative justice and recidivism, only one found an increase in recidivism (and only for a specific offense: driving under the influence (DUI)). The vast majority concluded that youths in these programs had lower recidivism rates than the control groups. One possible explanation for these results is selection bias, which is discussed next.

Selection Bias

Selection bias (or selection threat) occurs when participants are not randomly assigned to either the treatment or control groups. How subjects might be selected, or in these cases how the subjects make selections themselves, may affect the outcome of the study. Selection bias can threaten the quality of the findings (Gliner & Morgan, 2000).

These programs are almost always voluntary, which lends itself to criticisms of selection bias. Does the fact that a child (or his or her parents) selected participation in a restorative justice program alone influence recidivism? Does the fact that some member of the criminal justice system (police officer, prosecutor, judge, or correctional officer) selected or recommended diversion influence recidivism?

McCold and Wachtel (1998) studied 232 misdemeanor juvenile offenders who were eligible to participate in the Bethlehem, Pennsylvania, Restorative Justice Program. They randomly assigned the youths into two groups: control (103, 68 property/35 violent) and treatment (189, 113 property/76 violent). The treatment group was then broken into two groups based on participation choice: those who had been selected for conferencing and who

participated (80, 56 property /24 violent), and those who had been selected for conferencing and who refused to participate (109, 57 property /52 violent).

Although violent offenders were more than twice as likely as other offenders to refuse voluntary participation, they had lower recidivism rates (defined solely as re-arrest) than did the violent offenders who had refused to participate. The study found no statistically significant difference in recidivism between property offenders who participated and those who had been offered the program and refused to participate. This may be because of several factors. Juvenile offenders charged with property-related crimes generally reoffend at a lower rate than those charged with violent crimes, which in the small sample may have made the variation in recidivism between the groups too low to be considered significant (McCold & Wachtel, 1998).

On the topic of selection bias, the recidivism rates of the three groups were 20% for those who participated, 35% for those not offered participation, and 48% for those offered and refusing participation. Overall recidivism was lowest for those who participated in the program as compared to those who did not. Especially noticeable is that the recidivism rate observed more than doubled if participation was offered and declined. “Results suggest that recidivism was more a function of offender choice to participate than the effects of conferencing” (McCold & Wachtel, 1998, p. 4).

Facilitator Background

One variable in restorative as well as other diversion programs is the background of the facilitator. The major distinction in facilitator type is the difference between law enforcement (police) and civilian facilitators. In a study of 215 conferences, Hipple and McGarrell (2008) found the officers more likely to lecture youth offenders, but they found “no major differences

between conferences facilitated by civilians as opposed to police officers” (p. 553). They also reported that “youths who attended police-officer–facilitated conferences survived somewhat longer before reoffending than did youths who attended civilian-facilitated conferences, although these differences were not statistically significant” (p. 553).

Programs have also been run by other than police officers and civilian volunteers. Programs in the United States have been run by court employees and probation officers (De Beus & Rodriguez, 2007) and also by clergy. The demographics of the program coordinator are not nearly as important in terms of a correlation with reoffending as are the demographics of the offender. These demographics, the Offender Risk Propensity, are discussed after the application of the material above to the current study.

Application of Restorative Justice Literature and Theory on the Current Study

This section presents literature on academic findings and theories related to the relationship between participation in a restorative justice program and recidivism. A theoretical background for these findings is also presented. This section summarizes these findings and theories.

Participation in restorative justice programs as opposed to the formal court system has been found in most cases to reduce recidivism (Hayes & Daly, 2003; Latimer et al., 2005; McCold & Wachtel, 1998; Nugent et al., 2004; Rodriguez, 2005; Schneider, 1990; Sherman et al., 2000; Umbreit et al., 2001). A different impact on recidivism was seen for violent offenders, property offenders, and drug offenders. Several theories explaining this relationship are discussed.

Property offenders and those who had victimized businesses were not seen to have as dramatic a decrease in recidivism as was observed in violent offenders (Hayes & Daly, 2003), but some studies did observe reductions in recidivism for these groups after program completion (Umbreit et al., 2001). Violent offenders were more than twice as likely to refuse voluntary participation (McCold & Wachtel, 1998). But violent offenders who chose to participate had lower recidivism than did violent offenders who refused to participate and those not selected (Hayes & Daly, 2003; McCold & Wachtel, 1998; Sherman et al., 2000). One researcher found DUI offenders actually to have higher recidivism after participation as compared to a control group (Sherman et al., 2000). Ahlin, Zandor, Rauch, Howard, and Duncan (2011) studied this group of offenders and reported high rates of recidivism regardless of the sanctions imposed. Since DUI offenders make up a very small percentage of offenses by juveniles, this finding is not a material concern for the current study.

Although McCold and Wachtel (1998) suggested that the relationship between participation in restorative justice programs and recidivism is a function mostly of selection bias, several theories discussed in this section offer other explanations. Labeling Theory, Differential Association Theory, and Reintegrative Shaming Theory all explain why participants in restorative justice programs are observed to have lower recidivism rates than do offenders who did not participate (Dick et al., 2004).

Labeling Theory explains the relationship between self-identification as a delinquent and committing delinquent acts. The theory suggests as youths interact with the formal criminal justice system, they come to perceive themselves as delinquents and are likely to commit more delinquent acts (Dick et al., 2004). Application of this theory suggests that the less interaction a

youth has with any formal criminal justice system, the less self-identification as a delinquent occurs, and by extension the less recidivism will be observed.

Differential Association Theory explains the relationship between a youth's acquiring excessive self-definitions favorable to delinquency and the youth's engaging in delinquent acts (Dick et al., 2004). The application of this theory suggests that the process of denouncing the act without denouncing the actor reduces the likelihood of the offender reoffending. This theory may also explain the effectiveness of certain sanctions that do not reinforce excessive definitions favorable to delinquency, which is discussed later.

Reintegrative Shaming Theory explains the relationship between observed recidivism and sanctions imposed by community members (not a process). It also explains the relationship between recidivism and the "welcoming back" into the community after sanction completion (Dick et al., 2004). This "welcoming back" recognizes and builds on Social Capital, the trust that comes from social networks in a community (Putnam, 2000).

To apply these findings to the current study, the relationship of the indicator (completion of the restorative justice program) to the latent construct (Recidivism) was tested (see Figure 9 in Appendix B). This was accomplished using the measurement model described in the recidivism section above. It was expected from the theories and literature that variation in the construct (Recidivism) would be explained by variation in the indicator (completion). This is the relationship that was tested.

Offender Risk Propensity

Researchers have done a great deal of research on the demographics of offenders and reoffenders to answer questions such as who is likely offend, who is likely to reoffend, what

instruments/characteristics best predict recidivism, and, recently, what is the profile for certain types of crimes? This arena of research has generated heated controversy. If black males are arrested more often than other races for a certain offense, does that mean that a black male is more likely to commit that offense, or does it mean that black males are targeted by enforcement or are given less leniency? This section examines the relationship between two latent constructs, Recidivism and Offender Risk Propensity.

Many studies have created and tested complicated prediction models for juvenile recidivism. In a review of three such models for predicting juvenile recidivism (the Contra Costa County Model (Baird, 1982), the Orange County, California, risk assessment instrument, and the Arizona Juvenile Risk Assessment Form), Ashford and LeCroy (1990) reported that although the three models varied widely in their ability to predict juvenile recidivism, certain factors contributed to their accuracy. Age at which the first offense was committed was the strongest predictor, with younger offenders more likely to reoffend. “Prior referrals, prior parole violations, run-away behavior, offense type, school, peer associations, alcohol or drug abuse, and family dynamics” were the remaining eight characteristics used by the Arizona form, which was shown to be the most accurate model (Ashford & LeCroy, 1990).

Criminal Justice researchers are not the only ones studying the relationship of recidivism and demographics. Mental health professionals have also delved into this field but with far less predictive success. According to Lidz, Mulvey, and Gardner (1993) as cited in Levine (2009), “Early research has shown that mental health professionals are modestly more accurate than chance in predicting violence” (p. 166). This section discusses the findings of criminal justice recidivism studies, starting with those on criminal history.

Prior Arrest History

One strong predictor of people's future behavior is their past behavior. This section examines the relationship between prior record (a component of Offender Risk Propensity) and recidivism. Researchers have well established past criminal history as a predictor of future criminal activity. In some cases, the probability of reoffending was doubled for those with past arrest records, or the time to live of those with a past arrest record was less than half of those without prior arrests. Levine (2009), in a 20-year recidivism study of 413 young male offenders sentenced to a commitment program, found arrest history to be a significant predictor of time to live. For example, among non-Caucasians with married parents, he found that those with one or fewer prior arrests had a median time to live of 3.5 years, compared to 6 years for those with more than one prior arrest. For Caucasians, the results were even stronger; those who had two or fewer arrests had a mean time to live of 2.2 years, compared to 8.3 years for those with more than two arrests.

The Levine study presents a lack of comparability between the conclusions about Caucasians and those about non-Caucasians. For non-Caucasians, the prior arrest demarcation was one prior arrest, as opposed to two arrests for Caucasians. Further, for non-Caucasians, "median" was reported, as opposed to "mean" for Caucasians. Although the results for Caucasians and non-Caucasians are not comparable, Levine's findings (2009) do demonstrate the impact of prior arrests on the average time to live, regardless of the type of average used.

In the current study, almost all offenders were first-time offenders, so past arrest record may not seem relevant. The study considered school suspension as a predictor of recidivism. The link between school performance and recidivism has been well established (Rankin, 1980;

Simpson & Elis, 1995; Tobin & Sugai, 1999). Many of the causes for suspension could have resulted in criminal charges (e.g., theft, assault, battery, disorderly conduct, or disrupting a school function). Suspension history as a record of past criminal behavior was tested in this study as a predictor of recidivism. School discipline records as a predictor of recidivism is discussed in more detail in the school performance section below.

Gender

The next variable in Offender Risk Propensity to be discussed is gender. This section examines the relationship between gender and recidivism. The observed and anticipated efficacy of specific sanctions based on gender is discussed as well.

Females are arrested less often than males (FLDJJ, 2003), and in restorative justice diversion programs females traditionally reoffend less frequently than males do (Hayes & Daly, 2003). Females are arrested more often for running away (a status offense in most states) or for minor crimes like shoplifting. If females are arrested less frequently and reoffend less frequently, one might ask, why even study female offenders? In a program that focuses on serious offenses, males account for a majority of participants; however, when a program focuses on less serious crimes, females participate at a substantial rate (Elis, 2005). Females account for about 40% of the participation in the Ninth Judicial Circuit Court's Neighborhood Restorative Justice Program (NRJP), the program that is the data source for the current study.

Much of the juvenile justice literature on gender differences focuses on gender bias in arrest, prosecution, and punishment (Bishop & Frazier, 1992; Horowitz & Pottieger, 1991). The history of the juvenile justice system starts with females more often being prosecuted for status offenses and crimes related to "immorality and waywardness" and often receiving a penalty

more severe than males received for the same act (Elis, 2005, p. 376). According to Chesney-Lind and Shelden (2004) as cited in Elis (2005), such sanctioning practices were seen as the institutionalization of gender stereotypes to make women good wives and men good providers. The deinstitutionalization of status offenses by the Juvenile Justice and Delinquency Act of 1974 may have reduced the frequency of differential treatment, but sanctions specific to gender must be applied cautiously to avoid bias towards gender stereotypes and role institutionalization.

Sanctioning by gender begins with a better understanding of possible differences in values and communication by gender. Females, for the most part, place a greater value on connections with others, whereas males tend to place a greater value on autonomy, accomplishment, and independence (Gilligan, 1982 as cited in Elis, 2005). Sanctions that build social bonds may be “more compatible with the value orientation of women” and thus may be more effective in reducing recidivism among them (Elis, 2005, p. 375).

The Relational/Instrumental Theory as discussed by Hagan, Simpson, & Gillis (1988) builds on the theoretical framework for the findings discovered by Gilligan (1982) and Elis (2005). The theory suggests that males develop into more instrumental thinkers and communicators. Actions and communications among males are intended mostly for specific purposes: developing autonomy, independence, and a position in a hierarchy of other males. Females, in contrast, become more relational. Communication and thinking processes in females are more often intended to express emotions and build connections with others.

No-contact orders, forbidding the offender to associate with a codefendant or particular friends, are likely (based on the studies discussed above) to be a more effective sanction for males than for females. “Even in the presence of delinquent friends, girls are less likely to

engage in behavior if they believe it to be immoral. The presence of delinquent peers increases boy's involvement in delinquency" (Elis, 2005, p. 384). In laymen's terms, boys are more likely to do stupid things while surrounded by other boys doing stupid things. Furthermore, since males place more value on individualism and autonomy than females are likely to, curfews may in general be a more effective sanction for males than for females. These hypotheses, based on these theories, were tested later in this study.

Pressure placed on males by their friends to engage in risk-taking behavior may increase their delinquency and recidivism (Giordano, Cernkovich, & Pugh, 1986). This suggests that isolating males from delinquent friends may reduce recidivism. Females were found to socialize with less delinquent friends (Morash, 1986), less often encourage risk-taking behavior, and be less susceptible to encouragement from their friends to engage in criminal behavior (Mears, Ploeger, & Warr, 1989; Simpson & Elis, 1995). These assertions were tested against the sample of NRJP participants. Gender differences as they relate to poor school performance and recidivism are discussed further in the next section.

The literature on gender differences suggests that no-contact orders would be more effective for males than for females. Levine (2009), in contradiction to this expectation, found that the presence of a crime partner quadrupled the time to live for Caucasian males with two or more prior arrests. Possible explanations for this finding are discussed in a later section.

School Performance

Researchers have attempted to predict delinquency based on school performance, specifically academic and disciplinary records. In a longitudinal study, Tobin and Sugai (1999)

followed a sample (n = 526) of students who were in the sixth grade between the 1989 and 1992 school years, for six years. They concluded:

Referrals for violence involving fighting in graded 6 predicted similar referrals in grade 8. Frequency of discipline referrals in grade 6 predicted chronic discipline problems in later middle school, which predicted frequency of suspensions in grade 9. Three or more suspensions in ninth grade predicted school failure. Boys referred for fighting more than twice and girls referred even once for harassing, as sixth graders, were not likely to be on track for graduating high school. (p. 40)

The literature supports contentions that disruptive behavior, leading to poor school performance and delinquency, can be predicted as a function of poor social and behavioral skills (Boulden, 2010). Ashford and LeCroy (1990), in reviewing the Arizona Juvenile Risk Assessment Form, also observed that school referrals were a strong predictor of recidivism. This component is further complicated by gender differences.

Attitudes towards and performance in school were found by Rankin (1980) to be correlated with delinquency, but more strongly for females than for males. Simpson and Elis (1995) agreed with Rankin. Both found that poor performance in school was more strongly correlated with property crimes for females than for males. However, poor performance in school and violent crimes were associated for both males and females.

These authors found that placement in remedial math programs had the opposite effect. Females in remedial math programs were seen as more likely to offend violently, whereas males were seen as more likely to commit property-related crimes. This might suggest that school progress reports (or other efforts to encourage positive performance in schools) would be most effective for females charged with property crimes and for females in remedial math programs, but for males with poor school performance charged with any crime. However, school progress

reports are nevertheless seen to be an effective tool for any youth not performing well in school (Simpson & Elis, 1995). Age as a demographic linked to recidivism is discussed next.

Age

Researchers have studied the relationship between the age of the offender and recidivism. Age at first arrest has been shown to be an important factor in predicting recidivism, with younger offenders reoffending at a higher rate. Levine (2009) in a study of 413 young male offenders sentenced to a commitment program in California, found age significant ($p < 0.001$) for whites with fewer than two arrests who were not diagnosed with mental illness. In that group, those arrested while under 18 years of age ($n = 56$) had a median survival time of 5.5 years. Those over 18 years of age at the time of arrest ($n = 112$) had a median survival time of 10 years (almost double of those under 18 years of age). The longest median survival time was ten years for Caucasians with fewer than two prior arrests (0 or 1) who had married parents and were arrested after the age of 18 (Levine, 2009). Those arrested after 18 are no longer juveniles, but the point is no less important. Ashford and LeCroy (1990) examined three predictive models and concluded that the age at which the first offense was committed was the strongest predictor, with younger offenders more likely to reoffend.

Race

No demographic related to recidivism has received more public attention than race. Studies on the association with race seek to understand the relationship between race and recidivism and between race and the severity of the sanctions assigned. Studies have shown these relationships to be significant.

The relationship between race and recidivism is discussed first. Levine (2009), in a 20-year recidivism study of 413 young male offenders sentenced to a commitment program, observed “an increased risk of time to recidivism was significantly associated with not being Caucasian or Mexican [i.e., being black], ... not having married parents, ... more prior arrests, ... and having past hospitalizations ...” (p. 171). He reported, for example, that the median reoffending time for a non-Caucasian with an IQ less than 96 was 0.5 years (n = 12). The highest median time to live was 6 years for non-Caucasians with married parents and no prior arrests (n = 15). For members of that same group with prior arrests, the median time to live dropped to 3.5 years. For Caucasians, the longest time to live was 10 years for those with two or fewer prior arrests and married parents (n = 112). For Caucasians the shortest time to live was one year for those with more than two prior arrests and no crime partners (n = 15). Levine also reported that for some offender types, the presence of a crime partner was actually a negative predictor for recidivism. A discussion of criminal partners and the no-contact order is found in the sanction section below.

Next, the relationship between race and Multiplicity of Sanctions assigned is discussed. Wordes, Bynum, and Corley (1994) examined the decision to detain juveniles as it is made at three different decision points: officer arrest, court intake detention, and preliminary hearing detention. They sought to determine whether the dependent variable of detention rates was associated with the independent variable of race. They tested their empirical theory by studying felony cases from five counties in the same state, which may bring up issues of generalizability. They concluded that “African Americans and Latinos were more likely to be detained at each

decision point, even after controlling for offense seriousness and social factors” (p. 149). Further, blacks were

more likely to be younger, have more prior offenses on record, be charged with a drug offense, and be charged with offenses that had more serious injuries and in which a person was pregnant ... and detained independent of offense seriousness. (p. 156)

Race alone was not the only factor related to detention; so were a record of drug violations and a criminal record, both variables seen more frequently in black offenders (Wordes et al., 1994).

Other factors, such as the type of crime committed, also have been studied in relation to recidivism. The crime type is discussed next.

Crime Type

Aside from demographic differences, the actual crime committed by the youth has been studied to determine its impact on recidivism. This section explores those studies, especially those of youth offenders in diversion programs. It reviews recidivism of drug offenders (in this study DUI offenders are considered a subset of drug offenders), violent offenders, property offenders, and status offenders. Drug offenders are discussed first.

Drug Offenders

Drug offenders also have been the focus of many academic studies. This section examines the relationship between drug offenders and recidivism, a relationship that is tested later.

Ashford and LcCroy (1990), in a review of the Arizona Juvenile Risk Assessment Form, studied the model’s ability to predict juvenile recidivism. Alcohol and drug abuse was among eight characteristics determined to most accurately predict recidivism. Those who used drugs and

alcohol had a statistically higher probability of reoffending. This is no surprise to anyone who has reviewed drug recidivism studies. What may be surprising are the findings of Kim, Benson, Rasmussen, and Zuehlke (1993). They concluded that recidivism among drug offenders is influenced by incentives and constraints. Their study is discussed next.

Police and policymakers often claim that drug users somehow differ from the rational criminal presented in models that attempt to apply sanctioning to criminal behavior. Kim and colleagues (1993) concluded, contrary to this commonly held notion, that recidivism among drug offenders is indeed influenced by incentives and constraints. The size of the offender's local police department and imprisonment versus probation were both seen as significant predictors of recidivism for drug violators. A larger police department may be seen by the potential offender as increasing the probability of apprehension, which Robinson and Darley (2004) found is a strong incentive for not engaging in criminal behavior. They concluded that concepts such as incentives (or disincentives like punishments) and utility can appropriately be applied to drug criminals.

The authors defined recidivism as a criminal act resulting in identification, arrest, and prosecution, rather than as simply a subsequent criminal act, thus greatly reducing their reported recidivism rate. Also, they defined drug violators and those who commit both drug violations and other crimes separately. They found drug violators who also commit other crimes to have a much higher rate of recidivism than do drug offenders only. This finding may be explained by increases in criminal activity increasing the probability of apprehension. Their study of over 45,000 arrests for drug possession in Florida found that 76% of the offenders had no prior felony

arrest, 80% had no charges involving theft, and 90% had no charges for property-related crimes (Kim et al., 1993).

Probably the most surprising thing discovered by this study was that a vast majority of drug offenders were not involved in other crimes like burglary, theft, and stolen property (and vehicles). In fact many had no other non-drug-related criminal history at all. The authors' conclusions completely contradict the commonly held assumption in criminal justice that drug offenders commit a majority of property-related crimes in order to fund their drug use. It is upon that belief that many "get tough on drug crime" and mandatory sentence guidelines are based. This research demonstrated that those policies should be re-evaluated. The studies reviewed here support the notion that although drug offenders may also need drug treatment, sanctioning may deter future criminal acts by drug offenders.

A subset of drug offenders is those charged with alcohol-related driving offenses. Although in some states the laws distinguish between driving while impaired and driving while intoxicated as shown by blood alcohol levels, these terms are used synonymously here to mean an alcohol-related driving charge.

Ahlin and colleagues (2011) compared those charged with driving while intoxicated (DWI) to other licensed drivers in Maryland who were charged with other crimes. They reported that the probability of reoffending (defined as a new charge for DWI) was relatively high for those charged with a prior DWI, regardless of how they were sanctioned. Those sent to diversion had statistically similar recidivism to those not sent to diversion. In addition to being shown to be at risk for recidivism, DWI offenders have also been shown to be less-than-ideal candidates for restorative justice programs.

Sherman and colleagues (2000) studied an Australian restorative justice youth diversion program. They reported that as compared to those sanctioned in court, those originally charged with drunk driving who entered the diversion program had slightly higher recidivism “by 6 crimes per 100 offenders per year” (p. 3). No other study could be found to confirm or dispute their findings. Although first-time DUI offenders might be juveniles, their frequency is not sufficient to make this group material for the current study. In fact, none of the offenders in the sample used for this study (n = 218) had an original charge related to alcohol and driving.

Violent Offenders

The literature on violent offenders in restorative justice programs is promising. Sherman and colleagues (2000) found violent offenders to be particularly suited for restorative justice programs. In a study of an Australian youth program, they reported that the largest reduction in recidivism was among violent offenders. McCold and Wachtel (1998), in their study of a Pennsylvania restorative justice program, reported that although violent offenders were very likely to refuse participation in the program, those who did participate and completed it had significantly lower recidivism than did those who did not participate. Hayes and Daly (2003) reported that offenders who did not know their victims personally and offenders who victimized organizations accounted for 50% and 56%, respectively, of those who reoffended.

Status Offenders

Status offenses are crimes that are crimes only because of the age of the offender. Curfew violations and possession of alcohol or tobacco by minors are examples of status offenses. Status offenders have been particularly suited for restorative justice programs. Beus and Rodriguez

(2007) studied completion and recidivism and concluded that status offenders who participated in the restorative justice program were more likely to complete the program and less likely to reoffend than similar offenders in a control group.

Property Offenders

De Beus and Rodriguez (2007) studied program completion and recidivism and concluded that property offenders who participated in the restorative justice program were more likely to complete the program and less likely to reoffend than were similar offenders in a control group. Other studies discussed above suggest participants in restorative justice programs who knew the victim personally had lower recidivism than those who did not; the offenders in cases where the victim was a business (e.g., retail theft) had higher recidivism than did violent offenders who knew their victim personally.

Other Characteristics

Many Offender Risk Propensity characteristics were beyond the scope of this research and data available. Emerging research suggests, for example, that certain types of anxiety are risk factors for offending and reoffending (Kubak & Salekin, 2009). Two types of mental health assessment instruments as predictors of offending violently found them to be only slightly more accurate than pure chance (Levine, 2009; Lidz et al., 1993).

Another relationship between Offender Risk Propensity and Recidivism and between Multiplicity of Sanctions and Offender Risk Propensity surrounds the construct of respect. Hayes and Daly (2003) observed that offenders who treated officers with respect during a restorative justice conference had lower recidivism than those who did not. They found a strong and

significant relationship between respect and recidivism. Another relationship they observed was that between respect and sanction severity. They observed that those who showed disrespect to the sanctioning body received more severe sanctions, even when they controlled for the severity of the crime.

An observation of Worden and Myers (1999) that would hold no surprise for law enforcement officers is the relationship between suspects' display of disrespect towards officers and the probability of arrest. The authors also observed that the disrespect was more likely to be shown by black offenders than by white offenders. In cases where there was sufficient evidence to arrest, disrespectful behavior towards the officer increased the probability of arrest from 21% to 42%. The authors suggested that this was caused by two factors: either officers punishing offenders who failed to respect their authority, or officers rewarding offenders who were respectful. There is a third explanation.

It is possible that those who displayed disrespectful treatment towards officers also, in doing so, committed another crime (such as obstruction or resisting without violence) by disobeying officers' commands. Worden and Myers (1999) observed that 8% of the offenders were not simply disrespectful but also refused to obey commands. That behavior could make some of the disrespectful actions also additional crimes, changing it from an extra-legal to a legal factor.

Application of Offender Risk Propensity Literature and Theory to the Current Study

In the review of findings and theories related to Offender Risk Propensity, several relationships between indicators (race, age, gender, school performance, drug history, and crime

type) and recidivism were discussed. This section summarizes those findings and theories and applies them to the current study.

Offenders with a past history of offending are seen in the literature to be at a greater risk of reoffending than were those without a history of offending (Levine, 2009). School performance, specifically as a target of disciplinary action, is also seen as a predictor of delinquency (Rankin, 1980; Simpson & Elis, 1995; Tobin & Sugai, 1999). Younger offenders are seen to have higher recidivism than older offenders (Levine, 2009), males are seen to reoffend more than females (Giordano et al., 1986), and non-Caucasians are seen to reoffend more (and be sanctioned more harshly) than Caucasians (Levine, 2009; Wordes et al., 1994). Some theories have been formulated to explain these observations.

Elis (2005) proposed that gender differences in delinquency exist because females place a greater value on connections with others, whereas males place a greater value on autonomy, accomplishment, and independence. He suggested that sanctions that build social bonds may be better suited for females and be more effective in reducing recidivism among females. That conclusion reflects Relational/Instrumental Theory (Hagan et al., 1988), which suggests that males develop to be more instrumental in their thinking, communications, and actions, whereas females develop to be more relational. The theory thus suggests why sanctions that restrict autonomy, such as curfews and no-contact orders, are expected to be more effective for males. Sanctions are discussed later in this paper.

The studies discussed above looked at specific Offender Risk Propensity indicators separately and not all at the same time, which can provide an incomplete picture. Younger offenders reoffend more than older offenders, but does that hold true, for example, for female

property offenders without a record of school discipline? The current study applied the literature and theories reviewed to develop and test a measurement model for Offender Risk Propensity. That model was used to test relationships between Offender Risk Propensity, Multiplicity of Sanctions, and Recidivism.

The indicators for the construct of Offender Risk Propensity included suspension history, gender, age, drug history, race, and crime type (see Figure 10 in Appendix B). The relationship between the constructs of Offender Risk Propensity and Recidivism were also tested (see Figure 11 in Appendix B).

Multiplicity of Sanctions

This final section of the literature review chapter focuses on research studies that have examined the relationship between specific sanctions assigned and recidivism. Juvenile diversion programs are often criticized for “procedures that are arbitrary, subjective and often used without clear justification” (Dick et al, 2004, p. 1459). Use of theory to guide research can avoid many of those criticisms and help researchers understand and improve the programs.

The sanctions examined are limited to those commonly used in youth diversion: boot camps, community service, curfews, letters of apology, restitution, drug testing, mental health counseling, anger management, and no-contact orders. According to Muikuvuori (2001):

It is a problematic task to make comparisons between different sanctions in terms of how they affect recidivism: it is difficult to find completely identical control groups, and the material available often creates problems as well. A person’s criminal background, age, sex, and the nature of his offence, etc. are considered to be significant factors from the point of view of recidivism. ... (p. 72-73)

Those concerns notwithstanding, a review of the sanctions listed above and their impact on recidivism is the focus of the final section of this chapter. This review begins with boot camps.

Boot Camps

Boot camps became very popular in the field of juvenile justice through the 1990s.

By 1997, more than 27,000 teenagers were passing through 54 camps in 34 states annually. These programs subjected offenders to an intense military style discipline and physical training for short periods of time, usually from one day to one week. (Blair, 2000, p. 4)

Overwhelmed by allegations of abuse and misconduct, Maryland, Colorado, Arizona, and North Dakota eliminated their boot camps entirely. Florida and California reduced their enrollment (and length of time), and Georgia revamped its program completely. The relationship shown between assignment/completion of a boot camp and recidivism was not promising and is not likely to be applicable to the boot camps assigned in current youth diversion programs.

Studies on boot camps have shown one year recidivism rates of 64% to 75%, which is higher than the rate for those sentenced to adult prisons (Blair, 2000). The boot camps in which youths in diversion programs are sent differ greatly from those studied by Blair (2000).

Diversion programs rarely use boot camps with intense military styles and almost none last longer than a weekend, although some lead the youth to think they will be there longer. The boot camp used by the Ninth Judicial Circuit of Florida's NRJP program, as of the time when data were collected for this study, was only a one-day Saturday program. Since then, the program has been expanded to a second day and has included classes for parents on setting boundaries and enforcing rules at home.

Those critical of boot camps, pointing out the higher recidivism rates of their participants, assert their failure. But as Muikuvuori (2001) pointed out, finding comparable groups is difficult. In youth diversion programs, boot camps are one of the most severe sanctions. This means that only the worst offenders or those viewed as in the most need of an attitude adjustment are

sanctioned to boot camps. Therefore, with boot camps the most severe sanction available, those in diversion programs who are sanctioned to boot camps would be expected to have a higher recidivism rate than those not sanctioned to boot camps, all other factors being held constant. This expectation does not reflect the efficacy of the sanction but rather either the selection bias of those making sanctioning decisions or the characteristics of the offender.

Community Service

“Community service is the most underused intermediate sanction in the United States. It is inexpensive to administer, produces public value, and can be scaled to the seriousness of the crime” (Martin, 2003, p. 26). It has been used as an alternative sanction in criminal cases in the United States for over forty years (Bouffard & Muftic, 2007). The literature on community service as a sanction, and particularly on its impact on recidivism as compared to incarceration, has found no significant difference in recidivism (Bouffard & Muftic, 2007; McDonald, 1989).

The Vera Institute Community Service Sentencing Project, for example, found no statistically significant difference between the recidivism rates of those sent to jail and those sentenced to community service (McDonald, 1989). The problem with these findings, particularly as they apply to suburban youths, is that the analysis, aside from dating back over 20 years, did not include juveniles and was limited to the Bronx, New York.

In a study of adult offenders in Finland, Muikuvuori (2001) reported that those who completed community service (20–200 hours) had slightly lower recidivism than did those sentenced to prison for less than 8 months. This research, aside from being from Europe and focusing on adults (both generalizability limitations), has methodological limitations not the least of which are selection bias and lack of randomization (Bouffard & Muftic, 2007). Those

sentenced to community service are generally less severe offenders than those sentenced to prison. So any sanctions aside, those sentenced to community service would be expected to reoffend at a lower rate than would those sentenced to prison terms.

Since diversion programs never send youths to prison, comparison to prison sentences is inappropriate. Comparisons to other sanctions like fines are far more appropriate. Bouffard and Muftic (2007) examined community service as compared to fines for low-level offenders and reported “those who participate in CS [(community service)] sentences are less likely to experience post-program recidivism, controlling for several initial group differences” (p. 171). Martin (2003), in a review of which sanctions work, reported that “work crew/community service has the lowest rates of reconviction for all high/medium-risk offenders” (p. 27).

Proponents suggest that since the recidivism after community service is similar to that after incarceration and the cost is lower, community service is a good alternative to incarceration (Martin, 2003). In restorative justice programs community service is viewed as part of the restitution to the community and is assigned in almost all cases. In the City of Ocoee during the time period of this study, when youths were sanctioned to community service they were offered a discount on the number of hours to be served if all of their hours were served within that city.

Curfews

Curfews for youth date back to Alfred the Great of England, who died just before 900 A.D. (Ruefle, 1996). Most of the literature on curfews focuses on curfews for all youths in a particular area and during a particular time period and not on curfews as a sanction. In the 1990s curfews became a popular tool for attempting to reduce youth crime and were endorsed by both presidential candidates (McDowall, Loftin, & Wiersema, 2000). In a study of the 200 largest

cities in the United States between 1990 and 1995, Ruefle and Reynolds (1995, 1996), reported that 60% had created or revised curfew ordinances. Surveys reveal that a majority of adults support juvenile curfews (Crowell, 1996).

Nevertheless, the fact that curfews are supported by the public and politicians and date back over 1100 years does not demonstrate whether they are effective in reducing youth crime. This section examines the relationship between the assignment of a curfew and recidivism.

Ruefle and Reynolds (1996), pointing out the flawed logic concerning curfews, stated:

Curfew laws rest on a simple premise: Controlling the hours when young people may be in public will limit their opportunities to commit offenses or suffer victimization. ... [The problem with this logic is] curfews apply to only a few hours of the day. Although several cities have ordinances that cover the times when young people are in school, curfews generally begin in the late evening and end in the early morning. Juvenile violence arrests peak in the afternoon, however, immediately after school ends. Arrest rates then decrease through the rest of the day, settling at low levels through the periods when most curfew laws are in effect. (pp. 77-78)

The studies on evening youth curfews either suggest no significant impact or suggest methodological problems.

Males and Macallair (1999) studied the relationship between curfews and arrest rates in California for juveniles (for crimes other than the violation of curfew) between 1980 and 1997. They reported no significant relationship between curfews and juvenile crime.

McDowall and colleagues (2000) examined 57 U.S. cities with populations greater than 250,000. This sample included about 65% of the population of the United States. Using data from the Federal Bureau of Investigation's Uniform Crime Report, the authors reported "limited evidence that curfews are effective in preventing some types of crimes. In particular, juvenile arrests for burglaries, larcenies, and simple assaults show statistically significant decreases after counties revised existing curfew laws" (p. 84). However, their data did not report a significant

decrease in those crimes themselves after curfews were enacted or revised. So the rate at which juveniles were arrested for those crimes decreased, but the actual occurrences of these crimes did not. This discrepancy suggests that while youthful offenders were avoiding apprehension for curfew violations, they were also avoiding apprehension for other crimes (Ruefle & Reynolds, 1995).

McDowall and colleagues (2000) also suggested a likely limited impact of curfews due to limited police resources to enforce them. Enforcement of curfews, like enforcement of traffic laws, is not normally generated by calls from the public, but rather as the self-initiated activity of police officers. Worden and Myers (1999) compared police self-initiated encounters in Indianapolis, Indiana, and St. Petersburg, Florida. A much higher percentage of self-initiated police activity was observed in Indianapolis (57%) than in St. Petersburg (35%). Worden and Myers suspected that the existence of a curfew in Indianapolis contributed to this disparity. Youths in violation of the curfew were stopped by police, and other violations of law were then observed. This conclusion supports the conclusions by Ruefle and Reynolds (1995) that juvenile evening curfews reduce the rate of youth arrest for crime but not the probability of the youth committing a crime.

The literature on gender differences (discussed in the gender section above) speculates that a curfew would be more effective for males than for females, and common sense suggests that curfews would be more effective for older offenders than for younger offenders, who would naturally be at home in the evening regardless of a curfew. However, no literature on curfews as a sanction could be found.

For the participants in the youth diversion program that provides the data source for this study, the mean, median and mode times for youth crime/arrest were 3:11 pm, 3:32 pm, and 3:18 pm, respectively. Clearly, curfew starting at 6 pm, for example, would not have prevented youth crimes, so the effect of a curfew is likely to be minimal, if any, on the youth recidivism rates. The impact of a curfew on those who committed violations in the late evening/early morning was, however, examined along with gender and age differences. The literature suggests gender differences in the effect of letters of apology, which are discussed next.

Letters of Apology

As for curfews, no academic literature could be found that directly tests letters of apology as a sanction to reduce recidivism in youth diversion, or in any other setting for that matter. Kelly, Kennedy, and Homart (2003), examining sanction packages customized to specific offenders (all shoplifters), reported that those who received customized sanctions had significantly lower two-year recidivism rates than a control group (chi-square = 32.72, $p < 0.01$). Letters of apology were part of the customized sanction package, but so were “fines, community service, monetary restitution, written essays, anti-shoplifting videos, ... and individual and/or family counseling” (p. 725). These findings support not the use of letters of apology, per se, but rather their use in customized sanction packages.

Restitution

Fines and restitution are unique sanctions for youth offenders in that they are sanctions that can be satisfied by the youth's parents and not necessarily by the youth. Children who receive a higher allowance (or those whose parents simply pay the fines themselves) are less

affected than youths with jobs who pay the fines from their own incomes. Nonetheless, since restitution is part of the restorative justice repertoire, its impact on recidivism should be examined.

Few studies focus on restitution/fines in youth sanctioning, and those that do suggest it is effective. The Schneider (1990) study, for example, compared restitution (only part of what might be considered restorative justice), probation, and incarceration. The author reported that for every 100 youths who completed restitution programs, there was a decrease in 18 offenses per year as compared to the rate for those sentenced to probation or incarceration.

Just as with studies of boot camps and community service, selection bias is a possible flaw. Those sanctioned to pay fines (whether or not they ever did so) would be expected to have lower recidivism than those sanctioned to incarceration (whether or not the time was ever served). The characteristics of the youths (and the crime) that precipitate the sanction assignment may have more influence on recidivism than the sanction assigned, but these effects are much harder to measure.

Drug Testing

The next sanction examined is drug testing. Haapanen and Britton (2002) studied 1,958 paroled youths in California who had been randomly assigned frequencies of drug testing (from not at all to twice a month). They found that the frequency of drug testing had no impact on recidivism; however, “early positive tests ... indicated increased risk of recidivism” (p. 217). “To date, no studies have suggested that drug testing among regular offender populations helps to reduce recidivism” (p. 218).

Britt, Gottfredson, and Goldkamp (1992) evaluated the random assignment of random drug testing in two counties in Arizona. They found either no difference or a slight increase of failure in those assigned drug testing. The increase in failure could have been a function of a failed drug test. This Arizona study focused on adults and made no distinction between drug and non-drug-related offenders. The current study has examined the efficacy of drug testing on recidivism specifically for both drug and non-drug-related offenders.

Anger Management

Youth sanctioned for the commission of violent crimes are often sent to anger management counseling. The logic here is it that if the offender can better manage anger, there is likely to be fewer violent manifestations of it. Dowden and Serin (1999) examined matched pairs of 110 federal male prisoners. Matches were based on age, offense and admitting their violation.

Results revealed that for the lower-risk cases (n = 54), completion of the anger management program was not significantly associated with reduced levels of non-violent reoffending. However, when the analyses focused on higher-risk cases (n = 56), significant reductions in non-violent recidivism were found. This translated into a 69% reduction in nonviolent recidivism (i.e. 39.3% recidivism rate for the comparison group versus 12.5% of the anger management group) ... As expected, completion of the anger management program failed to produce significant reductions in violent recidivism among lower-risk cases (base rate = 7.4%). However, for the higher-risk group, completion of the anger management program was associated with significant reductions in violent reoffending. This translated into an 86% reduction in violent reoffending (25% violent recidivism rate for the control group versus 3.6% for the anger management group). (p. 3-4)

This study demonstrated some benefit from anger management for this population. Its benefit for a population of co-ed youthful offenders remains to be seen.

No-Contact Orders

Some diversion programs use no-contact orders to prevent violations, especially while in the program. If codefendants were arrested together, for example, it is the logic of the sanctioning bodies that their best chance for not reoffending would be to stay away from each other. Many states' probation/parole participation forbids the probationer from associating with known felons. The same logic is extended to youth diversion program sanctioning.

In the section above on gender, the theory of gender differences was cited as a basis for the assumption that no-contact orders would be more appropriate for males than for females (Elis, 2005). These assertions remain speculative because no research that tested no-contact orders could be found. Research on the presence of a crime partner actually suggests lower recidivism, at least for Caucasians.

Levine (2009), in a sample of 413 young male offenders sentenced to a commitment program in California, found that 236 (57%) had a partner during their original criminal act. In studying the relationship between presence of a crime partner and recidivism, Levine reported that "crime partners present predicts time to recidivism among Caucasians, [and] predicts incidents of recidivism among non-Caucasians" (Levine, 2009, p. 172). From the above findings (and common sense) one might expect the presence of a partner in the current charge to increase the probability of reoffending and reduce the time to live. That was not the case. Levine observed that for Caucasians with two or more prior arrests, the median time to live for offenders who had a crime partner was 4.5 years, but for those without a partner was 1 year. Perhaps the crime partner, serving as a lookout, reduced the probability of apprehension.

Intense Supervision

The Orange County (FL) Sheriff's Office has a youth program called Juvenile Arrest and Monitor (JAM). The JAM unit designates juveniles who on the basis of their past crimes are viewed as an extreme threat to public safety. They are youths who have committed multiple felonies (robberies, rapes, and attempted murders). The Orange County Sheriff's Office places these offenders under intense supervision. On average, JAM officers make contact with participants four to seven times a week, at home, school, work, or places they are known to frequent. The offender also receives frequent drug tests (Reynolds, Myers, & Dziegielewski, 2002).

This intense supervision provided JAM with dramatic success. Reynolds and colleagues (2002) found that 81% of participants were not arrested during the program, and that 80% remained arrest free for one year after completing the program. Although the JAM program was successful considering the clientele, the approach would be impractical and very expensive for use on a population of first-time, mostly non-violent misdemeanor offenders. It does, however, suggest the importance of frequent contact with offenders. The NRJP program requires offenders to call their sanction coordinator once a week. The use of time stamp and caller identification allows these calls to serve as a check of curfew compliancy as well, assuming a curfew is assigned.

Aftercare

A recent trend in the literature suggests the importance of ongoing supervision by and contact with service providers, called aftercare. Programs that offer aftercare have seen lower recidivism rates than for those that do not, all other things held constant (Fredrick & Roy, 2003;

Josi & Sechrest, 1999). Furthermore, the literature suggests that the longer youth services are provided as a choice of either the youth or the sanctioning body, the lower the recidivism.

For example, a review of the East Point Georgia Police Department Youth Program, discussed above, revealed that many youths chose to continue to participate after their minimum sentence had been completed. That program boasted a zero recidivism rate (Georgia 'juvenile diversions', 2001). Aside from the methodological problems with this study discussed above, selection bias comes under question. Youths who choose to participate in police-run programs would naturally be expected to have lower recidivism than that seen for those who do not. Lengths of service comparisons present another methodological issue, non-comparability of offenders.

For example, comparisons of residential commitment programs with non-residential treatment programs have comparability issues related to the type of offender sanctioned to each program. During 2003, for example, 8,388 juveniles completed residential programs in Florida at a cost just over \$1.1 million. With an average 11-day treatment, DJJ spent about \$12 per day and saw six-month and one-year recidivism rates of 16% and 23%, respectively. During that same year, 28,020 juveniles were released from non-residential programs costing \$46.7M that had provided an average treatment of 143 days. Those programs cost taxpayers \$0.35 less per day and saw six-month recidivism of 7% (FDOC, 2003; FLDJJ, 2003). One might conclude that the non-residential programs were more effective, or that length of service was crucial in reducing recidivism. The problem with conclusions drawn from these statistics, however, is that far more serious offenders are sanctioned to the residential programs, so higher recidivism rates for those participants should be expected, all other factors held constant.

In the current study, the youth were sanctioned for a 90-day period. Any period longer than 90 days would have indicated an extension for a violation or failure to complete an assigned sanction. For that reason, length of sanction time was not examined as a predictor of recidivism. Length of sanctions in restorative justice programs and their impact on recidivism should be examined, but because of a standardized time period, such an analysis was beyond the scope of this study.

Application of Literature and Theory on Sanctions to the Current Study

This section presented findings from the literature about specific sanctions and their relationship with recidivism. Boot camps were not shown to reduce recidivism, but that study raised questions about its comparability and selection bias (Blair, 2000). Community service was found to be effective at reducing recidivism for lower-level offenders as compared to fines (Martin, 2003), but not as compared to incarceration for more serious offenders (Bouffard & Muftic, 2007; McDonald, 1989). Anger management was found to reduce recidivism for adult males, but its impact on females and younger males is not well established (Dowden and Serin, 1999). Failing a drug test was found to be a significant predictor of recidivism, but the relationship between assigning drug testing and recidivism for all offenders in a program (not just drug offenders) was not found to be significant (Britt et al., 1992). The relationship between assigning drug testing and recidivism for drug offenders was tested in the current study and was expected to be significant.

Very little research has been done on sanctions like letters of apology and curfews. The literature on curfews suggests limited impact when used for all youths (Ruefle & Reynolds, 1995). Youths in the data set examined in the present study had committed a vast majority of

their crimes before the time when most curfews start, so the impact of curfews was expected to be minimal if any.

Because multiple sanctions are normally assigned at the same time, it is the relationship between the entire sanction package and recidivism that was examined. In order to accomplish this, the latent construct Multiplicity of Sanctions was developed. The indicators that were studied for how they contribute to the conceptualization of that construct were boot camps, community service, curfew, letters of apology, restitution, drug testing, anger management, and no-contact orders. The measurement model shown in Figure 12 in Appendix B was tested against the data for restorative justice program participants. The measurement models for Recidivism, Offender Risk Propensity and Multiplicity of Sanctions were later used to develop predictive models. These models tested the main, moderating and mediating effects noted in the introduction.

From previous research ((Elis, 2005) and the Relational/Instrumental Theory (Hagan et al., 1988)), it was expected that letters of apology would be more effective in reducing recidivism for older females than for younger males, but that no-contact orders and curfews would be more effective in reducing recidivism for males than for females. Anger management was expected to be more effective for males than for females, and drug testing was expected to be effective only for drug offenders. The conclusion of this chapter presents the predictive models tested.

Conclusions

This chapter presented literature and study findings related to the research project. The topics included punishment philosophies, the conceptualization of Recidivism, restorative

justice, Offender Risk Propensities, and Multiplicity of Sanctions. This review is the foundation for the research methodology discussed in the next chapter.

For the purpose of this paper, the mission of youth corrections is assumed to be rehabilitation. With that in mind, reducing recidivism is the objective. In order to meet the objective, recidivism must be measurable. Many different ways to measure, or more accurately conceptualize, recidivism have been presented. These measurement conceptualizations were then applied to studies of youth recidivism to better understand the impact of the offender's demographics as well as the impact of the sanctions assigned. The three models presented in Figures 13–15 in Appendix B represent three explanations of the relationships between the constructs.

The main effect (see Figure 13 in Appendix B) tested the relationship between Multiplicity of Sanctions and Recidivism. In the mediating effect model (Figure 14 in Appendix B) the Multiplicity of Sanctions was tested as a possible mediator between the Offender Risk Propensity and Recidivism. It has been well established in the literature that Offender Risk Propensity, the initial variable, is correlated with recidivism, the outcome. This model suggests that Multiplicity of Sanctions impacts Recidivism as an intervening variable based on the Offender Risk Propensity. The Moderating Effect (see Figure 15 in Appendix B) suggests that a combination of the Offender Risk Propensity and the Multiplicity of Sanctions impacts recidivism. It was expected that the mediating model would be the best model when tested against the data set. Now that the literature has presented a foundation and the models have been introduced, the methodology for testing these models and relationships will be discussed. The focus of the next chapter is methodology.

CHAPTER 3: METHODOLOGY

Chapter 1 introduced the study problem: the lack of evidence-based research to guide pre-trial juvenile court diversion programs in the sanctioning of youth offenders and the problems associated with the measurement of recidivism. The chapter defined the nomenclature to be used in this study, established the significance of this research problem, and pointed out weaknesses in the current research as it applies to that problem. Chapter 2 reviewed the literature and relevant theories and reported study findings about recidivism and restorative justice programs, Offender Risk Propensity, and Multiplicity of Sanctions. Relationships between constructs and their indicators and between constructs were explained, theorized, and proposed.

This chapter presents the methodology for testing these relationships. The chapter lists the hypotheses tested, discusses the specific analytical tools used to test the hypotheses, defines the endogenous, exogenous, and control variables, and addresses in advance how the results were evaluated. The data source, sampling strategy, unit of analysis, reliability, validity, and power are addressed. This chapter begins by introducing the hypotheses to be tested.

Specific Hypotheses Tested

This section lists the hypotheses tested. From previous research and the theories presented above, eight hypotheses were developed:

- 1) How should recidivism be measured? H_1 : Measuring recidivism as a latent construct (using the indicators of reoffended (y/n), severity of new charge, most severe sanction imposed, days charge free, total number of new charges, and change in crime level between the original violation and any new violation of law, if one exists) explains the variation in the

construct of Recidivism. The Recidivism measurement model was tested and modified as necessary, to determine its fit with the data. Model modification and testing are presented later.

- 2) H₂: The following conventional recidivism findings about youthful offenders apply to participants in restorative justice programs:
 - a. Males have higher recidivism than females.
 - b. Offenders who offend at younger ages (less than 15 years of age) have higher recidivism than those who offend later in life (at least 15 years of age).
 - c. Offenders with a history of offending (measured by suspension from school) have higher recidivism than do those without a history of offending.
- 3) H₃: Youth who complete a restorative justice program have significantly lower recidivism than do those who opt not to participate and those removed from the program.
- 4) H_{4A}: Drug offenders sanctioned to undergo random drug tests in restorative justice programs have lower recidivism than do drug offenders not assigned drug testing. H_{4B}: For non-drug-offenders, there is no significant relationship between the assignment of drug testing and recidivism. H_{4A} and H_{4B} are not reciprocals. H_{4A} contends that drug testing works for drug offenders. The reciprocal of that would be that drug testing does not work for drug offenders, which is not what H_{4B} contends. H_{4B} contends that drug testing does not work for non-drug offenders in general.
- 5) H_{5A}: There is no significant relationship between the assignment of curfews as a sanction for all offenders and Recidivism. H_{5B}: There is a significant relationship between the

assignments of curfews as a sanction for male offenders at least 15 years of age and Recidivism.

- 6) H₆: No-contact orders are an effective sanction for offenders who were charged with a codefendant, for drug offenders, and for males at least 15 years of age. No-contact orders are not an effective sanction in reducing recidivism for all offenders in general.
- 7) H₇: Letters of apology and essays are effective sanctions for females in reducing recidivism.
- 8) H₈: The relationship between Offender Risk Propensity (a latent construct with indicators of race, age, gender, crime type, and drug and school discipline history), Multiplicity of Sanctions (the total package of sanctions assigned), and Recidivism is significant.

The task of testing the hypotheses began with identifying the data source, defining and conceptualizing variables, defining the testing methods and approaches, and addressing the issues of validity, reliability, and power. The remainder of this chapter presents these steps, starting with the data source.

Data Source

The data for this study were collected by the Ninth Judicial Circuit's Neighborhood Restorative Justice Program as part of their internal record keeping and performance evaluations. The data set was composed of 221 cases of offenders who had been invited to participate in the program during a 26-month period of January 2002 through February of 2004. These cases included youths who participated in that program in the Florida cities of Ocoee, Winter Garden, Maitland, Eatonville, and Apopka. Demographic information about the offenders, offenses, and sanctions were collected by the program from the official charging documents submitted to the program. The recidivism information collected by the program as part of a program review from

the Florida Department of Juvenile Justice during the month of January 2009 was included in the data. Data from cases prior to 2002 were destroyed by the court.

The data, which were exported from the program's database, included information on program completion (yes/no), the specific program (city), sanctions imposed, demographic information, and information on subsequent offenses, if any. A great deal of the data was not used by this study; that data included results from two surveys the parents of offenders were asked to complete on the youths' behavior and home environment. Although this information, aside from having questionable reliability with many gaps, was not needed for this study, it was not removed for the data set in the event that it may be useful later.

Data Security

Because this data set included juvenile information, several methods were employed to ensure data security: limited data availability, encryption, and password protection using strong and complex passwords. This section discusses those steps.

The most important step was never to have any information that could be used to easily identify juvenile participants. Prior to when the data set was received each offender's name, address, height, weight, eye and hair color, and employer information were removed. The program's case number, arresting agency case number, court case number, corrections case number, and Department of Juvenile Justice offender identification numbers were removed as well. This step ensured that the individual offenders could not be identified. Cases were assigned study case numbers (from 1 to 221), which were based on their order in the dataset.

Second, password protection was used where possible. The data set was stored on the hard drive of a computer. Access to the computer required one password. The data set was

backed up to a Universal Serial Bus (USB) Kingston brand thumb drive, which used Data Traveler Vault Privacy, a hardware-based, 256-bit Advanced Encryption Standard. The drive is automatically reformatted after ten failed password attempts (Kingston, 2011).

The third step, used in conjunction with the first two, was the use of strong and complex passwords. Strong passwords require eight or more characters and the simultaneous use of capital letters, numbers and special characters (such as !@#\$^ and &) (Microsoft, 2011). The passwords for the computer, files, and encrypted thumb drive all had unique strong and complex passwords that were known only to this author.

These steps, which might appear unnecessary, nevertheless ensured complete security of the data set. The data set was used to test the hypotheses listed above through the analytical steps described below after the unit of analysis is explained.

IRB Determinations

The author completed the Collaborative Institutional Training Initiative (CITI) Human Research Curriculum Social / Behavioral Research Investigators and Key Personnel on November 3, 2011 (Reference number 6972368) as well as Human Subject Research from the National Institute of Health. This study was submitted to the University of Central Florida's Institutional Review Board (IRB). The IRB determined that "the proposed activity is not human research as defined by DHHS regulations at 45 CFR 46 or FDA regulations at 21 CFR 50/56." The IRB notification letter can be found in Appendix C.

Unit of Analysis

The subject of a study is the unit of analysis, which limits its boundaries. A study can focus on “such units as groups of persons or organizations, key decisions, public programs, or organizational change. In each situation, the corresponding unit of analysis is different (the group, the decision, the program, or the change)” (Yin, 1997, p. 237). In this case, the individual (and his/her characteristics, sanctions received, and recidivism) was the focus of the study, and therefore, the unit of analysis was the individual youth who participated in the Ninth Judicial Circuit Court’s Restorative Justice Program from January 2002 through February of 2004.

Hypothesis Testing Methods

The listing of the hypotheses and the identifying the data set to be used having been completed; the next step is to explain how each hypothesis was tested. As explained later, many of these hypotheses were tested first through the use of structural equation modeling (SEM) and Confirmatory Factor Analysis (CFA). To accomplish this, indicators were coded and their contributions to the latent construct were tested in terms of both their amount of contribution and whether that contribution is significant. Correlated errors were added to the models, and the models were tested against a data source. Then measurement models were developed into predictive models to which exogenous variables (e.g., the indicator of program completion’s relationship with the measurement model of Recidivism) or other measurement models (such as the relationship between the constructs of Offender Risk Propensity, Multiplicity of Sanctions, and Recidivism) were added to test the relationships between them.

In the event that overall models were not found to be a good fit with the data, the related hypotheses were tested by examining the relationships between the individual measures of each

construct. For example, if the measurement model for Multiplicity of Sanctions was found not to be a good overall fit with the data, the relationship between apology letter assignment (a categorical independent variable) and recidivism as defined by one of its indicators, for example, the number of days without a new charge (a continuous dependent variable), would have been tested using the Mann-Whitney U test. If both the independent and dependent variables were categorical, the chi-square tests were used to test the hypothesis. The tests used are illustrated in Table 2 in Appendix A. The next several sections explain these processes, starting with an example to illustrate the power of SEM.

Structural Equation Modeling (SEM)

Measuring a person's weight and height are relatively simple tasks. Using measurement instruments such as a bathroom scale and measuring tape produces measures that in conventional units most would understand. Measures such as six-foot, 195 pounds, for example, can be easily made, conveyed, and understood. These are unidimensional methods, as only one measurement instrument for height and weight is necessary. There are conditions, however, for which accurate and complete measurement requires multiple simultaneous measures. Consider attempting to measure a male's level of fitness. A single measurement instrument (like the bathroom scale) would fail to capture an accurate measure of health, especially for a bodybuilder (who may appear overweight to the scale alone), for example. That bodybuilder might be strong yet at the same time in poor cardiovascular health. Accurately conceptualizing this man's fitness level would require a multidimensional approach where multiple factors (such as weight-to-height ratio, blood pressure, cholesterol levels, maximum bench press, and two-mile run time) all

simultaneously contribute to the measurement of the latent construct of fitness. This concept of measurement was applied to the conceptualization of Recidivism.

Structural equation modeling, which is used in confirmatory research to better understand latent constructs such as the quality of health care (Wan, 2002; Zhang & Wan, 2007), can be applied as well to the understanding of Recidivism. “With seemingly few exceptions, current practices in structural equation modeling aims at establishing close rather than exact fit between hypothetical models and observed data” (McIntosh, 2007, p. 859). According to Wang, Wan, Clement, and Begun (2001):

The SEM uses a two-step process: the measurement model and the structural equation model. Briefly speaking, the measurement model specifies how the latent (observed) variables or hypothetical constructs are measured in terms of the observed variables. Based on results derived from the measurement models, the structural equation model specifies the casual relationships among the exogenous and endogenous variables and describes the amount of unexpected variances among them. (p. 184)

The SEM process is further explained below as it applies to the construct and measurement model of recidivism.

Measurement Model of Recidivism

The most important construct of this study and the endogenous variable under scrutiny is Recidivism. This section explains the selection of the indicators and the construct formulation for the model. Understanding the target variable of Recidivism is crucial before attempting to understand what variables affect it.

“All across the country, juvenile justice agencies are judged to be successful or not based on recidivism rates that indicate the extent to which youths commit crimes after receiving juvenile justice services” (Stoodley, 2010, p. 86). Measuring recidivism can be problematic.

According to Myers and Travis (2004) as cited in Stoodley (2010), recidivism, when measured only by re-arrest, can be as high as 66%, but, when defined as re-adjudication or conviction within a few years of receiving services, can be as low as 33%. Even though the lack of standardized definitions of recidivism hampers research communication, it is still by far the most frequently used measure for assessing the success of programs.

“About 59% of Florida's juvenile offenders never return to the juvenile justice system after first arrest” (FLDJJ, 2010, para. 3). This would make the state's recidivism rate 41%. The statistic fails to answer such questions as: Of those who reoffend, how soon do they reoffend? Of those who reoffend, how severely do they reoffend (i.e., do the crimes get more or less severe)? Of those who reoffend, how severely do they reoffend (i.e., what was the most severe sanction assigned to the offender)? To more fully understand (and measure) recidivism a more complete measurement model is needed. This section explains the proposed measurement model for Recidivism, which was conceptualized in this study by six indicators.

Indicator “Reoffended (y/n)”

The first variable, Reoffended (y/n), is a dichotomous variable that indicated whether the juvenile reoffended (as defined by any new charge by the police, regardless of action taken by the State Attorney or the courts). Throughout this paper this variable may also be referred to as re-arrested or recharged; however, whether or not an actual physical arrest was ever made, as opposed to the filing of charges, is not known. This variable indicates only that a criminal charge was referred to the courts by an arrest by the police and/or by the filing of charges by a state attorney. In most recidivism studies this is the only indicator of recidivism. This variable was coded as 0 or 1 (see Table 3 in Appendix A).

Indicator “Severity of New Charge”

The second variable, Severity of New Charge, is an indication of the level of the crime with which the reoffender was charged as a new offense (as defined by Florida State Statute 775.081) (Florida Senate, 2011). The coding for these crime levels, which is listed in Table 3 in Appendix A, is a continuum of offenses from no new offense to a capital felony offense as defined in Florida law.

In the event of multiple criminal charges (either in the same criminal incident or in separate incidents), the hierarchy rule implemented by the Federal Bureau of Investigation for collecting crime statistics through the Uniform Crime Report, was used. According to Roberg, Novak, and Cordner (2005), this approach limits “reporting to the most serious offense even though multiple offenses [may have been] committed” (p. 129). If the offender had not reoffended, the variable was set at zero (0), and violation of probation for non-criminal activity (such as violation of curfew) was treated like a status offense (e.g., running away or possession of alcohol by a minor) and was given a score of one (1).

Later in this analysis some of these categories needed to be condensed to comply with the requirement of the chi-square test that at least 80% of cells have an expected frequency of five or more. Of those charged with a new offense, only two had new status offenses, only three had second degree misdemeanors, and there were no life or capital felonies. So of the nine categories in this variable, four categories had frequencies (not expected frequencies) of less than five. Thus for the purpose of the chi-square test (and specifically to satisfy the assumption that at least 80% of cells have an expected frequency of five or more), this indicator was reduced to three categories: 0 (no new offense), 1 (all misdemeanors including status offenses), and 2 (all

felonies). The data set contained 146 offenders with no new charge, 26 with new misdemeanor charges and 46 with new felony charges. This new variable was called Severity of New Charge Category.

Indicator “Crime Level Change”

Following the same ordinal categories listed for the severity of new charge, the indicator Crime Level Change is a measure of the difference between the original charge and the most severe new charge. This indicator measured recidivism as a function of the change in crime from the original violation to the most severe of any new charge. For example, if an offender had originally been charged with assault, a second-degree misdemeanor (coded as 1), and was later charged with burglary to a vehicle, a third-degree felony (coded as 4), the change in crime level would be 3 (4 - 1). If an offender had originally been charged with vehicle burglary (coding of 4) and did not reoffend at all (coding of zero), his or her change in crime level would be -4. This indicator demonstrated both the change in crime level and the direction of that change.

Just like the indicator Severity of New Charge, Crime Level Change was also condensed. This method of calculation could produce a possible range from -8 (originally charged with a capital felony (8) and no new charge (0)) to 7 (originally charged with a status offense (1) and reoffended as a capital felony (8)). The range from a possible -8 to 7 is 16, which is technically possible but also extremely improbable. The data set had an actual range from -5 to 3, or nine possible categories. With eight of the possible 16 categories having actual frequencies of zero, the need to reduce the number of categories in this variable also became apparent. The number of categories was reduced to two categories: (-1) did not reoffend at a higher level and (1) did reoffend at a higher level. This new variable was called Crime Level Change Sign.

Indicator “Most Severe Sanction”

The next indicator, the most severe sanction given for a subsequent criminal act, is a continuum between no new charge and execution. It is unlikely that a youth offender from this program would ultimately be executed (or even assigned that sanction), but it is in the spectrum of sanctions available from the courts.

This indicator can be seen as a measure of many different factors, including criminal history, severity of the crime, and quality of evidence against the offender. For example, if an offender was charged with a crime and the state declined to prosecute, it is possible that the crime was minor, that the evidence submitted by law enforcement was insufficient to prove the charge, and/or that the offender did not have a lengthy criminal record. If, in contrast, that offender was sanctioned to high security confinement or was transferred to adult court, this would indicate a more severe charge, an abundance of evidence, and/or a lengthy criminal history. This variable was coded from zero (indicating no new charge) to 11 (indicating execution), as shown in Table 3 in Appendix A.

Again, this indicator had to have the categories reduced for use with the chi-square test. This category had twelve possible scores, from 0 (no new charge) to 11 (execution). Of the twelve possible categories, the data set contained only ten. This was reduced to 2 categories: (0) sanctioned to less than probation and (1) sanctioned to probation or higher. This new variable was called Most Severe Sanction Category.

Indicator “Number of New Charges”

The next indicator of Recidivism is the number of new charges the offender received. Other researchers have used similar measures of recidivism, such as the number of new offenses

per 100 offenders per year (Sherman et al., 2000). This continuous indicator was used because it distinguishes among offenders who reoffend at different frequencies. There is a big difference between an offender who reoffended only once and another who reoffended 50 times. Simply using an indicator of new charge would fail to capture that variation in the construct of Recidivism.

Indicator “Days Charge Free”

The last indicator in the Recidivism measurement model is a measure of how soon (in number of days) an offender reoffended following the start of participation in the program. Because completion date was not in the data set, contract date was used: that is the date when the youth formally started the program and the sanctions were set (or the date when the youth refused to participate). The standard participation period for this program was 90 days, but because the cities have their programs meet on the same days of the week (Ocoee, for example, met on Tuesday evenings), participants are never in the programs for exactly 90 days. This continuous indicator was calculated by subtracting the date of the contract from the date of the first new charge. If, for example, an offender’s contract date was 9 October 2002 and reoffending occurred on 27 May 2003, then that offender went 230 days without a new charge.

In the event of no new charge, the indicator was calculated by subtracting the date of the contract from the date of the check for new violations. The checks were done over a two week period ending Thursday, January 15, 2009. Because it is unknown exactly when during that two-week period each offender was checked for new violations, that date was used. So an offender who had a contract date of 24 March 2003 and did not reoffend as of the check date would have been charge free for 2,124 days as of the date of the check.

In order to have all indicators of recidivism use the same direction, this indicator was made negative by multiplying it by -1. For the indicator number of new charges, the higher the number the more Recidivism had occurred. Thus -230 is greater than -2,124, indicating more Recidivism for the offender who lasted only 230 days, as opposed to the offender who lasted 2,124 days.

Admittedly, this variable presents methodological problems because of censored data, in that it is unknown whether or not the offender reoffended beyond the time period examined. Because of this, an artificial re-offense date was created, the date of the criminal history check. Using this approach, there is no difference between someone who reoffended on that date and one who did not. Censorship of data actually occurred for all indicators, and in cases of those who did not reoffend during the examination period, the censorship for many of these indicators was in favor of not reoffending. Because it is unknown whether or not the offender reoffended after the examination period, for the purposes of indicators like Reoffended (y/n), it is assumed that the offender did not reoffend. Surely, some of these offenders reoffended after the examination period; that is discussed in more detail later with regard to the study's limitations.

Recidivism Measurement Model

The measurement model for the construct of Recidivism was conceptualized through a combination of the indicators listed above and coded as noted. Those indicators sought to answer the following questions: Did the youth reoffend? How soon did the youth reoffend? How severely did the youth reoffend? How often did the youth reoffend? If the youth reoffended, did he or she reoffend more or less seriously than shown by the original charge? The plan was to test

this model against the data set, and if it were shown to be significant, to use it to build predictive models to test other hypotheses.

Measurement Model of Offender Risk Propensity

Offenders cannot be described accurately with only one variable, and therefore studies that consider only one variable at a time may develop incomplete conclusions. For example, studies that examine only gender relationships with recidivism may fail to uncover further differences that are apparent when age, race, crime type, and past history are examined simultaneously. The literature suggests a positive relationship between age of first arrest and recidivism (Ashford & LcCroy, 1990; Levine, 2009), but does that hold true, for example, for white females charged with property offenses? This section discusses the development of the latent construct of Offender Risk Propensity and the indicators that contribute to it, starting with crime type.

Indicator “Crime Type”

The relationship between the crime with which an offender has been charged and recidivism has received a great deal of attention in the literature, as discussed above. The literature has focused on violent offenders, property offenders, drug offenders, and status offenders. In keeping with these categories, the indicator Crime Type was given the following categorical coding: (1) drug offense (e.g., possession of drugs or DUI), (2) property offense (e.g., theft or burglary), (3) violent offense (such as assault and battery), and (4) status and other offenses (e.g., possession of alcohol by a minor and probation violations). This variable served as in indicator of Offender Risk Propensity.

Indicator “Race”

Race also has received a great deal of attention in the literature. Race was tested as an indicator of the construct Offender Risk Propensity. This data set defined race only as white (including Hispanic) or black, so the following coding was used: (1) Caucasian (including Hispanic) and (2) black.

Indicator “Age”

Age at first arrest/charge has been seen as a significant factor in recidivism studies. Age as an indicator of Recidivism was a continuous variable, simply the age of the youth as reported at the time of arrest multiplied by -1. Because younger offenders are expected to have greater recidivism than older offenders, age was made negative to prevent a negative factor loading. Age categories were also used in non SEM tests, as described later. The coding methods for age groups are discussed later.

Indicator “Gender”

Gender, as discussed in the literature review, has received a great deal of attention in relation to recidivism. Unlike age, gender cannot be coded as a continuous variable. Gender was coded as: (-1) male and (-2) female. Because male offenders are expected to have greater recidivism than female offenders, gender was coded this way to prevent a negative factor loading. Some methods of analysis require a 0/1 coding. For those tests male received a coding of 0 and female received a coding of 1.

Indicator “Suspended”

The literature suggests a relationship between past history of delinquency and recidivism. Because this data set comprises almost entirely first offenders, a history of school suspension was used as an indicator of past delinquency. The data set indicated only whether there were suspensions and, if so, the reasons. It did not indicate the number of suspensions. The coding of this indicator was (0) never suspended and (-1) having been suspended. Because offenders with a history of suspension are expected to have greater recidivism than offenders who were never suspended, the variable suspended was coded this way to prevent a negative factor loading.

Indicator “Drug Offender”

The link between drug use and recidivism is well supported by the literature, as discussed above. Therefore, the variable drug offender was used as an indicator of the construct Offender Risk Propensity. This variable was no, coded as zero (0) if the offender was not charged with a drug-related crime and had not admitted to drug use. It was yes, coded as negative one (-1), if the offender had a drug-related charge or admitted to drug use. Because drug offenders are expected to have greater recidivism than non-drug offenders, the variable drug offender was coded this way to prevent a negative factor loading.

Measurement Model

The construct of Offender Risk Propensity was conceptualized as a function of the indicators discussed above. These indicators were supported by the literature as used

individually, but the current study tested these constructs simultaneously. Next, the construct of Multiplicity of Sanctions and its indicators is discussed.

Measurement Model of Multiplicity of Sanctions

The measurement model for Multiplicity of Sanctions was conceptualized as a function of the individual sanctions. This section discusses these indicators and the coding method employed. The contribution of each indicator to the construct was evaluated when the measurement model was compared to the data.

Indicators “Yes/No” Sanctions (Dichotomous Coded Sanctions)

Unlike many of the indicators discussed above, which could have been coded as categorical or continuous variables, many of the indicators of Multiplicity of Sanctions are dichotomous. These indicators included the assignment of no-contact orders, anger management, drug testing, letters of apology, essays, assignment of community service, and boot camps, and they were not assigned as continuous variables, such as 20 hours of boot camps or 8 random drug tests. They had been either assigned or not assigned to the offender in the program; therefore their coding was either zero (0) for not assigned or one (1) for having been assigned.

Indicators (Continuous or Ratio Coded)

Some of the indicators could have been coded as continuous variables. The number of community service hours assigned, the amount of restitution assigned, and the numbers of letters of apology required were all coded as continuous ratio variables. These indicators could have been tested to determine their contribution to the construct, but not with the data set used in this

study. These variables had too much missing information for an adequate analysis. Originally, the assignment of a curfew was coded as a dichotomous variable, either assigned or not assigned; however, because almost all offenders were assigned a curfew, the assigned curfew time was used. For the SEM analysis, the curfew time (in military time) was multiplied by -1 to prevent negative factor loading.

Measurement Model

The measurement model for Multiplicity of Sanctions (see Figure 11 in Appendix B) was conceptualized as a function of the indicators or specific sanctions assigned to the offender. Many of these sanctions have been discussed in the literature review, and their relationship with recidivism has been tested in previous studies. Many sanctions have not been tested on populations of youthful offenders, however, especially in diversion programs. Examining multiple sanctions simultaneously may provide a better understanding of the impact sanctioning has on recidivism, if any.

Control Variables

The data set for this analysis was provided by the Neighborhood Restorative Justice Program from the Ninth Judicial Circuit of Florida. As of the dates of data collection, that program operated in the cities of Apopka, Winter Garden, Ocoee, Maitland, and Eatonville. Each city ran its program slightly differently, especially with regard to the roles of the police department representative and the volunteers, both of which could have served as control variables in the current study.

The five cities had major program differences that changed over time. In the city of Ocoee, at least until late 2003, civilian volunteers acted as the offenders' sanction coordinators. In that capacity a volunteer collected proof of sanction completion, monitored the youths' progress, and in some cases did home visits to check on the youth. The Ocoee program, only during the time period examined, also accepted many cases of youths who lived in the neighboring city of Winter Garden, which had not yet established a program. During one year the Ocoee program had as many participants from Winter Garden as from its own city.

In a small city like Eatonville, it was not uncommon for the Chief of Police himself to sit in on the conferences, whereas in other cities, like Maitland, the police representative was a patrol officer. Because these programs underwent so many changes during the examination time period, the city of participation was not used as a control variable.

Hipple and McGarrell (2008) compared police- and civilian-run conferences to discover whether there were differences in procedures, agreements (or sanctions), over-all recidivism rates, and time to failure. They found "no major differences between conferences facilitated by civilians as opposed to police officers" but observed "police officers seemed to lecture offenders more" (Hipple & McGarrell, 2008, p. 553). They also found time to failure (or time to reoffending) slightly longer for youths who participated in police facilitated programs as compared to those in civilian-run programs, but these results were not statistically significant.

The individual city programs during the period studied underwent major changes. The status of the sanction coordinators, the structures of the sanctioning bodies (in which the officers had varying degrees of involvement), and even the areas served by the programs underwent

drastic changes. For these reasons, the status of the sanction coordinator was also not used as a control variable.

A benefit of structural equation modeling is that when multiple indicators are used simultaneously, the model has the effect of controlling for one variable while examining another. Therefore, the final models (main, moderating, and mediating effects) examined one relationship while controlling for others. The process is discussed in more detail when the analysis of curfew time (as an effective sanction while controlling for the offenders' age) is presented.

Predictive Models

Once measurement models are developed and tested, predictive models can be developed and tested. Predictive models seek to predict the variation in one construct based on the variation in another construct or indicator. In the current study the latent constructs of Offender Risk Propensity and Multiplicity of Sanctions were predictors of the construct of Recidivism. This produced a variety of predictive models, which were tested later in the analysis.

Explanation of Statistical Design for Testing Hypotheses

This analysis was accomplished through an analysis of descriptive statistics and correlation using SPSS and Covariance Structure Modeling and multi-group Confirmatory Factor Analysis using AMOS. Model fit was evaluated through chi-square, relative chi-square, *p*-value, GFI and AGFI (goodness of fit and adjusted goodness of fit), RMSEA (root mean square error of approximation), and CFI (comparative fit index).

In drawing conclusions from the measurements discussed above, a low chi-square value with a *p*-value greater than 0.05 indicates that the model developed is a good fit for the data

observed. Chi-square is a comparison (or difference) between the data and the model, so the lower the chi-square, the better the model fit. A relative chi-square value (chi-square/degrees of freedom) less than 5 and GFI, AGFI, and CFI values greater than 0.9 also indicate a good fit between the model and the data. Lastly, a RMSEA value of less than 0.05 indicates a good model fit (Wang et al., 2001).

Model Modification

When modifying any model to attempt to improve the fit between the data and the model, several steps were followed. First, the strength of all indicators was determined. Indicators that did not materially contribute to the strength of the model (basically, those found not to be significant) were removed to produce a more parsimonious model. Next the errors of each indicator were checked to determine whether any were correlated. Representing correlation between highly correlated errors improved the model fit.

Lastly, as the point of such a study was to determine the effectiveness of sanctions and not simply to confirm a specific model, others were tested as predictors for Recidivism for the entire data set as well as for subgroups like drug offenders, age group and gender. The additional sanctions were boot camps, community service, restitution, and anger management. The results of these tests are discussed in the exploratory section.

Chi-Square Test

Most of the hypotheses were tested through the use of structural equation modeling, but some were not suitable for this method. The intent was to develop measurement and predictive models and use confirmatory factor analysis to conduct hypothesis testing between a latent

construct of Recidivism and the constructs of Offender Risk Propensity and Multiplicity of Sanctions. If, for example, the overall measurement models for Recidivism and/or Multiplicity of Sanctions were not found to have good fit with the data, specific indicators of the construct, such as reoffended (y/n) or number of new charges, could have been tested separately against characteristics such as gender, specific sanctions, and age. In this case, a different measure of Recidivism, such as the number of new charges or crime level change, served as the dependent variable, and an indicator of Multiplicity of Sanctions served as the independent variable. The specific tests were based on the type of variables used as the dependent and independent variable, as are listed in Table 2 in Appendix A. The tests are described next.

Chi-square test can be used to determine whether or not two categorical variables are related. For example, are males more likely than females to be re-arrested after participation in a program? Gender served as the categorical independent variable and recharged (yes/no) served as a categorical dependent variable. The null hypothesis for this type of test is that the groups are independent, and not until the null is rejected can it be concluded that the groups are related (Spatz, 2005).

In evaluating the test results in SPSS, there are two important statistics: the Pearson chi-square value and the significance. If each variable had only two categories (such as male/female for gender and yes/no for recharged), then a two by two table was prepared by SPSS. In this case the Continuity Correlation value presented is the Yates Continuity Correlation, “which compensates for the overestimate of the chi-square value when used in a 2 by 2 table” (Pallant, 2005, p. 290). Using 95% confidence, if the significance reported (presented as Asymp Sig) is less than or equal to 0.05 (alpha), the conclusions are significant (Pallant, 2005).

This test has an additional assumption, the minimum cell frequency, which has to be checked. It is assumed that there is a frequency in each cell of five or more. It is acceptable if at least 80% of the cells have a frequency of 5 or more (Pallant, 2005). A check of this assumption is noted with the results of any chi-square test used in this study.

Structural equation modeling offers researchers the ability to test many predictors of a latent construct simultaneously. In this study, however, the hypotheses are tested separately because many used different subsets of the data. For example, letters of apology were tested as a predictor of recidivism for females while drug testing was tested as a predictor of recidivism for drug offenders, because that is what the theories and literature suggested. Having both drug testing and letters of apology as predictors for both data subsets would have been inappropriate because there was no reason to suspect, for example, that the assignment of letters of apology would have any influence on the recidivism for drug offenders. Because of the non-overlapping data subsets used to test the individual hypotheses, each hypothesis is tested and discussed separately.

Mann-Whitney U Test

The Mann-Whitney U test is a nonparametric test used to test the differences between two independent groups when a continuous measure dependent variable is used. For example, do males and females (two independent groups) have different recidivism when Recidivism is defined as the number of new charges (a continuous measure)? The Mann-Whitney U test converts the continuous measure dependent variable to a ranking and compares the rankings in the two groups to discover whether they are statistically different. The null hypothesis for this

type of test, just as with the chi-square test, is that the groups are independent, and not until the null is rejected can it be concluded that the groups are related (Spatz, 2005).

In evaluating the results of the Mann-Whitney U test in SPSS, two important statistics are presented: Z (a measure of correlation) and significance (p). Using a confidence level of 95%, if p is less than or equal to 0.05, the results are significant. In tests with samples larger than 30, SPSS provides the Z-approximation, which is the measure of correlation (Pallant, 2005).

Hypothesis Testing

This section discusses how each of the specific hypotheses was tested, matching each hypothesis with an appropriate test or tests. The results will be discussed in the next chapter.

Hypothesis 1: Recidivism Measurement Model

The first hypothesis was that measuring recidivism as a latent construct (using the indicators of reoffended (y/n), severity of new charge, most severe sanction imposed, days charge free, total number of new charges, and change in crime level between the original violation and any new violation of law, if one exists) explains the variation in the construct of Recidivism. The null hypothesis was that the model did not accurately conceptualize the construct of Recidivism. The alternative hypothesis was that the model accurately conceptualized the construct. This was tested through confirmatory factor analysis, error correlation, and model testing against the data. Non-significant indicators were removed, correlated errors were added to the model, and the model was evaluated for fit with the data. Model fit was evaluated through the model evaluation steps described above (p -value, chi-square, RMSEA, GFI, AGFI and CFI). In order to simplify the discussion, for the remainder of this paper these steps will be referred to as

evaluating the model. If the indicators were significant and the mode was significant, the null was rejected and the model was used in the other tests.

Hypothesis 2: The Data Fit Common Conventions About Recidivism

Before tests of hypotheses that compared the relationship between Recidivism and other constructs or indicators, the data in this study were examined against several commonly held conventions about recidivism. Recidivism studies have repeatedly concluded that males reoffend more often than females; that offenders charged with their first crime at a younger age reoffend more often than older offenders; and that those with prior histories of deviancy are more likely to be charged with a crime than those without such histories. Those three statements were tested against the data examined for this study, to determine whether this group of offenders behaved similarly to those examined in other studies. Thus, Hypothesis 2 states: The following conventional recidivism findings about youthful offenders apply in particular to participants in restorative justice programs. Those conventions are that males have higher recidivism than females; offenders who offend at younger ages (15 years of age or younger) have higher recidivism than those who offend later in life (after 15 years of age); and offenders with a history of offending (measured by suspension from school) have higher recidivism than do those without a history of offending.

In these hypotheses, the null hypothesis was that there was no difference between males and females, younger (less than 15 years of age) and older (at least 15 years of age), and between those with a history of school suspension and those without a history of suspension when it comes to Recidivism. The alternative hypothesis was that there is a significant difference between these groups when it comes to Recidivism.

The first step in testing this hypothesis was to evaluate the model seen in Figure 16 in Appendix B. Then the relationship between the three categorical variables of Offender Risk Propensity was tested separately against the categorical and continuous indicators of Recidivism, using the test described in Table 2 in Appendix A (with a confidence of 95%). It was expected that the conventions of Recidivism hold true for a majority of these tests.

Hypothesis 3: Completion of Restorative Justice Program and Recidivism

Hypothesis 3—that youth who complete a restorative justice program have significantly lower recidivism than do those who opt not to participate and those removed from the program—tested the relationship between completion of NRJP and Recidivism. The null hypothesis was that completion of NRJP did not have a significant impact on reducing Recidivism. The alternative hypothesis was that those who complete NRJP reoffend less than those who opt not to participate and those who were removed from the program. This was tested in two ways: by evaluating the predictive model (seen in Figure 9 in Appendix B) and with the appropriate tests listed in Table 2 in Appendix A. It was expected that offenders who complete the NRJP reoffend less (less often, less severely, are sanctioned less severely, and take longer to reoffend if they reoffend at all) than offenders who failed to complete or chose not to participate in the program; i.e., the indicator of program completion is a strong predictor of the latent construct of Recidivism, as displayed in the model.

Hypothesis 4: Drug Testing and Recidivism

The fourth hypothesis was that drug testing is an efficient sanction for offenders with a drug history, which was defined as having a drug-related charge or admitting drug use.

Hypothesis 4A states that drug offenders sanctioned to undergo random drug tests in restorative justice programs have lower recidivism than do drug offenders not assigned drug testing.

Hypothesis 4B states: for non-drug-offenders, there is no significant relationship between the assignment of drug testing and recidivism. It was expected that drug offenders who were assigned random drug testing would have lower Recidivism than drug offenders who were not assigned random drug testing. The null hypothesis was that there is no difference in Recidivism between drug offenders who were assigned random drug testing and drug offenders who were not assigned drug testing. The alternative hypothesis was that there is a difference in Recidivism between drug offenders who were assigned random drug testing and drug offenders who were not assigned drug testing. This test was repeated for non-drug offenders. It was expected that for non-drug offenders, the assignment of drug testing does not significantly impact Recidivism. This hypothesis was first tested by evaluating the model (seen in Figure 17 in Appendix B) against drug and non-drug offenders. Then the specific tests in Table 2 in Appendix A were completed. It was hypothesized that assigning drug testing to non-drug offenders (as defined above) has no significant impact on Recidivism (following the same testing procedure as for the previous hypothesis).

Many programs, including the NRJP, assign random drug testing to a majority, if not all, of offenders regardless of drug offender status (and regardless of whether or not actual testing is ever done). This practice allows sanction coordinators who suspect drug usage to require testing at any time. If the assignment of drug testing to non-drug offenders was not found to affect Recidivism, these programs might still assign drug testing for that reason, but might choose to avoid the cost of actually testing most non-drug offenders unless actual drug use was suspected.

Hypothesis 5: Curfews and Recidivism

Next the relationship between the assignment of curfews and Recidivism was tested. Hypothesis 5A states: There is no significant relationship between the assignment of curfews as a sanction for all offenders and Recidivism. Hypothesis 5B states: The assignment of curfews reduces Recidivism for male offenders at least 15 years of age. The null hypothesis was that there is no significant difference in Recidivism between offenders who are assigned curfews and offenders who are not assigned curfews, regardless of their subgroup. The alternative hypothesis was that there is a significant difference in Recidivism between offenders who are assigned curfews and offenders who are not assigned curfews. The steps described above for testing the relationship between drug testing and Recidivism were repeated for curfews as it applies to the age of offenders (those under 15 years of age and those at least 15 years of age) and to gender (see Table 2 and Figures 18 in appendixes A and B, respectively). It was expected that curfews have no significant relationship to recidivism for all offenders in the data set. The literature and theories discussed above suggest, however, that if there is a significant relationship, it is with older male offenders.

Hypothesis 6: No-Contact Orders and Recidivism

Next the relationship between the assignment of no-contact orders and recidivism was tested. Hypothesis 6 states: No-contact orders are an effective sanction for offenders who were charged with a codefendant, for drug offenders, and for males at least 15 years old. The null hypothesis was that there is no significant difference in recidivism between offenders who are assigned no-contact orders and offenders who are not assigned no-contact orders, regardless of their subgroup. The alternative hypothesis was that there is a difference in recidivism between

offenders who are assigned no-contact orders and offenders who are not assigned no-contact orders. The steps described above for testing the relationship between the assignment of curfews and recidivism were repeated for no-contact orders as it applies to the age of offender, gender, drug offenders, and those charged with a codefendant (see Table 2 and Figure 18 in appendixes A and B, respectively). It was anticipated that the assignment of no-contact orders has no significant impact on recidivism overall, but may for older male drug offenders.

Hypothesis 7: Letters of Apology and Essays on Recidivism

Next the relationship between the assignment of letters of apology and essays and recidivism was tested. Hypothesis 7 states: Letters of apology and essays are effective sanctions for females in reducing recidivism. The null hypothesis was that there is no significant difference in recidivism between offenders who are assigned letters of apology and essays and offenders who are not assigned letters of apology and essays, regardless of their subgroup. The alternative hypothesis was that there is a difference in recidivism between offenders who are assigned letters of apology and essays and offenders who are not assigned letters of apology and essays.

The steps described above for testing the relationship between the assignment of curfews and recidivism were repeated in a similar manner for no-contact orders as it applies to the age of offenders (offenders under 15 years of age and offenders at least 15 years of age) and to gender. The specific, non-SEM testing methods for this hypothesis are listed in Table 2 in Appendix A. It was anticipated that letters of apology and essays have no significant impact on recidivism overall, but may for older female offenders.

Hypothesis 8: Over-All Main, Moderating, and Mediating Models

Lastly, the main, moderating, and mediating models, which combined the constructs of Recidivism, Multiplicity of Sanctions, and Offender Risk Propensity into predictive models, were evaluated against the data set, as seen in Figures 13–15 in Appendix B. Hypothesis 8 states: The relationship between Offender Risk Propensity (a latent construct with indicators of race, age, gender, crime type, and drug and school discipline history), Multiplicity of Sanctions (the total package of sanctions assigned), and Recidivism is significant. The null hypothesis was that these models are not significant. The alternative hypothesis was that the models are significant.

The models were tested through the model evaluation steps outlined above. It was expected that these overall models are significant. These models were tested and the results of testing these models are discussed in the next chapter. Reliability, validity, and power must first be addressed and are discussed next.

Measurement Reliability and Validity of Study Variables:

Reliability and validity are crucial aspects of any study. This section examines those concepts, defines the terms and outlines the steps taken to address these issues in the current study. The discussion begins with validity.

Validity

Model validation, an important step, includes examining the validity and the reliability of any measurement processes. Reliability as basically defined is the ability of a measurement method to produce consistent results. The bathroom scale is reliable if it consistently gives the person standing on it about the same reading. Validity is the ability of an instrument to provide

an accurate measure. The bathroom scale is reliable if it consistently provides a reading of 100 lbs., but would be invalid if the person standing on it weighed 200 lbs. This section examines various types of validity and reliability that is important to understand in order to evaluate the model presented in this paper.

Face Validity

First, face validity was considered. Face validity is presumed if a measurement system reasonably appears to measure what it intends to measure. Gliner and Morgan (2000) likened face validity to a professor who chooses a textbook based on its table of contents, because it appears to cover the important topics. Although face validity is considered the weakest form of validity, the model presented in this study appears to have face validity. The constructs and the measurement coding all have face validity since they all appear to measure the construct in a reasonable way.

Content Validity

The next criterion, content validity, assesses whether the components of an instrument is actually representative of the concept under examination. Like face validity, this is a judgment call, and “no statistic demonstrates content validity” (Gliner & Morgan, 2000, p. 320). The process of assessing it includes the definition of the concept measured, review of the literature, and generation of the measurement concept.

For example, the latent construct of Recidivism could have been measured as a function only of re-arrest. It is proposed that the model presented develops the construct better. Its

indicators, at least as they relate to youth criminal sanctioning, do appear to measure the latent construct, giving the model content validity.

Predictive Validity

Next, researchers consider how well a measurement instrument or model can be used to make predictions. Predictive validity is criterion related; it is a measure's ability to predict something it should be able to predict. This assessment takes a fully developed and refined model, in this case including the individual characteristics of an offender as well as the sanctions received, and compares a predicted recidivism rate with the offender's actual recidivism. The closer the predicted recidivism rate is to the actual recidivism rate, the stronger the predictive validity the model can be said to have.

Reliability Measures

Reliability is also a crucial aspect of any measurement instrument. According to Crombach (1960) as cited in Gliner and Morgan (2000), reliability "always refers to consistency throughout a series of measurements." (p. 311). If an outcome measure is not reliable the study's results cannot be relied on. This section addresses reliability.

Many forms of reliability do not apply directly to this study. Inter-rater reliability (the degree to which different raters give consistent measurements), test-retest reliability (the degree to which the same group of people give consistent measurements at different times), and parallel forms reliability (the degree to which the consistency of the results of two tests constructed in the same way from the same content domain give consistent measurements)—none of those apply to this study, as there are no raters, no retesting, and no multiple forms of an instrument.

Power

In statistics and probability, power is defined as $\text{Power} = 1 - \beta$, where β is the probability of not making a type II error. More simply stated β is the probability of not rejecting the null hypothesis when the null hypothesis is actually false. “The more power the analysis [has], the more likely you are to detect a false null hypothesis” (Spatz, 2005, p. 218). The power of an analysis can be influenced by a variety of factors, including effect size, standard error of a difference, sample size, sample variability, and confidence or alpha (α) (Spatz, 2005).

The greater the effect size the greater likelihood that H_0 will be rejected. The larger the difference, however measured, between those who received a treatment compared to those who did not (or who received placebo), the higher the likelihood of rejecting the null (Spatz, 2005). Researchers are seeking to determine if a large effect size actually exists.

Sample size is a critical component of power analysis. The larger the sample, the lower the standard error of the differences will be. The lower the standard error of the differences, the more likely it is that H_0 will be rejected (Spatz, 2005). In determining the appropriate sample size for structural equation modeling, the rule of thumb is that for every variable to be measured there should be five to ten individuals (the unit of analysis). With twenty variables in the largest prediction model, a sample of between 100 and 200 would have been acceptable. The data examined were 221 cases of offenders who participated or were invited to participate in the NRJP during a 26-month period from January 2002 through February 2004. For some tests, slightly smaller samples were used to ensure there were no missing data, but every SEM test had a sample with an appropriate sample size. The additional sample size adds to the power of this analysis. Standard error can also be reduced by reducing sample variability through “using

reliable measurement instruments, recording data correctly, being consistent, and in short, reducing the ‘noise’ or random error in your experiment” (Spatz, 2005, p. 218).

Lastly, power is affected by how confident the researcher seeks to be in any conclusions. The probability of making a Type I error (i.e., rejecting a true null hypothesis) is denoted by α . The larger α , the more likely it is that H_0 will be rejected. The conventional although sometimes arbitrary limit for alpha is 0.05 (Spatz, 2005). For this analysis, alpha levels of 0.05 or smaller were used.

According to Kaplan (2011),

In the framework of structural equation modeling the assessment of power is complicated. Unlike simple procedures such as the t -test or ANOVA wherein alternative hypotheses pertain to only a few parameters, in structural equation modeling there are considerably more parameters. Each fixed parameter in the model is potentially false and each can take on, in principle, an infinite number of alternative values. Thus, each fixed parameter needs to be evaluated, in principle, one at a time. (para. 2)

Now that the concepts of power, reliability, and validity have been introduced and their concepts applied to the process of hypothesis testing, the next step is to present the hypothesis testing. The results of that testing appear in the next chapter.

Addressing Missing Data

This study’s data set, like those in most studies, was not perfect. There were some missing data, and for structural equation modeling missing data must be addressed before models can be tested. The missing data here fell into three categories: 1) missing data that could be accounted for through calculation, 2) missing data that required removal of some cases from the data set, and 3) missing data that affected only certain tests. This section explains the approach to missing data.

The first group of missing data comprised data that could be calculated. The age of offender 169 was missing. Her abridged date of birth was in August 1987 (exact date redacted from the data set to reduce the probability of identification), and the date of her initial charge was 21 April 2003, which made her 15 years old at the time of the charge. The missing date of birth was replaced with 15 in the data set.

The second set of missing data required removal of several offenders from the data set. For offender 146, gender and race were missing. For offender 85 the number of new charges and the disposition were missing. He had been charged with at least one crime (domestic battery) as indicated by the notes. Offender 58 had been recharged at least once with robbery, but the total number of new charges and the most severe sanction assigned were missing. For these reasons, these three offenders were removed from the data set, leaving a data set of 218 offenders.

The third group of missing data affected only certain tests. It was specific to only two variables: admitted drug use and whether a codefendant had also been charged. Four offenders' data (75, 99, 100, and 201) were missing the variable 'admitted drug use.' None of the four offenders had prior, current, or subsequent drug-related charges, and none had notes indicating a failed drug test. For those reasons and because answering no to admitted drug use was by far the favored mode, these missing data were replaced with no, coded as zero.

For offenders 19, 37, 146, and 169 the variable that indicated whether a codefendant had also been charged was missing. Since this variable was to be used in only one hypothesis, the missing data were addressed specifically (by removal) at that point. As indicated above, offender 146 had already been removed because of missing critical data. With the missing data addressed, descriptive statistics could be calculated and are discussed in the next section.

Descriptive Statistics

Prior to conducting hypothesis testing, descriptive statistics were prepared and are discussed in this section. Following the descriptive statistics, the hypotheses will be tested.

Descriptive Statistics Related to Hypothesis 1

The first indicator of Recidivism examined is the indicator Reoffend (y/n). This variable was coded as 1 if the offender had received an additional charge, regardless of the severity, timing, or disposition of that charge. If the offender had not received an additional charge, Reoffend was coded as 0. Of the 218 cases examined for the recidivism model, 146 (67.0%) had not received a new charge during the period examined. Of that same group, 72 (33.0%) had received an additional charge after starting or declining to participate in the program. This was the Recidivism rate not for the NRJP but for the entire data set. The Recidivism of those who had completed the program is discussed when addressing Hypothesis 3, below.

Next, the severity of any new charge was examined. The variable Severity of New Charge was coded as defined in Table 3 in Appendix A. Of the 72 who had received a new charge, 26 (36.1 % of those who had reoffended) had misdemeanors as their most severe charge while 46 (63.9% of those who had reoffended) had felonies as their worst new charge. Of the entire sample, 67.0% had not received a new charge, 11.9% had received misdemeanor charges, and 21.1% had received felony charges (see Figure 22 in Appendix B). Figure 22 shows that although a majority of the sample had not received an additional charge, those who had were charged with felonies at a rate almost twice that of those charged with only misdemeanors. For a complete breakdown of the severity of new charges, see Table 4 in Appendix A.

It could be argued that the severity of any subsequent charge is relative to the initial charge. The next indicator of Recidivism is Crime Level Change. This variable was calculated by subtracting the level of the initial charge from the level of any new charge. The data set comprised nine categories ranging from -5 through 3. The most frequently occurring Crime Level Change for the entire data set examined was -2, which was observed 66 times (30.7%). Level -2 would occur if, for example, the initial charge had been a second degree misdemeanor (e.g. retail theft less than \$100 in value, or assault) and there had been no new charge ($0 - 2 = -2$). For a complete breakdown of the levels, their frequencies, and percentages see Table 5 in Appendix A.

The next indicator of Recidivism examined is Number of New Charges. This indicator simply counted the frequency of new charges after the initial violation. The Number of New Charges for the entire data set ranged from 0 (the mode) to 125, with a mean of just over 3 new offenses per person for the entire sample. However, the 72 who had reoffended had received on average just over 9 new charges each. As a group, they had received 662 new charges, or 303.7 new offenses per 100 offenders. The worst ten offenders had received 433 new charges, which accounted for 65.4% of the new charges received by the group. Of the 218 in the entire sample, 51 (23.4%) had received more than one new charge. See Table 6 in Appendix A for offense number, frequencies, and percentages.

The next indicator of Recidivism is Most Severe Sanction, which was coded as per Table 3 in Appendix A. This variable ranged from no new sanction (the mode) to adult prison sentence and was coded as per Table 3 in Appendix A. Of the 72 charged by police, 6 were non-file, indicating the state had declined to prosecute, and 6 were Nolle pros, indicating the state had

initially pursued charges but later had withdrawn or dropped them. These two groups, coded as 1 and 2, respectively, accounted for 5.5% of the entire group and 16.6% of those who had received an additional charge. Of those who had received an additional charge, probation was the most frequently observed disposition, comprising 15.1% of the total sample and 45.8% of those who received an additional charge. The most severe sanctions observed in the sample were adult jail and adult prison, which were seen in 0.9% and 1.4% of the entire sample and in 2.8% and 4.2%, respectively, of those who had received additional charges. For a complete breakdown of sanction frequencies and percentages, see Table 7 in Appendix A.

The last indicator of Recidivism examined is Days Charge Free, a calculation of the days between the conference and the first new charge. In cases where there was no new charge, the date of the criminal history checks (15 January 2009) was used. This approach produced methodological problems that are addressed later. This variable ranged from 15 to 2,544 days (about 7.0 years), with a mean of 1,572.7 (about 3.2 years) and a standard deviation of 845.3.

The program lasted about 90 days, yet 14 offenders had received new charges even before it was completed. Almost all of them had subsequently been charged with additional crimes and almost all of those new charges had been felonies. Those 14 included 2 who ultimately had been sentenced to adult prison, 3 sentenced to high security detention, 1 sentenced to medium security detention, 1 sentenced to low security detention, and 6 sentenced to probation. For the 14th offender, the most severe sanction assigned was listed only as detained. Those 14 offenders had been charged with a total of 284 new charges, an average of 20.3 new charges per offender. At first glance it would appear that those who had reoffended soonest had

also reoffended most severely, but that observation was tested to determine if it is significant, as described later.

Descriptive Statistics Related to Hypothesis 2

Gender and Recidivism

Of the 218 offenders in the data set, 133 (61.0%) were males and 85 (39.0%) were females. For Reoffended (y/n), the traditional recidivism rates (defined solely as having received an additional charge) for males and females was 36.8% and 27.1% respectively. For Severity of New Charges, males had reoffended 17.2% more often than females had, when recidivism is defined as receiving a new felony charge (see Table 12 in Appendix A). For Most Severe Sanction, when recidivism was defined as a new charge resulting in a sanction of at least probation, the recidivism rates for males and females were 27.1% and 18.8%, respectively (see Table 13 in Appendix A). For Crime Level Change, males had reoffended more severely than their original charge 9.0% more often than females had (see Table 14 in Appendix A). For Number of New Charges, the mean is 4.5 per male and only 0.8 per female, and for frequency of more than 5 new charges, males had 23 (17.3%) and females had 4 (4.7%). Lastly, for Days Charge Free, males had gone an average of 1,513.7 days (about 4.2 years) before a new charge had been received (standard deviation of 878.4), and for females the mean is 1,665.1 (about 4.6 years) (standard deviation of 786.80) (see Table 15 in Appendix A).

Whether or not these differences are statistically significant is discussed in the hypothesis testing section. Now that the descriptive statistics for gender have been examined, the focus shifts to age differences.

Age Group and Recidivism

There were two age groups: younger offenders (less than 15 years old when initially charged (n = 84)) and older offenders (at least 15 years of age when initially charged (n = 134)). The traditional recidivism rates for younger and older offenders are 48.8% and 23.1%, respectively. For Severity of New Charge, when recidivism is defined as receiving a new felony charge, the Recidivism rates for younger and older offenders are 34.5% and 12.7%, respectively (see Table 12 in Appendix A). For Most Severe Sanction, 16.7% of younger offenders and 3.7% of older offenders had eventually been sanctioned to some form of incarceration (see Table 13 in Appendix A).

For Crime Level Change, 39.3% of younger offenders and 14.9% of older offenders had been charged with new charges more severe than their original offenses (see Table 14 in Appendix A). The mean Number of New Charges is 6.5 (standard deviation of 17.0) for younger offenders and 0.9 (standard deviation of 2.3) for older offenders (see Table 15 in Appendix A). For Days Charge Free, younger offenders have a mean of 1,353.2 (about 3.7 years) with a standard deviation of 888.0; and older offenders have a mean of 1,710.4 (about 7.0 years) with a standard deviation of 790.0 (see Table 16 in Appendix A). The differences based on age group were examined to determine whether they are statistically significant, as discussed in the hypothesis testing section.

Suspension and Recidivism

The last predictor related to Hypothesis 2 is history of suspension from school. It had been suggested that a history of suspension would be seen to have a positive relationship with

recidivism. This section examines the descriptive statistics of school suspension as it pertains to the six indicators of the construct of Recidivism.

For the indicator Reoffended (y/n), of the 68 offenders who had never been suspended, 9 (13.2%) had received new criminal charges. Of 150 offenders with histories of suspension, 63 (42.0%) had received new criminal charges. For Severity of New Charge, of those who were never suspended, 3 (4.4%) reoffended and received a new misdemeanor charge, and 6 (8.8%) had reoffended and received at least one felony charge. Of the 150 with histories of suspension, 23 (15.3%) reoffended and received new misdemeanor charges, and 40 (26.7%) had reoffended and received at least one new felony charge (see Table 12 in Appendix A).

For Most Severe Sanction, 1.5% of those never suspended and 12.0% of those with a suspension history had eventually been sanctioned to some form of incarceration; 30.0% of those suspended and 10.3% of those never suspended had reoffended and been sanctioned to at least probation (see Table 13 in Appendix A). For Crime Level Change, 8.8% of those never suspended and 31.3% of those suspended had reoffended at more severe levels than that of their initial charges (see Table 14 in Appendix A). For Number of New Charges, those not suspended had a mean of 0.5 with a standard deviation of 1.5. Offenders with suspension histories had received an average of 4.2 new charges per offender with a standard deviation of 13.1 (see Table 15 in Appendix A). For Days Charge Free, those not suspended had a mean of 1,908.3 (about 5.2 years) with a standard deviation of 645.7. Offenders with suspension histories had gone an average of 1,420.6 days (about 3.9 years) before receiving a new charge (standard deviation of 882.3) (see Table 16 in Appendix A). All the differences discussed in this section from descriptive statistics were examined to determine whether they are statistically significant.

Descriptive Statistics Related to Hypothesis 3

All indicators of Recidivism show higher Recidivism for those who did not complete the program than for those who did. For Reoffend (y/n), of those who did not complete the program (n = 36), 80.6% were charged with new crimes. Of those who completed the program (n = 182), only 23.6% received new charges. For Severity of New Charge, 58.3% of those who did not complete and 13.7% of those who did complete the program received new felony charges (see Table 22 in Appendix A). For Most Severe Sanction, 75.0% of those who did not complete and 13.7% of those who did complete reoffended and were sanctioned to probation or more. For those who did not complete and those who did, the rates of those who had reoffended and sanctioned to some form of incarceration are 33.3% and 3.8%, respectively (see Table 23 in Appendix A). For Crime Level Change, 63.9% of those who did not complete and 16.5% of those who did complete the program reoffended at higher levels than their initial charges (see Table 24 in Appendix A). For Number of New Charges, of those who did not complete the program, 41.7% received more than 5 new criminal charges and only 6.6% of those who completed it received more than 5 new criminal charges (see Table 25 in Appendix A). Lastly, for Days Charge Free, the mean for those who did not complete the program was 621.9 days (1.7 years) with a standard deviation of 811.9, and the mean for those who did complete it was 1,760.8 days (4.8 years) with a standard deviation of 716.7 (see Table 26 in Appendix A). Despite showing drastic recidivism differences based on completion status, it remained to be determined whether the differences are statistically significant. That is described in the next section.

Descriptive Statistics Related to Hypothesis 4

Of the 218 youths in the data set, 54 (24.8%) were drug offenders. Drug offenders comprised two not mutually exclusive groups, those with a drug-related charge, 27 of the 218 (12.4%), and those who had admitted drug use, 41 of the 218 (18.8%). It was established for Hypothesis 3 that completion of the restorative justice program is significantly associated with a reduction in recidivism. Of the 54 drug offenders, 12 (22.2%) had not completed the program, and of these 12, 8 (66.7%) had reoffended. This section compares Recidivism indicators for those drug offenders assigned and those not assigned drug testing.

Of the drug offenders ($n = 54$), 46 were assigned drug testing and only 8 were not. The recidivism rates as defined by receiving a new criminal charge are 29.6% for all drug offenders and 50.0% and 26.1% for those not assigned and for those assigned drug testing, respectively. In addition to a higher rate of receiving new charges for those not assigned drug testing, they also reoffended more severely. Of those who reoffended, 25.0% of those not assigned drug testing and 66.7% of those assigned drug testing were charged with at least one felony (see Table 30 in Appendix A).

Comparison for Most Severe Sanction revealed that 37.5% of drug offenders who were not assigned drug testing had eventually been sanctioned to at least probation, compared to 21.7% of the drug offenders who were assigned drug testing. Further, 12.5% of drug offenders who were not assigned drug testing had eventually been sanctioned to some form of incarceration, compared to 4.3% for drug offenders who were assigned drug testing (see Table 31 in Appendix A). For Crime Level Change, 12.5% of drug offenders who were not assigned drug testing and 6.5% of drug offenders who were assigned drug testing reoffended at the same

levels as their original charges and 25.0% of drug offenders who were not assigned drug testing and 19.6% of drug offenders who were assigned drug testing reoffended at more severe levels than those of their original charge (see Table 32 in Appendix A).

For Number of New Charges, 12.5% of drug offenders who were not assigned drug testing and 17.4% of drug offenders who were assigned drug testing had reoffended and received at least five new charges (see Table 33 in Appendix A). For Days Charge Free, drug offenders who were not assigned drug testing had a mean of 1,324.0 days (3.6 years) with a standard deviation of 828.0, and drug offenders who were assigned drug testing had a mean of 1,652.6 days (4.5 years) with a standard deviation of 803.9 (see Table 34 in Appendix A).

The drug offenders not assigned drug testing had reoffended at a higher rate, received more severe sanctions, and reoffended on average sooner than their counterparts who were assigned drug testing. Drug offenders who were assigned drug testing had reoffended at a more severe level as compared to those who were not assigned drug testing (without controlling for severity of the charge). The next section describes the test of whether differences are significant.

Descriptive Statistics Related to Hypothesis 5

It was noted that with regard to Hypothesis 4 that only 8 of the 54 drug offenders had not been assigned drug testing. That fact makes conclusions from the hypothesis testing somewhat questionable. Originally the intention was to test the mere assignment of a curfew against recidivism, but when the data were examined, only one of the 218 offenders in the data set was found to not have been assigned a curfew. He did not reoffend, but obviously no conclusions can be drawn from that observation. One other offender had been assigned a curfew, but the assigned time was missing from the data set. For these reasons the two cases were eliminated, leaving 216

cases for this analysis. The descriptive statistics for Hypothesis 5 are presented in two parts. First the entire data set is examined; then the analysis of older male offenders is presented.

The assigned curfew times for the data set ($n = 216$) ranged from 12:30 pm to midnight, with a mean of about 7:38 pm and a mode of 7:00 pm. The mean serves as the demarcation line between those assigned an early (before 7:38 pm) and those assigned a late (after 7:38 pm) curfew. A new dummy ordinal variable, Curfew Code, was created and coded as 0 for those assigned an early curfew and 1 for those assigned a late curfew.

An important distinction must be made between the continuous variable Curfew Time (the actual assigned curfew time) and the ordinal variable Curfew Code (a dummy variable indicating whether the assigned time was either early or late as previously defined). The two terms cannot be used interchangeably.

Of the 216 offenders, the rates of receiving new criminal charges for those assigned an early curfew and for those assigned late curfew are 40.5% and 25.7%, respectively. The rates of receiving new felony charges for those assigned an early curfew and those assigned a late curfew are 27.0% and 15.2%, respectively (see Table 36 in Appendix A). For Most Severe Sanction, the rates of receiving some form of incarceration for those assigned an early curfew and for those assigned a late curfew are 12.6% and 4.8%, respectively (see Table 37 in Appendix A). For Crime Level Change, the rates of receiving a new charge more severe than the original charge for those assigned an early curfew and for those assigned a late curfew are 35.1% and 17.1%, respectively (see Table 38 in Appendix A). For Number of New Charges, the rates of receiving more than five new charges for those assigned an early curfew and for those assigned a late curfew are 17.1% and 7.6%, respectively (see Table 39 in Appendix A). For Days Charge Free,

the means for those assigned an early curfew and those assigned a late curfew are 1,452.3 days (about 4.0 years) with a standard deviation of 849.7 and 1,688.1 (about 4.6 years) with a standard deviation of 831.1, respectively. Those assigned an early curfew reoffended more often, more severely, at a greater frequency, and sooner than those assigned a later curfew. Later in the analysis these differences were examined to determine whether they are statistically significant.

Offenders charged with their first crime at earlier ages typically have greater recidivism than do those first charged at an older age. Hypothesis 2 confirmed that those who receive their first charge before the age of 15 have significantly higher recidivism than do those who receive their first charge after the age of 15. It is possible that in the data set younger offenders were simply assigned earlier curfews than the older offenders were, which would cause it to appear that assigned curfew time has a strong impact on recidivism, but correlation does not mean causation. This possibility required further investigation.

Those assigned an early curfew had an age range from between 9 and 17 and a mean age of 14.1 with a standard deviation 1.8. Those assigned a late curfew had an age range from 11 to 17 and a mean age of 15.5 with a standard deviation of 1.4. To test whether differences are significant, a new dummy ordinal variable (Curfew Code) was created and coded as 0 for those assigned a curfew before 7:38 pm (n = 111) and 1 for those assigned a curfew after 7:38 pm (n = 105). The Mann-Whitney U test was conducted to test a null hypothesis that the distribution of age was the same across the category of Curfew Code. The test has a significance of 0.000 and suggests rejecting the null hypothesis at the 0.05 level. This test demonstrated a statistically significant difference between younger and older offenders when it comes to their assignment of

either an early or a late curfew. This difference is discussed and controlled for later in the analysis.

When older male offenders (at least 15 years of age) were examined separately, the differences became less apparent. Of the 79 older male offenders, the rates of receiving new criminal charges for those assigned an early curfew (n = 29) and those assigned a late curfew (n = 50) are 27.6% and 26.0%, respectively. The rates of receiving new felony charges for those assigned an early curfew and for those assigned a late curfew are 13.8% and 22.0%, respectively (see Table 36 in Appendix A). For Most Severe Sanction, the rates of receiving some form of incarceration for those assigned an early curfew and for those assigned a late curfew are 0.0% and 4.0%, respectively (see Table 37 in Appendix A). For Crime Level Change, the rates of receiving new charges more severe than the original charge for those assigned an early curfew and for those assigned a late curfew was 13.8% and 20.0%, respectively (see Table 38 in Appendix A). For Number of New Charges, the rates of receiving more than five new charges for those assigned an early curfew and for those assigned a late curfew are 7.6% and 12.0%, respectively (see Table 39 in Appendix A). For Days Charge Free, the means for those assigned an early curfew and those assigned a late curfew are 1,598.0 days (about 4.4 years) with a standard deviation of 784.8 and 1,667.3 (about 4.6 years) with a standard deviation of 875.9, respectively.

Males at least 15 years of age when they received their initial charge who were assigned an early curfew reoffended on average slightly more often and sooner than their counterparts who were assigned a later curfew. But those of that age group who were assigned an early curfew reoffended less frequently and less severely and received less severe sanctions than did

their counterparts who were assigned a later curfew. These differences are examined later to determine whether they are statistically significant.

Descriptive Statistics Related to Hypothesis 6

Of the 218 cases in the data set, no-contact orders were assigned in 42 cases (19.3%) and not assigned in 176 cases (80.7%). The rates of receiving new charges for those not assigned a no-contact order and for those assigned a no-contact order are 30.7% and 42.9%, respectively. For Severity of New Charge, the rates of receiving new felony charges for those not assigned a no-contact order and for those assigned a no-contact order are 18.2% and 33.3%, respectively (see Table 43 in Appendix A). For Most Severe Sanction, the rates of receiving new charges and some form of incarceration for those who were not assigned a no-contact order and for those assigned a no-contact order are 4.7% and 14.3%, respectively (see Table 44 in Appendix A). For Crime Level Change, the rates of receiving new charges at the same or more severe levels than the original charges for those who were not assigned a no-contact order and for those assigned a no-contact order are 27.8% and 40.5%, respectively (see Table 45 in Appendix A). For Number of New Charges, the rates of receiving at least five new charges for those not assigned a no-contact order and for those assigned a no-contact order are 11.4% and 16.7%, respectively (see Table 46 in Appendix A). For Days Charge Free, the mean for those who were not assigned a no-contact order is 1,619.2 days (4.4 years) with a standard deviation of 824.4, and for those assigned a no-contact order is 1,378.2 (3.8 years) with a standard deviation of 912.5 (see Table 47 in Appendix A).

In this data set, those assigned a no-contact order reoffended more often, sooner, more frequently, and at a more severe level, and were incarcerated at a higher rate than those not

assigned a no-contact order. It is possible that those who associated with other delinquents were simply assigned no-contact orders more often and by their nature were more likely to reoffend. This data set, unfortunately, did not allow for any method to test that possibility. In the next section these differences are examined to determine whether they are significant.

Descriptive Statistics Related to Hypothesis 7

The descriptive analysis began with a comparison of Recidivism indicators between those in the data set who were not assigned or who were assigned apology letters and essays. Of the 218 cases in the data set, 29 (13.3%) were not assigned letters of apology, and 189 (86.7%) were assigned letters of apology. Of the same data set, 52 (23.9%) were not assigned essays and 166 (76.1%) were assigned essays (see Table 50 in Appendix A). This section compares Recidivism indicators between these subgroups.

The rates of receiving a new charge for those not assigned and for those assigned apology letters are 34.5% and 32.8%, respectively. This difference may appear minimal, but for Severity of New Charge the rates of receiving new felony charges for those not assigned and for those assigned apology letters are 10.3% and 22.8%, respectively (see Table 51 in Appendix A). For Most Severe Sanction the rates of receiving some form of incarceration for those not assigned apology letters and for those assigned apology letters are 10.3% and 8.5%, respectively (see Table 52 in Appendix A). For Crime Level Change, the rate of receiving a new charge more severe than the original charge for those not assigned apology letters and assigned apology letters are 10.3% and 26.5% respectively (see Table 53 in Appendix A). For Number of New Charges, the rates of receiving more than five new charges, for those not assigned apology letters and for those assigned apology letters are 6.9% and 13.2%, respectively, and the rate of receiving more

than ten new charges for those not assigned apology letters and for those assigned apology letters are 3.4% and 5.8%, respectively (see Table 54 in Appendix A). For Days Charge Free, the mean is 4.3 years for both those not assigned apology letters and for those assigned apology letters (see Table 55 in Appendix A).

Those not assigned letters of apology appear to have reoffended slightly more often and to have been incarcerated at a higher rate than were those assigned letters of apology. Those not assigned letters of apology appear to have reoffended less often at the felony level and reoffended less frequently than did those assigned apology letters. In the next section, these differences are examined to determine whether they are statistically significant. First the descriptive statistics of Recidivism indicators based on assignment of essays are discussed.

The rates of receiving new charges for those not assigned essays and for those assigned essays are 34.6% and 32.5%, respectively. For Severity of New Charge, the rates of receiving a new felony charges for those not assigned essays and for those assigned essays are 26.9% and 19.3%, respectively (see Table 51 in Appendix A). For Most Severe Sanction, the rates of receiving some form of incarceration for those not assigned essays and those assigned essays are 13.7% and 7.2%, respectively (see Table 52 in Appendix A). For Crime Level Change, the rates of receiving new charges more severe than the original charge for those not assigned essays and those assigned essays are 26.9% and 23.5%, respectively (see Table 53 in Appendix A). For Number of New Charges, the rate of receiving more than five new charges is 10.8% for both those not assigned essays and those assigned essays, and the rate of receiving more than ten new charges is 4.2% for both those not assigned essays and those assigned essays (see Table 54 in Appendix A). For Days Charge Free, the means for those not assigned essays and those assigned

essays are 1,586.3 (4.4 years) with a standard deviation of 913.6 and 1,568.5 (4.3 years) with a standard deviation of 825.6, respectively (see Table 55 in Appendix A).

Thus for Days Charge Free, Number of New Charges, and Reoffended (y/n) there is basically no difference between those assigned essays and those not assigned essays. Those not assigned essays, however, reoffended at the felony level more often, reoffended at higher levels than the original charges more often, and were incarcerated more often than were those assigned essays. The next section examines these differences to determine whether they are statistically significant and also examines the relationship between the assignment of letters and essays and the Recidivism indicators for a specific group, females at least 15 years old.

Descriptive Statistics Related to Hypothesis 8

The main, moderating, and mediating effect predictive models are basically four measurement models (Recidivism, Offender Risk Propensity, Multiplicity of Sanctions, and the interaction construct) with slightly different relationships between them. The different relationships among these models represent different relationships among the constructs. The first step in evaluating the predictive models was to evaluate the individual measurement models. The measurement model for Recidivism had already been evaluated and after modification had been found to be a good fit with the data. The remaining three measurement models were evaluated next.

The measurement model for Offender Risk Propensity was evaluated first. This construct was hypothesized as a function of six indicators working simultaneously: suspension history, gender, age at initial charge, race, charge type, and drug offender. These variables were coded as indicated in Table 3 in Appendix A.

It was established by testing Hypothesis 2 that age and suspension history are significant predictors of the latent construct of Recidivism for this data set. Younger offenders with histories of suspension were seen to have significantly higher recidivism, and gender was not seen as a significant predictor of recidivism when these variables were tested as predictors of the latent construct Recidivism. This section examines their contribution to a different construct, Offender Risk Propensity. Tests of the relationship between Offender Risk Propensity and Recidivism are presented in the next section.

Just as for Hypothesis 1, the measurement model evaluation began with an analysis of correlation. The indicators of Offender Risk Propensity were analyzed for their correlation (see Table 59 in Appendix A). The highest correlation for indicators of this construct is between drug offender and charge type (0.371), which is significant at the 0.01 level (two-tailed). This finding should not be surprising, since drug offense (a contributor to the indicator drug offender) is one of the charge types. The only other correlation that is significant at the 0.01 level (two-tailed) is between drug offender and age (0.204). The correlations between race and drug offender, gender and charge type, suspended and drug offender, and between suspended and race are -0.162, -0.136, 0.158, and 0.176, respectively. These correlations are significant at the 0.05 level (two-tailed) but are not very strong in magnitude. No other indicators of Offender Risk Propensity were significantly correlated.

When the indicators of the construct Multiplicity of Sanctions were analyzed for correlation, again few were found to be significantly correlated. See Table 60 in Appendix A. The correlations between boot camp and anger management and between community service and restitution are significant at the 0.01 level (two-tailed) and have magnitudes of 0.181 and -0.215,

respectively. The correlations between drug testing and no-contact orders and between community service and drug testing are significant at the 0.05 level (two-tailed) and have magnitudes of 0.154 and -0.173, respectively.

It is important to remember that this was not an experimental design, but rather data collected from youth cases where sanctions were actually assigned to youths as deemed appropriate by the sanctioning body. One might therefore intuitively expect a positive and significant correlation between the assignment of boot camp and anger management, as well as between no-contact orders and drug testing. Youths seen by the sanctioning body to be in need of an attitude adjustment or ones charged with violent crimes may be likely to be assigned both anger management and a boot camp. Drug offenders could also reasonably expect to be sanctioned with drug testing as well as to be assigned no contact with other drug users or drug suppliers. One might also intuitively expect a negative and significant correlation between community service and restitution. It is reasonable to expect that offenders less able to pay restitution or fines would be more likely to be assigned community service.

The last sets of correlations to be examined are those for the interaction construct. To select the indicator combinations, four of the strongest indicators of Offender Risk Propensity (charge type, drug offender, age, and suspension history) were multiplied with four of the strongest indicators from Multiplicity of Sanctions (no-contact orders, drug testing, community service, and essays). This process produced sixteen indicators for the interaction construct (see Table 61 in Appendix A). When these indicators were analyzed for correlations, some correlations were expected. For example, the indicators no-contact and age was highly correlated (0.908) with the indicator no-contact and charge type, and that correlation is significant at the

0.01 level (two-tailed). This correlation should not be surprising, since both indicators contain the assignment of a no-contact order.

Of the 120 possible correlations, 57 are significant at the 0.01 level (two-tailed). Of those, 25 correlations are between indicator combinations where no common indicator exists. Of those 25, the highest correlations are between letters and drug offender and drug testing and age (0.518) and between letters and drug offender and drug testing and suspended. All the remaining correlations that are significant at the 0.01 level (two-tailed) and do not contain common indicators have correlations below 0.4.

For the indicators of Recidivism analyzed for Hypothesis 1 (see Table 8 in Appendix A), the correlations are much higher (ranging from 0.393 to 0.944, with half over 0.8). Also, all the correlations are significant at the 0.01 level (two-tailed). For this reason, the measurement models of Offender Risk Propensity and Multiplicity of Sanction were not expected to be as strong as the measurement model of Recidivism.

Now that the correlations have been examined, the testing of the models and their relationships is presented.

CHAPTER 4: RESULTS

Testing Hypothesis 1

The first hypothesis was that measuring recidivism as a latent construct (using the indicators of reoffended (y/n), severity of new charge, most severe sanction imposed, days charge free, total number of new charges, and change in crime level between the original violation and any new violation of law, if one exists) explains the variation in the construct of Recidivism. The first step in testing this hypothesis was to determine correlations between the indicators. All indicators were highly correlated, with significance at the 0.01 level (two-tailed). This result was expected because all the indicators contribute to the understanding of the latent construct Recidivism. See Table 8 in Appendix A for the Pearson correlation table.

Next, the recidivism measurement model was run in AMOS, with factor weights as shown in Figure 23 in Appendix B. Immediately, the variable Number of New Charges stood out because of a relatively low standardized regression coefficient of 0.468. That indicates that this variable's contribution to the model is modest and that it should be removed or modified in order to improve the model. All indicators' contributions to the model are significant at the 0.001 level (two-tailed), as seen in Table 9 in Appendix A.

Modifications were made to the model in order to reach the strongest and most parsimonious model. First, the indicator Number of New Charges was removed and replaced with a modified form of that variable. The variable Multiple Charges (y/n) was created as an indicator of the latent construct Recidivism. This variable was coded as 0 if the offender had received one new charge or fewer and as 1 if the offender had received 2 or more new charges.

The standardized regression coefficient is 0.468 for Number of New Charges, but is 0.833 for Multiple Charges (y/n). This change also improved the overall fit of the model. See Tables 10 and 11 in Appendix A.

Next, the modification indices were examined. The highest modification index, 52.2, was between e2 and e6. These error terms, between the errors associated with Severity of New Charge and with Crime Level Change, were correlated in the model. After the model was rerun, the next highest modification index was 12.1, between e1 and e5. These error terms, between the errors associated with Reoffended (y/n) and with Multiple Charges (y/n), were correlated in the model. After the model was rerun, the only remaining modification index was 8.1 and was between e1 and e3. These error terms, between the errors associated with Reoffended (y/n) and with Most Severe Sanction were correlated in the model. The steps outlined here for the identification and graphic representation of correlation between error terms were followed for the ensuing models, but are not specifically explained for each model.

This measurement model of recidivism had a chi-square value of 4.2 with 6 degrees of freedom, a relative chi-square of 0.7; GFI, AGFI and CFI of 0.993, 0.977 and 1.000, respectively, RMSEA of 0.000; and a *p*-value of 0.645. See Table 11 in Appendix A for goodness of fit measures of the Recidivism measurement model.

Although this measurement model is a good fit with the data according to all goodness of fit statistics used, one additional modification, though perhaps unnecessary, was to remove the indicator Multiple Charges (y/n) from the model. This indicator, although significant, has the lowest standardized regression coefficient of all the indicators. Once this indicator had been removed, the modification indices were added using the procedure described above. The

resulting measurement model of recidivism had a chi-square value of 0.4 with 3 degrees of freedom; a relative chi-square of 0.1; GFI, AGFI and CFI of 0.999, 0.997 and 1.000, respectively; RMSEA of 0.000; and a *p*-value of 0.949. See Figure 24 and Table 11 in appendixes B and A, respectively, for the regression weights and goodness of fit measures of the revised recidivism measurement model, respectively.

For Hypothesis 1, this analysis supports rejecting the null hypothesis and concluding that the revised measurement model of Recidivism is a good fit with the data. The revised measurement model of Recidivism had a chi-square value of 0.4 with 3 degrees of freedom; a relative chi-square of 0.1; GFI, AGFI and CFI of 0.999, 0.997 and 1.000, respectively; RMSEA of 0.000; and a *p*-value of 0.949.

Testing Hypothesis 2

Hypothesis 2 reformulated conventional recidivism findings about youthful offenders to determine if they apply to participants in restorative justice programs, specifically that males have higher recidivism than females, offenders who offend at younger ages (less than 15 years of age) have higher recidivism than those who offend later in life (at least 15 years of age), and offenders with a history of offending (measured by suspension from school) have higher recidivism than do those without a history of offending. The hypothesis was tested by first determining whether gender, age, and suspension history are significantly correlated.

None of the variables are correlated with each other at any significant level (see Table 17 in Appendix A). Unlike the indicators for Recidivism (which all contribute to the same construct), these predictors were not expected to be highly correlated.

Next, the measurement model for recidivism was made into a predictive model with the addition of gender, age group, and suspension history as exogenous variables. The model seen in Figure 25 in Appendix B was then tested against the data. All the exogenous variables with the exception of gender were found to significantly contribute to the model (see Table 18 in Appendix A for factor weights). This finding was not unexpected, as the gender differences in the descriptive statistics section above are not as drastic as those for age group and suspension history.

This predictive model has a relative chi-square below 5 (2.32); GFI, AGFI and CFI all above 0.9 (0.963, 0.911, and 0.989, respectively); but with *p*-value less than 0.05 (0.003) and RMSEA above 0.05 (0.78). The model was modified by the removal of gender, since it is not significant. The revised model, seen in Figure 26 in Appendix B, is an adequate fit with the data according to all criteria used (*p*-value (0.067), chi-square (18.7), relative chi-square (1.7), GFI (0.977), AGFI (0.940), CFI (0.996) and RMSEA (0.057)). See Table 19 in Appendix A for goodness of fit statistics for the original and revised predictive models of Recidivism based on age group, gender, and suspension history.

Age group and suspension history were both found to be significant predictors of recidivism. The overall revised model is significant and was found to be a good fit with the data. Because gender had been removed from this model, the non-SEM tests as described in Chapter 3 (chi-square and Mann-Whitney U tests) were completed next.

From the SEM analysis above, age group and suspension history were both found to be significant predictors of the latent construct of Recidivism with five indicators. It should be no surprise that for all six indicators of Recidivism, the results of this analysis suggest rejecting the

null hypothesis and concluding that the distribution of these predictors is not the same across the indicators of Recidivism. Many of the chi-square tests violate the assumption that at least 80% of the cells have an expected cell frequency of 5 or more. This issue is addressed later in this section.

Also from the SEM analysis above, gender was not found to be a significant predictor of the latent construct of Recidivism when five indicators are used. For that reason, it should be no surprise that in half of the tests comparing the distribution of recidivism based on gender, the analysis suggests failing to reject the null hypothesis and concluding that the distribution (of the specific indicator of Recidivism) is the same across the category of gender.

The tests for the indicators Crime Level Change, Days Charge Free, and Reoffended (y/n) have significances of 0.094, 0.358, and 0.177, respectively. These tests suggested not rejecting the null hypothesis that the distribution of these specific indicators of recidivism is the same across the category of gender. The remaining three indicators did not have the same result. Number of New Charges (which was removed from the model), Severity of New Charge, and Most Severe Sanction have significances of 0.045, 0.032, and 0.008, respectively. These tests suggest rejecting the null hypothesis and concluding that the distribution of these specific indicators of recidivism is not the same across the category of gender (see Table 20 in Appendix A).

The chi-square tests specific to gender and the Recidivism indicators Crime Level Change, Severity of New Charge, and Most Severe Sanction cannot be conclusive because of the requirement that 80% of the cells have expected frequencies of 5 or more (see Table 20 in Appendix A for the number and percentage of cells with expected frequencies less than 5). This

failure to meet the assumption is seen in every chi-square test that was not a two-by-two test, probably because so many categories made it difficult to satisfy this prerequisite. The finding was addressed by reducing the number of categories in these Recidivism indicators.

The variables Most Severe Sanction Category, Crime Level Change Sign, and Most Severe Sanction Category were thus created as explained in the methodology section above. Then the distribution of Recidivism among the predictors (gender, age group, and suspension history) was tested to determine whether the distribution was the same among the categories in those predictor variables (see Table 21 in Appendix A).

The chi-square tests specific to gender and the Recidivism indicators Crime Level Change Sign, Days Charge Free, Reoffended (y/n), and Most Severe Sanction Category all had results that suggest failing to reject the null hypothesis. The tests show that the distribution of gender is the same across these measures of Recidivism. Only the Recidivism indicators Days Charge Free and Most Severe Sanction Category are significant and suggest rejecting the null hypothesis. The tests indicate that there is a difference in the distribution of gender among these Recidivism indicators. Four of the six tests suggest no difference in the distribution of gender among Recidivism indicators for this data set when the modified Recidivism indicators were used.

When the chi-square test to test the distribution of age group and suspension history among the modified Recidivism indicators Crime Level Change Sign, Severity of New Charge Category, and Most Severe Sanction Category was conducted, none of the conclusions changed. The null hypothesis is rejected in all tests. Moreover, the assumption that at least 80% of cells have an expected frequency of five or more was satisfied.

The conclusions from these tests specific to Hypothesis 2 are that males in the data set had higher Recidivism than females had (although not a statistically significant difference). Those with an initial charge before the age of 15 had significantly higher Recidivism than did those initially charged after the age of 15. Those with histories of suspension from school had significantly higher recidivism than those with no histories of suspension from school.

The gender differences related to Recidivism observed in other studies were not observed in this study. A possible explanation relates to the data set source. The data came from a program that accepts first-time and mostly misdemeanor offenders. Since more felony crimes are committed by males, a higher proportion of the males may have been deemed unsuitable for the diversion program and thus were not included in this data set. It is further possible that the generation of youths who participated or were offered the opportunity to participate in this program do not behave in a way consistent with the relational/instrumental theory. Generational gender differences are beyond the scope of this study.

Testing Hypothesis 3

The third hypothesis was that youth who complete a restorative justice program have significantly lower recidivism than do those who opt not to participate and those removed from the program. To test this hypothesis, the predictor variable Completion was added to the measurement model of Recidivism developed and tested in Hypothesis 1. This predictive model, seen in Figure 27 in Appendix B, was run against the data set. The factor weights and their significance were then examined. The standardized regression weight between the predictor Completion and the construct of Recidivism is 0.47 and is significant at the 0.001 level (two

tailed). All the Recidivism indicators have significant standardized factor weights between 0.90 and 0.95 (see Table 27 in Appendix A).

This predictive model has a chi-square of 16.0 (with 7 degrees of freedom); a relative chi-square of 2.29; and GFI, AGFI and CFI values of 0.997, 0.932, and 0.995, respectively. These statistics show a good fit between the model and the data. The *p*-value is below 0.05 (0.025), and the RMSEA is above 0.05 (0.077). These statistics indicate an adequate fit between the data and the model (see Table 28 in Appendix A).

Because the model did not pass all SEM goodness of fit tests, the non-SEM tests described in Chapter 3 were conducted. Again, as for hypothesis 2, many of the chi-square tests could not be conclusive because of failure to meet the assumption that at least 80% of cells have an expected frequency of 5 or more. Just as for hypothesis 2, the modified Recidivism indicators with fewer coded possibilities were then used. All the tests have significance of 0.000, which suggest rejecting the null hypothesis and concluding that there is a statistically significant difference in Recidivism indicators between those youths who had completed the restorative justice program and those youths who had not (see Table 29 in Appendix A).

On the basis of the large differences in Recidivism statistics between those youths who had completed and those who had not completed the program, the significant standardized regression weight of the predictor completed in the predictive model, and the fact that all Recidivism indicators showed statistically significant distributions based on completion status, the null hypothesis for hypothesis 2 is rejected. For the second research question, this analysis suggests that completion of a juvenile restorative justice program reduced recidivism, as

compared to having been offered participation and either declining participation, being deemed unsuitable, or being removed from the program.

Initially a further comparison was suggested of those who had not completed the program in relation to the reasons for the failures. The data necessary to conduct that analysis were unfortunately not available. Such analysis could have addressed selection bias, which was raised in the methodology chapter and is discussed below in the limitations section. The next several hypotheses tested specific sanctions on specific offender types to determine effectiveness.

Testing Hypothesis 4

The testing the Hypothesis 4, that assigning drug testing to drug offenders reduces recidivism, began with a predictive model. The Recidivism measurement model developed and tested in Hypothesis 1 was developed into a predictive model with the addition of the predictor drug testing (see Figure 28 in Appendix B). This model has a good fit with the data (chi-square of 7.7 with 7 degrees of freedom; *p*-value of 0.358; GFI, AGFI and CFI of 0.956, 0.867, and 0.998, respectively; and an RMSEA of 0.044). However, the regression weight of drug testing in the prediction of Recidivism is not significantly different from zero at the 0.05 level (two-tailed).

For a test of that predictive model against the entire data set ($n = 218$), the results are similar. The model is again a good fit with the data (chi-square of 10.6 with 7 degrees of freedom; *p*-value of 0.157; GFI, AGFI and CFI of 0.984, 0.953, and 0.998, respectively; and an RMSEA of 0.049). That being said, the regression weight of drug testing in the prediction of Recidivism is not significantly different from zero at the 0.05 level (two-tailed). Although for drug offenders the predictor drug testing is more significant, neither is significant at the 0.05 level (two-tailed). These results suggest failing to reject the null hypothesis.

Because the SEM tests noted above are not significant, the non-SEM tests described in Chapter 3 were conducted next. The differences in individual Recidivism indicators between drug offenders who were assigned and those not assigned drug testing were examined (see Table 35 in Appendix A). The results of all tests using the original and modified Recidivism indicators suggest failing to reject the null hypothesis. All the chi-square tests fail to meet the assumption that 80% of cells have expected frequencies of 5 or more. This shortfall was probably caused by the fact that very few drug offenders were not assigned drug testing. As the study was a non-experimental design, it should be expected that almost all drug offenders had been assigned drug testing.

Because the regression weight of drug testing in the prediction of Recidivism is not significantly different from zero and because the individual Recidivism indicators show no significant difference between the drug offenders assigned and those not assigned drug testing, this analysis suggests failing to reject the null hypothesis. The results suggest that the assignment of drug testing to drug offenders (and all offenders in the data set) had no significant impact on Recidivism.

Testing Hypothesis 5

The testing of Hypothesis 5 began with the development of a predictive model. The predictor Curfew Time was added to the Recidivism measurement model developed and tested for Hypothesis 1. This model was first tested against the data set of the 216 offenders whose data included the assigned curfew time (see Figure 29 in Appendix B). All the standardized regression coefficients are significant at the 0.05 level (see Table 41 in Appendix A). All the goodness of fit measures indicate a good fit between the data and the model (chi-square of 2.4

with 7 degrees of freedom; p -value of 0.931; GFI, AGFI, and CFI of 0.996, 0.989, and 1.000, respectively; and RMSEA of 0.000) (see Table 42 in Appendix A).

This model is a good fit with the data and suggests a statistically significant relationship between the assigned curfew time and recidivism. As discussed in the descriptive analysis above, there is a significant relationship between Age Group (less than 15 years of age or at least 15 years of age) and Curfew Code. Furthermore, for Hypothesis 2, a significant relationship between age group and the latent construct Recidivism was proven.

To control for the effects of Age on the relationship between Curfew Time and Recidivism, the variable Age was added to this predictive model. The new model (seen in Figure 30 in Appendix B) is also an adequate fit with the data (chi-square of 19.0 with 11 degrees of freedom; p -value of 0.061; GFI, AGFI, and CFI of 0.976, 0.939, and 0.996, respectively; and RMSEA of 0.057). Once Age had been controlled for, the regression weight of Curfew Time in the prediction of the latent construct Recidivism is not statistically different from zero at the 0.05 level (two-tailed).

On the basis of the large differences in Recidivism indicators between those assigned an early curfew and those assigned a late curfew and on the basis of the good fit between the predictive model of Recidivism (based on assigned curfew time) and the data, this analysis suggests rejecting the null hypothesis and concluding that those offenders assigned a late curfew had lower recidivism than did those assigned an early curfew.

However, conclusions cannot guide the actions of those assigning curfew times, because it was also shown that younger offenders, who by their nature have greater recidivism than older

offenders do, were assigned earlier curfews. Once Age Group is controlled for in the analysis, there is no statistically significant difference in recidivism based on assigned curfew time.

When the predictive model of Recidivism based on assigned curfew time (seen in Figure 27 in Appendix B) was run against the 79 males who were at least 15 years of age, the regression weight of Curfew Time in the prediction of recidivism was not statistically different from zero at the 0.05 level (two-tailed). The model has a chi-square of 9.7 with 7 degrees of freedom; *p*-value of 0.204; GFI, AGFI, and CFI of 0.890, 0.955, and 0.996, respectively; and RMSEA of 0.071.

From these results and the small differences in Recidivism indicators between older male offenders assigned an early curfew vs. those assigned a late curfew, discussed above, this analysis suggests failing to reject the null hypothesis. There is no significant difference in observed recidivism based on the assigned curfew times for males at least 15 years of age. Furthermore there is no significant difference in observed recidivism based on the assigned curfew times for all offenders once the model controls for the offender age.

Testing Hypothesis 6

Testing Hypothesis 6 began with the addition of the predictor no-contact order to the Recidivism measurement model developed and tested for Hypothesis 1 (see Figure 31 in Appendix B). When tested against the entire data set ($n = 218$), the model is a very good fit with the data. All the goodness of fit measures indicate a good fit between the data and the model (chi-square of 3.7 with 7 degrees of freedom; *p*-value of 0.819; GFI, AGFI, and CFI of 0.994, 0.983, and 1.000, respectively; and RMSEA of 0.000). However, the regression weight for the predictor no-contact order in the prediction of the latent construct Recidivism is not statistically

different from zero at the 0.05 level (two-tailed). This result suggests failing to reject the null hypothesis and concluding that for all offenders in the entire data set, recidivism for those assigned and not assigned a no-contact order is not statistically different at the 0.05 level (two-tailed).

This predictive model was then tested against data for all males ($n = 113$) and females ($n = 85$), older offenders (at least 15 years of age) ($n = 134$) and younger offenders (less than 15 years of age) ($n = 84$), older males ($n = 81$), older females ($n = 53$), and drug offenders ($n = 54$). The goodness of fit indicators and regression weight for no-contact orders for each of these models can be found in Table 48 in Appendix A. The model is a good fit with the data (as defined by compliance with all goodness of fit indicators) for the entire data set, for all males, for both older and younger offenders, and for drug offenders. The model is not a good fit for all female offenders and for male offenders over 15 years of age. The only model where the regression weight of no-contact order in the prediction of Recidivism is statistically different from zero (at the 0.05 level, two-tailed) is for all females, but that overall model is not a good fit with the data. Where the model was a good fit with the data, the regression weight of no-contact order in the prediction of Recidivism was not statistically different from zero, and where the regression weight of no-contact order in the prediction of Recidivism is statistically different from zero, the model is not a good fit with the data.

Lastly, in the analysis for Hypothesis 6 the non-SEM tests as described in Chapter 3 were completed (see Table 49 in Appendix A). For most indicators of Recidivism (including modified indicators), the tests suggest failing to reject the null hypothesis and concluding that the distribution of the Recidivism indicators is the same for those assigned and for those not

assigned a no-contact order. Only three of the nine tests had results that suggest rejecting the null hypothesis. Crime Level Change, Crime Level Change Sign, and Days Charge Free all have significance levels below 0.05, which suggests rejecting the null hypothesis.

The chi-square test for the distribution of the Recidivism indicator Crime Level Change across those assigned and those not assigned a no-contact order cannot be relied upon because of the failure to meet the requirement that at least 80% of cells have an expected frequency of 5 or more. The condensed variable Crime Level Change Sign does have significantly different distributions for those assigned and those not assigned a no-contact order and met the prerequisite for the use of the chi-square test. The Recidivism indicator Days Charge Free is significantly different from those assigned and those not assigned no-contact orders, but this indicator has a censorship problem as discussed earlier. Thus the only indicators that show a significantly different distribution of any Recidivism indicator between those assigned and those not assigned a no-contact order are the indicators with either a censorship problem or one that required coding to be condensed to comply with the requirements of the statistic test used.

The results of the tests of the predictive models and the chi-square and Mann Whitney U-test results suggest failing to reject the null hypothesis for all parts of Hypothesis 6. There is no statistically different recidivism for any subgroup of this data based on the assignment of a no-contact order.

The Instrumental and Relational gender theory discussed in the literature review as well as the findings of Elis (2005) suggested that autonomy-limiting sanctions such as curfews and no-contact orders are likely to have the strongest impact on older males. Further, one might intuitively predict that drug offenders, who rely on other criminals to provide their drugs, would

benefit from the assignment of no-contact orders. Those predictions are not supported by the analysis. In fact, despite the failure of the overall model to be a good fit with the data, it is only in the analysis of older females that the assignment of a no-contact order has a significant (although very small) impact on recidivism. It is possible that the mere assignment of a no-contact order did not limit the offender's exposure to other delinquents or to any possible supply of narcotics. The problem of assessing adherence to assigned sanctions is discussed in the limitations section.

Testing Hypothesis 7

The testing of Hypothesis 7 began with the addition of the predictors, apology letters and essays, to the Recidivism measurement model developed and tested for Hypothesis 1 (see Figure 32 in Appendix B). When tested against the entire data set ($n = 218$), the model is a good fit with the data. All the goodness of fit measures indicate a good fit between the data and the model (chi-square of 16.3 with 11 degrees of freedom; p -value of 0.132; GFI, AGFI, and CFI of 0.979, 0.948, and 0.997, respectively; and RMSEA of 0.047) (see Table 56 in Appendix A).

However, the regression weight for the predictors, apology letters and essays, in the prediction of the latent construct of Recidivism are not statistically different from zero at the 0.05 level (two-tailed). This result suggest failing to reject the null hypothesis and concluding that for all offenders in the entire data set, recidivism for those assigned and not assigned apology letters and essays is not statistically different at the 0.05 level. This is the expected result.

When the model was run against all females ($n = 85$), the model fit was weakened (chi-square of 23.5 with 11 degrees of freedom; p -value of 0.015; GFI, AGFI, and CFI of 0.938, 0.842, and 0.982, respectively; and RMSEA of 0.116) (see Table 56 in Appendix A). Further, the

regression weight for both of the predictors, apology letters and essays, in the prediction of the latent construct Recidivism is not statistically different from zero at the 0.05 level (two-tailed). This result suggests failing to reject the null hypothesis and concluding that for all female offenders in the entire data set, recidivism for those assigned apology letters and essays and those not assigned apology letters and essays is not statistically different at the 0.05 level.

Lastly, the model was run against females who were at least 15 years of age when they received their initial charges ($n = 53$). This is the group hypothesized to be most affected by the assignment of these sanctions. The model fit does not support those predictions.

The model is a poor fit with the data (Chi-square of 24.6 with 11 degrees of freedom; p -value of 0.011; GFI, AGFI, and CFI of 0.898, 0.740, and 0.969, respectively; and RMSEA of 0.154) (see Table 56 in Appendix A). Further, the regression weight for both of the predictors, apology letters and essays, in the prediction of the latent construct Recidivism is not statistically different from zero at the 0.05 level (two-tailed). This result suggest failing to reject the null hypothesis and concluding that for older female offenders in the entire data set, recidivism for those assigned apology letters and essays and those not assigned apology letters and essays are not statistically different at the 0.05 level. In addition to the factor weights for the predictors of apology letters and essays not being significantly different from zero, it is surprising that as the analysis went from the entire data set to all females and finally to older females, the model became a progressively poorer fit with the data.

As with the other Recidivism predictors that were not determined to be significant in the predictive model, essays and letters as predictors were examined separately using the non-SEM tests described in Chapter 3 (see 57 and 58 in Appendix A). First, the distributions of the

Recidivism indicators were tested to determine whether they are the same across the predictor of apology letters, by using the chi-square and Mann Whitney U tests for the entire data set (n = 218).

Of the nine Recidivism indicators (including the modified indicators), seven suggest failing to reject the null hypothesis and concluding that the distributions of the individual Recidivism indicators are the same across the predictor of apology letters. Only two suggest otherwise. The chi-square test for Crime Level Change has a significance of 0.009, but 8 cells (44.4%) have an expected frequency of less than 5, making the validity of the results questionable. When the modified indicator Crime Level Change Sign is used, none of the cells has an expected frequency less than 5, but the significance is 0.099, which suggests failing to reject the null hypothesis.

The other indicator with significance below 0.05 is Severity of New Charge Category (0.049), another modified indicator, which passes the assumption test with only 1 cell (16.7%) with an expected frequency less than 5. This suggests there is a significant difference between the distributions of Severity of New Charge Category in relation to the assignment of letters of apology (see Table 57 in Appendix A).

In the descriptive analysis above for Severity of New Charge, the rates of receiving a new felony charge for those not assigned apology letters and for those assigned apology letters are 10.3% and 22.8%, respectively (see Table 51 in Appendix A). Those offenders who were assigned letters of apology have significantly higher recidivism than do those not assigned letters of apology, but only when recidivism is defined by the one indicator, Severity of New Charge Category.

Next, the distributions of the indicators for the latent construct Recidivism were tested to determine whether they were the same across the predictor of essays. This was accomplished again through the chi-square and Mann Whitney U tests for the entire data set ($n = 218$). All nine indicators have significance levels above 0.05, which suggests failing to reject the null hypothesis and concluding that the distributions of the indicators are the same across the predictor essays. Three of the seven chi-square tests fail the prerequisite test that 80.0% of cells have expected frequencies of five or more. Four chi-square tests and the two Mann Whitney U tests have results that both can be relied upon and suggest failing to reject the null. These results again were expected since they were tested with the data for all offenders in the data set.

Next the same tests were conducted against only the females over the age of 15 ($n = 53$) in the data set (see Table 58 in Appendix A). The results of all Recidivism indicators for both predictors (essays and letters of apology) have significance levels above 0.05, which suggest failing to reject the null hypothesis. All chi-square tests have more than 20% of cells with expected cell frequencies of less than five.

The third research question was to determine which sanctions work, specifically whether the use of specific sanctions on specific offender groups reduces recidivism. It has been hypothesized that assignment of essays and letters of apology reduces recidivism, specifically for older female offenders. The Instrumental Relational Gender theory suggests that sanctions that build relationships may be effective for females. Based on the small and insignificant factor weights for both of the predictors (essays and apology letters) in the Recidivism predictive models and based on the results of the non-SEM tests (all but one suggest failing to reject the null), the analysis of the data suggests failing to reject the null hypothesis and concluding that

there is no significant difference in recidivism between those assigned letters of apology and essays and those not assigned letters of apology and essays for the entire data set and for older female offenders. The Instrumental Relational Gender theory suggests that older female offenders would benefit from the assignment of sanctions that sought to repair relationships. The analysis does not support that assertion.

Offenders in this data set were not assigned sanctions randomly, so those assigned or not assigned these sanctions may have had other similarities that contributed to these results. Further, although the study's data set of 218 was sufficient for many of the tests conducted, when small subsections of the data set were examined the smaller sample sizes may have contributed to the fact that many of the chi-square tests fail to meet the requirement that 80% of cells have expected cell frequencies of five or more. These and other limitations are addressed in the limitations section in the next chapter.

Testing Hypothesis 8

The Offender Risk Propensity measurement model, seen in Figure 10 in Appendix B, was evaluated against the data set. The model, when tested against the entire data set ($n = 218$), is a good fit with the data (chi-square of 11.8 with 8 degrees of freedom; p -value of 0.161; GFI, AGFI, and CFI of 0.983, 0.956, and 0.940, respectively; and RMSEA of 0.047) (see Table 62 in Appendix A). The problem is that none of the six indicators of Offender Risk Propensity are significant at the 0.05 level (two-tailed). See Table 63 in Appendix A.

When the construct Offender Risk Propensity was tested as a predictor of Recidivism, the predictive model was shown to be a poor fit with the data (chi-square of 101.4 with 31 degrees of freedom; p -value of 0.000; GFI, AGFI, and CFI of 0.924, 0.866, and 0.962, respectively; and

RMSEA of 0.102) (see Table 62 in Appendix A). Interestingly, in this predictive model all of the indicators of Offender Risk Propensity have factor weights that are significant at the 0.05 level, which is not the case when the measurement model of Offender Risk Propensity is examined separately. Also, the factor weight between Offender Risk Propensity and Recidivism is 0.30, which is significant at the 0.05 level (two-tailed). Those relationships are moot, however, because the overall model is not a good fit with the data.

The same process was followed for the construct Multiplicity of Sanctions. The construct Multiplicity of Sanctions was tested as a function of the assignment of no-contact orders, anger management, drug testing, restitution, apology letters, community service, boot camps, and assigned curfew time. The model for this construct can be seen in Figure 12 in Appendix B. In order to test this hypothesis, the data set again had to be truncated because of missing data. Case number 100 had restitution missing, cases 24 and 78 had community service missing, and cases 18 and 41 were missing the curfew time. These five cases were removed from the analysis leaving a sample of 213 cases, which is still of adequate size to test the model.

When the Multiplicity of Sanctions measurement model was tested against the data, the model was shown to be a good fit with the data (chi-square of 17.3 with 16 degrees of freedom; *p*-value of 0.366; GFI, AGFI, and CFI of 0.980, 0.955, and 0.959, respectively; and RMSEA of 0.020) (see Table 62 in Appendix A). None of the factor weights, however, is significant at the 0.05 level (two-tailed) (see Table 64 in Appendix A).

With the Multiplicity of Sanction construct as a predictor of Recidivism, the predictive model is a poor fit with the data (chi-square of 91.8 with 58 degrees of freedom; *p*-value of 0.003; GFI, AGFI, and CFI of 0.993, 0.903, and 0.981, respectively; and RMSEA of 0.052) (see

Table 62 in Appendix A). None of the factor weights for the indicators of Multiplicity of Sanctions is significant.

The last measurement model tested alone in this section is for the interaction construct between Offender Risk Propensity and Multiplicity of Sanctions. When the interaction measurement model was tested against the data, the least significant indicators were systematically removed from the model. Only four indicators are significant at the 0.001 level (two-tailed): no-contact order and charge type, no-contact order and suspended, no-contact order and drug offender, and no-contact order and age. Once the significant modification indices were represented in the model, this measurement model was shown to be a good fit with the data (chi-square of 1.1 with 1 degree of freedom; *p*-value of 0.290; GFI, AGFI, and CFI of 0.997, 0.947, and 1.000, respectively; and RMSEA of 0.024) (see Table 62 in Appendix A).

With the interaction measurement model used as a predictor of Recidivism, that predictive model is a good fit with the data (chi-square of 23.1 with 22 degrees of freedom; *p*-value of 0.400; GFI, AGFI, and CFI of 0.978, 0.955, and 1.000, respectively; and RMSEA of 0.015) (see Table 62 in Appendix A). The problem with this model is that the probability of getting a regression weight estimate as high as that observed between Recidivism and the interaction construct is 0.186. This indicates the regression weight of the interaction construct in the prediction of Recidivism is not statistically different from zero at the 0.05 level (two-tailed).

Finally the relationships between all of these constructs were tested together. The main effect model, seen in Figure 33 in Appendix B, is the outcome of retaining only the indicators that contributed the most to the model, although some were not significant. Despite having mostly significant weights, the model is a poor fit with the data (chi-square of 108.0 with 47

degrees of freedom; p -value of 0.000; GFI, AGFI, and CFI of 0.927, 0.878, and 0.967, respectively; and RMSEA of 0.078) (see Table 64 in Appendix A). The most insignificant factor weights in this model are between Multiplicity of Sanctions and Recidivism (P of 0.387) and between Offender Risk Propensity and Recidivism (P of 0.385), indicating that the regression weight of the constructs Multiplicity of Sanctions and Offender Risk Propensity in the prediction of the latent construct of Recidivism is not statistically different from zero at the 0.05 level (two-tailed).

When the moderating effect model, seen in Figure 34 in Appendix B, was tested against the data, the model had the same goodness of fit statistics as the main effect model. Further, the regression weights of Offender Risk Propensity and Multiplicity of Sanctions in the prediction of Recidivism are not significantly different from zero at the 0.05 level (two-tailed).

Lastly, the interaction construct was added to the predictive model to produce the moderating predictive model, seen in Figure 35 in Appendix B. When tested against the data, it is a poor fit (chi-square of 342.6 with 93 degrees of freedom; p -value of 0.000; GFI, AGFI, and CFI of 0.844, 0.771, and 0.931, respectively; and RMSEA of 0.113) (see Table 65 in Appendix A).

The third research question of this study was which sanctions, if any, reduce recidivism for specific offenders. To test Hypothesis 8, the main, moderating, and mediating models, which combine the constructs of Recidivism, Multiplicity of Sanctions, Offender Risk Propensity, and a construct that represented an interaction between indicators of Offender Risk Propensity and Multiplicity of Sanctions into predictive models, were evaluated against the data set. The null hypothesis was that these models would not be significant. The alternative hypothesis was that

the models would be significant. The main, mediating, and moderating effect predictive models of Recidivism are all a poor fit with the data set. For that reason, this analysis suggests failing to reject the null hypothesis.

For this data set, there is no significant difference in recidivism based on specific sanctions for any offender group. Earlier, assigned curfew time was seen as significantly correlated with recidivism, with those assigned earlier curfews as having greater recidivism than did those assigned later curfews. This difference is not significant once age of the offender was controlled.

Exploratory Research

As stated in the methodology chapter, one of the main objectives of this study was to determine the effectiveness of specific sanctions on specific offender types. The research and theories discussed in the literature review laid the foundation for the hypotheses tested above. These tests however, covered a limited number of sanctions traditionally employed by diversion programs. For that reason, other sanctions were tested against the data to determine if the assignment of other sanctions has an impact on recidivism.

Thus far, this analysis has examined curfews, no-contact orders, essays, letters of apology, and drug testing. The remaining sanctions tested constitute the six next most commonly assigned by the Neighborhood Restorative Justice Program. They were: jail tours (assigned in 155 cases), community service (assigned in 100 cases), home chores (assigned in 64 cases), anger management (assigned in 56 cases), school progress reports (assigned in 52 cases), and boot camps (assigned in 42 cases).

One way to test such relationships could have been to build a predictive model, such as the one seen in Figure 21 in Appendix B, where these sanctions would serve as predictors of the Recidivism construct developed and tested for Hypothesis 1. Such an approach would be inappropriate because structural equation modeling should be used only in confirmatory research. This section is exploratory in nature, since no theories that describe the relationships between these constructs have been proposed, and no specific hypotheses were tested.

Here, a non-SEM tests as described in Chapter 3 (specifically the chi-square test) was performed to determine whether the distribution of one indicator or Recidivism (reoffended (y/n)) is the same across the categories of the sanction (assigned or not assigned). Here, the null hypothesis was that the distribution of Reoffended (y/n) is the same across the categories of the sanction (assigned or not assigned). This test was conducted for each of these six sanctions independently.

The tests were conducted using the entire data set ($n = 218$) followed by only older offenders (where older was defined as at least 15 years of age at the time of the initial charge) ($n = 134$), older females ($n = 54$), older males ($n = 81$), and drug offenders ($n = 54$), where complete data were available (see Tables 66 through 70 in Appendix A). Among these 25 individual tests, none has a significance of 0.05 or lower. For every test the null hypothesis is not rejected. The distribution of Reoffended (y/n) is the same across the categories of the individual sanctions (where categories are the specific sanctions that were either assigned or not assigned).

In many of the tests for groups that had small sub-samples, the requisite criterion for the chi-square test (that at least 80% of cells have an expected frequency of 5 or more) is not satisfied. As the subgroups get more specific, and by that nature smaller, this problem is more

likely. These were 2 by 2 tests, as both variables were coded as dichotomous (reoffended was either yes or no, and the sanction was either assigned or not assigned). For that reason, condensing variable coding for these tests is not possible. This aspect is addressed in the limitations section. That having been said, this analysis suggests no significant differences in recidivism among these groups based on the assignment of any of these sanctions, when recidivism is defined only as receiving a new criminal charge.

Summary of Findings

This study sought to answer three research questions: What is the best way to measure Recidivism? Does completion of a restorative justice program reduce recidivism? Which sanctions, if any, reduce Recidivism for specific offender types? To answer the first question: a multi-indicator latent construct of Recidivism did a very good job of measuring variation in Recidivism. Multiple indicators analyzed simultaneously produced a robust tool that can be used in other recidivism studies and help to reduce comparability issues between studies.

The Recidivism construct, when tested as a function of completion of the restorative justice program, was seen to produce a significant model having an overall good fit with the data. Thus to answer the second research question: the offender's completion status for the restorative justice program was shown to be a significant predictor of the latent construct of Recidivism at the 0.05 level (two-tailed), with those who failed to complete (or chose not to participate) having higher recidivism than did those who completed the program. To answer the third research question: the assignment of specific sanctions (both those suggested by research and theory and those traditionally assigned by this and similar programs) on the entire data set (and on various

subsets) of this study have no statistically significant impact on recidivism at the 0.05 level (two-tailed).

CHAPTER 5: DISCUSSION

In this last chapter, this study's limitations, as well as the policy, theoretical, and ethical implications are discussed. The study proposed a new tool to measure recidivism, proposed suspension history as a measure of prior deviancy to predict future delinquency, proved a strong correlation between completing a restorative justice program and lower recidivism, and found that the sanctions most commonly used in these programs have no significant impact on recidivism. These findings have many implications. First, however, the limitations will be identified.

Limitations

The limitations of the current study include a historical (non-experimental) research design, possible generalizability problems, factors not considered, limited use of control variables, data censorship, the distinction between sanction assignment and sanction completion, and selection bias. This section discusses these limitations. Research design is addressed first.

Research Design

The strongest limitation of this study is the historical research design. The study was conducted by analyzing a data set of actual offenders, the sanctions they were assigned, their demographics, and their recidivism indicators. This was not an experimental design, there was no random assignment of offenders or sanctions, and no control group was employed. It is not possible to know, for example, whether offenders assigned an earlier curfew had greater recidivism because of the sanction assignment or whether the fact that younger offenders (who by their nature have greater recidivism) were simply assigned earlier curfews by the sanctioning

body. All that can be stated in that example is that once the age of the offender is controlled, the assigned curfew time is not a significant predictor of recidivism.

Selection Bias

Part of the research design limitation further relates to selection bias. Selection bias (or selection threat) occurs when participants in a study are not randomly assigned to either the treatment or control groups. How subjects might be selected, or in the cases here how the subjects make selection choices themselves, may impact the outcome of the study. McCold and Wachtel (1998) in their study of restorative justice reported that recidivism was mostly a function of selection bias. Those who were offered and refused participation had the highest recidivism rates, followed by those not offered participation and then by those who were offered participation and completed the program.

Originally this study included an examination of the factor, but the data necessary to determine the reason for removal from the program (e.g., the difference between the program administrators' refusing participation and the offender or their parents' declining participation) were not available in the current data set. Selection bias is a limitation to the current study because those selected to participate in the program may have had lower recidivism than those not selected, all other things held constant.

Removal Reason Unknown

This study compared recidivism between those who completed and those who did not complete the restorative justice program. As discussed above, no distinction was made for the different reason for not completing the program, as sufficient data was not available. There were

14 offenders in the data set that started the program and reoffended within 90 days, which is the length of the program. It is not known if that reoffense was the sole reason for failure to complete the program or if the offender had already been removed or withdrawn from the program for another reason, like failure to comply with other program requirements. It is likely that the new offense was, at least in part, a reason for failure to complete the program for some of these offenders. However, of those 14 who reoffended within 90 days, 12 received multiple charges (ranging from 3 to 125). So even if the offense which occurred within 90 days of starting the program was not counted as a new offense, almost all of those offenders would still be counted as having reoffended. This factor is still a potential limitation worth mentioning.

Generalizability

Limitations to this study could also include a lack of generalizability to juvenile offenders as a whole, because only youths who live in small cities in Florida who were charged with first-time, mostly misdemeanor charges were used. Inner-city repeat offenders may have totally different recidivism characteristics, different outcomes from the same punishment combinations, and different outcomes from participation in a restorative justice program.

In some states, only children who have charges filed against them by the state and who are not physically arrested are eligible for diversion programs. This is not true for Florida and can cause generalizability problems in comparison with other states. For the most part, the literature has not suggested generalizability problems between countries, at least not those with similar legal systems. Studies have compared similar restorative justice programs in the United States, Canada, Australia, and New Zealand, with similar findings.

Factors Not Considered

Absent from the criminal justice literature is the use of criminal record suppression as an incentive for not re-engaging in criminal behavior. Many court diversion programs offer to participants an incentive of having no criminal record if they complete the program or if they complete the program and do not reoffend within a set time period. Again, because of the absence of academic literature on the topic, this author suggests, from experience administering a diversion program, that criminal record suppression is a strong incentive for an offender to complete a diversion program and remain offense free. In any area where record suppression is not used, generalizability issues may arise, as youths in these programs may not have as much to gain by participating in and completing the program.

The issue of criminal record suppression can be complicated by the distinction between arrest and conviction criminal history. If an offender is charged (but not physically arrested) for a criminal act, then the incentive of no criminal record is a stronger motivating force than if the offender is physically arrested, at least in Florida. That is because if an offender is arrested, a record of that arrest will exist forever, and the offender may have to explain it to employers, colleges, and the military. However, if the offender only has charges filed with no physical arrest, and he or she completes a diversion program, then there is often no record of the deviant act to explain. This would suggest that if officers plan to recommend a diversion program for a criminal act, they should file charges rather than make a physical arrest so as to keep criminal record suppression an incentive for program completion.

Limited Use of Control Variables

A weakness of the current study is the limited use of control variables. This is a limitation only for those relationships found to be significant, e.g., between participation in a restorative justice program and recidivism. For relationships found not to be significant (or part of models with an overall poor fit), further controlling for other factors would have been unnecessary. Further tests, such as comparing program completion as it impacts recidivism, could have controlled for economic status had that information been available.

Sample Size

Although the sample size here was sufficient for the specific tests employed, sample size could be considered a weakness. For some of the chi-square tests, the requirement that 80% of cells have an expected cell frequency of 5 or more was not satisfied, which required some variable coding to be condensed. Therefore either there were too many categories or the sample was not large enough. Further, analysis of very specific offender type combinations, for example, older male drug offenders charged with a violent crime, could not have been conducted. This was because as the groups became more specific, the number of offenders in the data set that fit that description shrank. Further testing of this statistical type should include a very large data set to allow for examination of specific subgroups.

Data Censorship

Data censorship is caused when examining some phenomenon for a limited time period, data either before or after the examination are missing, or censored. In the current study, the offenders were monitored between their contract dates and the dates of their criminal history

check (on or about 15 January 2009). Because the analysis was conducted more than two years after data collection, it is possible, even likely, that some offenders who had not reoffended during the examination period reoffended after the examination period. Most of the study's recidivism indicators favor lower recidivism because of this censorship issue. Days Charge Free, a recidivism indicator that was calculated by subtracting the reoffense date (or criminal history check date if no reoffense occurred) from the contract date, actually favored higher recidivism because of censorship. This is so because there was no difference between someone who reoffended on the 15th of January 2009 and someone who had never offended at all (because that was the date used for the checks). Censorship is always going to be a limitation in research unless offenders are followed to their deaths, which is simply not practical for the study of juvenile diversion programs.

Sanction Assignment vs. Sanction Completion

Just because a sanction was assigned, it is not known whether the sanction was completed, enforced, or monitored. For sanctions that required the offender to produce some proof, such as essays or letters of apology, it can be assumed that if the offender completed the program, those assignments were completed. It is unknown for this data set if drug tests were ever administered. Nor do the data include results of these tests if administered. It is also not known for this data set whether curfew checks, for example, were ever conducted and if so how many. For those who did not complete the program, it is not known how many of their assigned sanctions, if any, they actually completed. All of these limitations need to be addressed and accounted for in future studies.

Evaluation: Ethical, Punishment Philosophy, and Economic

Ethically, different arguments could be made for different aspects of offender-specific sanctioning. Consequentialists, specifically utilitarianists, would support a public policy seeking the greatest good (producing the most utility) for the greatest number. A policy that has a positive consequence is good regardless of the intent of policymakers. They might suggest that as long as recidivism decreases, the policy is good. Moral relativists would support action that sought to fit the particular situation, as offender-based sanctioning would do. Opposition to sanction standardization based on offender characteristics would come from libertarianists, who believe that the most important value is freedom and choice; to standardize sanction packages, even by offender and based on evidence-based research, would reduce choice, if only by the sanctioning body. Such actions would be opposed by anyone with a strict libertarian view (Narveson, 2002).

Sanctioning theory has several perspectives: retribution, rehabilitation, deterrence, and incapacitation. Incapacitation (jail terms or death penalty) rarely applies to juveniles in diversion programs. Rehabilitation, which is the underlying perspective in juvenile justice, would support any sanctioning that reduced recidivism, even if different offenders received different sanctions for the same violation.

The very word “penitentiary” suggests that the prison was not to be a place where offenders were merely warehoused or suffered their just deserts, but rather that the experience or incarceration was to transform their very spirit and habits of living. (Cullen & Gendreau, 2000, p. 6)

Deterrence is a philosophy of sanctioning that believes that the punishments should deter future violations by the offender sanctioned (specific deterrence) and also reduce the probability that others would break the law (general deterrence). If a sanction were perceived by the offender

as so undesirable as to prevent future violations, then those who subscribe to deterrence would support it, even if others might not perceive the sanction as undesirable. Retribution argues that the offender should simply be punished, and those that follow this view might argue that the same violations of law should result in the same sanctions regardless of the offenders' characteristics (Lab et al., 2003, pp. 5–6).

Other bases for evaluating policy actions include costs. Here, cost–benefit analysis, break-even analysis, and return on investment could be used. The Florida Department of Juvenile Justice claims that a 1% reduction in juvenile crime could save an estimated \$10.3 million in criminal justice and \$5.3 million in victim costs (FLDJJ, 2003). Based on that, it would be difficult to argue against an investment in understanding sanctioning and program effectiveness.

Policy Relevance and Theoretical Contributions

With an increasing number of juvenile offenders sent to court diversion programs, a firm understanding of the effectiveness of these programs is crucial to the success of the juveniles who participate in them. In order to have that understanding, we must first have a good measuring stick with which to evaluate participants and programs, and that tool has to be used to tease out the specific characteristics of the programs that have been observed to have a significant impact on the youths' success.

The findings from the current study have many implications. This section discusses those implications in relation to the research question they sought to answer, specifically, the conceptualization of recidivism, the effectiveness of restorative justice programs in reducing recidivism, and the effectiveness of sanctions commonly employed in these programs to reduce recidivism for specific offender typologies.

Consistency is all but nonexistent in recidivism measurements in the academic literature and in program review studies. These studies use different definitions, time periods, data sources, and reporting measures, all of which make comparability between studies difficult. Using a multi-indicator latent construct of Recidivism, such as the one proposed and proven effective in this study, as the endogenous variable to evaluate programs and their practices could reduce the problem of study comparability, which could lead to a better understanding of program characteristics and their impact on offender success.

Further, a multi-indicator latent construct of Recidivism provides a more complete picture than simply conceptualizing recidivism by one dummy variable, such as the receipt of a new criminal charge (regardless of the outcome, time lapse between treatment and reoffense, and severity of the new charge) or re-incarceration (which fails to consider criminal activity that did not result in incarceration). If the tool is going to be able to detect the most variation, the multi-indicator latent construct is the best approach.

With regard to program review, this study found that completion of the Neighborhood Restorative Justice Program was a significant predictor of recidivism for the data set examined, yet none of the eleven most commonly assigned sanctions was seen to have a significant impact on recidivism for any subgroup examined. Proponents of restorative justice might argue that it is the programs' characteristics and not their specific activities that make the programs successful. This assertion is supported by the theories discussed in Chapter 2.

The Reintegrative Shaming Theory as proposed by Braithwaite (1989) seeks to explain that sanctions imposed by members of the offender's community are more effective than those set by members of the criminal justice system, who are seen as anonymous. This is especially

true when the “shaming involves a conscious effort to shame the action of the offender but not the offender as a person” (Hipple & McGarrell, 2008, p. 557). By not chastising offenders directly (only their actions) and by limiting offenders’ contacts with the formal criminal justice system, these programs may reduce the potential for the youth to begin to see themselves as delinquents.

Labeling Theory suggests that the less youth perceives themselves as a delinquent, the less likely they are to engage in delinquent acts (Dick et al., 2004). These theories explain why in this study, participation in the Neighborhood Restorative Justice Program was shown to significantly reduce recidivism for participants, while at the same time the sanctions assigned by the program had no significant impact on recidivism.

CHAPTER 6: CONCLUSIONS

With an increasing number of youthful offenders diverted from the courts to non-traditional programs for which volunteers are assigning sanctions intuitively, a firm understanding of these programs and their characteristics is needed. The need is significant because of the increasing popularity of these programs and because of the high costs of juvenile crime. Traditional research on recidivism has focused on the offenders most likely to reoffend, on sanction assignment, and on program reviews, but not on the sanction package efficacy for specific offender types. Moreover, the recidivism studies available offer comparability issues because of different conceptualizations of recidivism.

This study offered an in-depth review of the theories and literature on recidivism in youth diversion with a specific focus on restorative justice programs, and it proposed and conducted hypothesis testing and exploratory research to answer three research questions: 1) What is the best way to measure recidivism? 2) Does completion of a restorative justice program reduce recidivism? 3) What sanctions, if any, reduce recidivism for specific offender types? A data set of actual participants in a restorative justice program was used to test measurement and predictive models in order to test these hypotheses.

To answer the first question: a multi-indicator latent construct of Recidivism that was proposed in this study did a very good job of measuring variation in Recidivism. Multiple indicators analyzed simultaneously produced a robust tool that could be used in other recidivism studies and could help to reduce comparability issues between studies. Those indicators found to most significantly contribute to the conceptualization of the latent construct of Recidivism were whether the offender reoffended, the severity of the new charge, the most severe sanction

imposed, the number of days the offender went without a new charge, and the level of change between the original and any subsequent charge. This model, which was found to be a good fit with the data, was then used to test the study hypotheses.

When completion of the restorative justice program was tested as a predictor of the Recidivism construct, the predictive model was an overall good fit with the data. So to answer the second research question: the offender's completion status of the restorative justice program was seen as a significant predictor of recidivism at the 0.05 level (two-tailed), with those who had failed to complete (or chosen not to participate) having higher recidivism than did those who had completed the program.

To answer the third research question: the assignment of specific sanctions (both those suggested by research and theory and those traditionally assigned by the Ninth Judicial Circuit's Neighborhood Restorative Justice Program and similar programs) on the entire data set (and on various subsets) had no significant impact on recidivism at the 0.05 level (two-tailed).

Further investigation into sanction efficacy should be conducted using larger data sets, an experimental design, and a wide variety of offender demographics; the data should include sanction completion and reasons for failure to complete. These limitations were all discussed as they related to this study.

Sanctioning bodies assign punishment packages based on experience and hunches, rather than on proven research as to the effectiveness of sanction packages on certain types of offenders. Clearly more research is needed to assist sanctioning bodies because of the importance of what they are trying to accomplish: keeping youths from entering or becoming more deeply involved in the criminal justice system.

APPENDIX A: TABLES

Table 1

Time Periods Used in Juvenile Recidivism Studies

Years	Source
0-.83	Guerra & Slaby, 1990
0.25	Josi & Sechrest, 1999
0.5	Bouffard & Bergseth, 2008; Fagan, 1990; Fredrick & Roy, 2003; Leeman et al., 1993; Roskey et al., 2004
0.92-2.00	Guerra & Slaby, 1990; Sontheimer & Goodstein, 1993
1	Bank et al., 1991; Fredrick & Roy, 2003; Josi & Sechrest, 1999; Leeman et al., 1993; Roskey et al., 2004; Tarte et al., 2007
1.3	Herzfield et al., 2008; Peters et al., 1997
1.42	Smith & Monastersky, 1986
1.5	Wiebush, 1993
1.7	Kahn & Chambers, 1991
1.75-4.08	Borduin et al., 1990
2	Bank et al., 1991; Fagan, 1990; Hagan & Cho, 1996; Roskey et al., 2004
2.9	Botcher & Ezell, 2005
2.04	Borduin et al., 1995
2.2	Peters et al., 1997
2.5	Gottfredson & Barton, 1993
3	Fagan, 1990; Bank et al., 1991
3.08	Borduin et al., 1990
4	Borduin et al., 1995
5	Hagan & Cho, 1996; Roskey et al., 2004
5.42	Borduin et al., 1995
5.5	Herzfield et al., 2008; Roskey et al., 2004
7.5	Botcher & Ezell, 2005
13.7	Schaeffer & Borduin, 2005
20	Levine, 2009

Note. Time period reviewed was in some cases converted from months into years.

Table 2

Tests Appropriate According to Variable Types

Types of variables	Appropriate tests
Categorical (Independent), Categorical (Dependent)	Chi-Square
Categorical (Independent), Continuous (Dependent)	Mann-Whitney U Test
Indicator (Independent), Construct (Dependent)	Structural Equation Modeling
Construct (Independent), Construct (Dependent)	Structural Equation Modeling

Note. Tests are done with confidence levels of 95%. Variable types are defined in Table 3.

Table 3

Conceptualization, Scale, Type, and Coding for Study Variables

Indicator	Scale	Type	Coding
Age	Continuous/ Ratio	Exogenous	The age of the offender as reported on the charging affidavit at the time of the initial charge multiplied by -1. (0.0) Under 15 years of age (1) At least 15 years of age
Age group	Categorical	Exogenous	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Anger management	Dichotomous	Endogenous/ Proximal	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Boot camp	Dichotomous	Endogenous/ Proximal	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Community service	Dichotomous	Endogenous/ Proximal	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Completion	Dichotomous	Exogenous	(0.0) Did not complete the NRJP (-1) Did complete the NRJP
Crime level change	Continuous/ Ratio	Endogenous/ Distal	Calculated by subtracting the severity of the initial charge from the severity of any new charge as coded in the Severity variable. (-1) Did not reoffend at a higher level (1) Reoffended at a higher level
Crime level change sign	Categorical	Condensed	(1) Drug offense (e.g. possession of drugs and DUI) (2) Property offense (e.g. theft & burglary) (3) Violent offense (e.g. assault and battery) (4) Status and other offenses (e.g. possession of alcohol by a minor & probation violations)
Crime type	Nominal Categorical	Exogenous	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Curfew	Dichotomous	Endogenous/ Proximal	The assigned curfew time in military time from 0000 to 2400 hours multiplied by -1 (0.0) Assigned a curfew before 7:38 pm (early) (1) Assigned a curfew after 7:38 pm (late)
Curfew time	Continuous/ Ratio	Exogenous	Calculated by subtracting the contract date from the date of the first new charge (or criminal history check date in cases where there was no new offense). (0.0) No drug charge or admitted drug use (-1) Drug charge or admitted drug use
Curfew code	Categorical	Condensed	
Days charge free	Continuous/ Ratio	Endogenous/ Distal	
Drug offender	Categorical	Exogenous	

Indicator	Scale	Type	Coding
Drug testing	Dichotomous	Endogenous/ Proximal	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Essay	Dichotomous	Endogenous/ Proximal	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Gender	Categorical	Exogenous	(-1) Male (-2) Female
Jail tour	Dichotomous	Endogenous/ Proximal	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Home chores	Dichotomous	Endogenous/ Proximal	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Letter	Dichotomous	Endogenous/ Proximal	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Most severe sanction	Ordinal	Endogenous/ Distal	(0.0) No new charge (1) Non-file (The state declined to prosecute.) (2) Nolle pros (The state started to prosecute, but dropped the charge.) (3) Judicial warning (4) Probation (5) Secure detention (6) Committed to DJJ (Low Security) (7) Committed to DJJ (Medium Security) (8) Committed to DJJ (High Security) (9) Adult jail (10) Adult prison (11) Execution
Most severe sanction category	Ordinal	Condensed	(0.0) Sanctioned to less than probation (1) Sanctioned to probation or higher
Multiple charges (y/n)	Ordinal	Endogenous/ Distal	(0.0) Offender received one or fewer new charges (1) Offender received more than one new charge
No-contact	Dichotomous	Endogenous/ Proximal	(0.0) Not assigned as a sanction (1) Assigned as a sanction
Number of new charges	Continuous/ Ratio	Endogenous/ Distal	Calculated by adding the total number of new criminal charges received by the offender between the contract date and the date of the criminal history check.
Race	Categorical	Exogenous	(1) Caucasian (including Hispanic) (2) Black
Restitution	Dichotomous	Endogenous/ Proximal	(0.0) Not assigned as a sanction (1) Assigned as a sanction

Indicator	Scale	Type	Coding
Reoffended (y/n)	Ordinal	Endogenous/ Distal	(0.0) No (1) Yes
Severity of new charge	Ordinal	Endogenous/ Distal	(0.0) No new offense (1) Status offense (2) Second degree misdemeanor (3) First degree misdemeanor (4) Third degree felony (5) Second degree felony (6) First degree felony (7) Life felony (8) Capital felony
Severity of new charge category	Ordinal	Condensed	(0.0) No new offense (1) New misdemeanor charge (2) New felony charge
Suspended	Categorical	Exogenous	(0.0) Never suspended (-1) Has been suspended

Table 4

Severity of New Charge Frequencies and Percentages by Category for the Data Set and for Those Who Reoffended

Category	Frequency	Percentage	
		Date set (n = 218)	Reoffenders (n = 72)
No new charge	146	67.0	N/A
Status offense/other	2	0.9	2.8
Second degree misdemeanor	3	1.4	4.2
First degree misdemeanor	21	9.6	29.2
Third degree felony	15	6.9	20.8
Second degree felony	25	11.5	34.7
First degree felony	6	2.8	8.3

Table 5

Crime Level Change Observed: Frequencies and Percentages for Entire Data Set

Crime level change	Frequency	Percentage
-5	1	0.5
-4	16	7.3
-3	60	27.5
-2	66	30.3
-1	9	4.1
0	13	6.0
1	22	10.1
2	14	6.4
3	17	7.8

Note. n = 218

Table 6

Number of New Charges, Observed Frequencies, and Percentages for Entire Data Set

Number of new charges	Frequency	Percentage
0	146	67.0
1	21	9.6
2	7	3.2
3	10	4.6
4	6	2.8
5	1	0.5
6	5	2.3
7	4	1.8
8	2	0.9
9	3	1.4
10	1	0.5
12	2	0.9
16	1	0.5
22	1	0.5
31	1	0.5
36	2	0.9
38	1	0.5
39	1	0.5
42	1	0.5
48	1	0.5
125	1	0.5

Note. n = 218

Table 7

Most Severe Sanctions, Observed Frequencies, and Percentages for Entire Data Set

Most severe sanction (case disposition)	Frequency	Percentage
(0.0) No new charge	146	67.0
(1) Non-file	6	2.8
(2) Nolle pros	6	2.8
(3) Judicial warning	8	3.7
(4) Probation	33	15.1
(5) Secure detention	5	2.3
(6) Committed to DJJ (low security)	1	0.5
(7) Committed to DJJ (medium security)	3	1.4
(8) Committed to DJJ (high security)	5	2.3
(9) Adult jail	2	0.9
(10) Adult prison	3	1.4

Note. n = 218

Table 8

Pearson Correlations for Indicators of the Recidivism Construct

	Re-offended (y/n)	Days charge free	Number of new charges	Most severe sanction	Crime level change	Severity of new charge
Days charge free	.944					
Number of new charges	.393	.411				
Most severe sanction	.851	.850	.586			
Crime level change	.889	.852	.469	.799		
Severity of new charge	.942	.901	.482	.852	.936	
Multiple charges (y/n)	.797	.805	.453	.777	.745	.816

Note. Correlation for all indicators was significant at the .01 level (2-tailed).

Table 9

Path Analysis of Recidivism Indicators for Measurement Model Using the Entire Data Set

Indicator	Standardized regression coefficient	Unstandardized regression coefficient	Standard error	Critical value	P
Severity of new charge	0.973				*
Reoffended (y/n)	0.972	0.232	0.005	42.367	*
Days charge free	0.945	404.706	11.781	34.353	*
Most severe sanction	0.879	1.077	0.044	24.578	*
Crime level change	0.930	0.982	0.031	31.434	*
Number of new charges	0.468	2.615	0.341	7.673	*

Note. * indicates significance at .001 level (two-tailed); n = 218

Table 10

Path Analysis of Recidivism Indicators for Revised Measurement Model Using the Entire Data

Set

Metric	Ideal model	Recidivism model with all indicators	When number of new charges replaced with multiple charges (y/n)	Modification indices added	When multiple charges (y/n) was removed
Indicators		6	6	6	5
Chi-Square	Low	152.5	101.5	4.2	0.4
Degrees of freedom		9	9	6	3
Relative Chi Square	<5	16.948	11.283	0.695	0.119
<i>p</i>	> .05	0.000	0	0.654	0.949
GFI	> .9	0.809	0.864	0.993	0.999
AGFI	> .9	0.554	0.683	0.977	0.997
CFI	> .9	0.922	0.953	1.000	1.000
RMSEA	< .05	0.271	0.218	0.000	0.000
Interpretation	Good fit	Poor fit	Poor fit	Good fit *	Good fit *

Note. * Good fit determined because all goodness of fit measures indicated a good fit between the model and the data; n = 218.

Table 11

Goodness of Fit Statistics for Recidivism Measurement Model

Metric	Ideal Model	Recidivism model with all indicators	Replaced number of new charges with multiple charges (y/n)	Modification indices added	Removed multiple charges (y/n)
Indicators		6	6	6	5
Chi-Square	Low	152.5	101.5	4.2	0.4
Degrees of freedom		9	9	6	3
Relative Chi-square	< 5	16.948	11.283	0.695	0.119
<i>p</i>	> .05	0.000	0	0.654	0.949
GFI	> .9	0.809	0.864	0.993	0.999
AGFI	> .9	0.554	0.683	0.977	0.997
CFI	> .9	0.922	0.953	1.000	1.000
RMSEA	< .05	0.271	0.218	0.000	0.000
Interpretation	Good fit	Poor fit	Poor fit	Good fit *	Good fit *

Note. * Good fit determined because all goodness of fit measures indicated a good fit between the model and the data; n = 218.

Table 12

Severity of New Charge Categories and Frequencies and Percentages by Gender, Age Group, and Suspension History

Category	Gender				Age group				Suspension history			
	Males (n = 133)		Females (n = 85)		Under 15 (n = 84)		At least 15 (n = 134)		Never suspended (n = 68)		Suspended (n = 150)	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
No new charge	84	63.2	62	72.9	43	51.2	103	76.9	59	86.8	87	58.0
Status offense/other	1	0.8	1	1.2	1	1.2	1	0.7	0	0.0	2	1.3
Second degree misdemeanor	0	0.0	3	3.5	1	1.2	2	1.5	1	1.5	2	1.3
First degree misdemeanor	11	8.3	10	11.8	10	11.9	11	8.2	2	2.9	19	12.7
Third degree felony	12	9.0	3	3.5	8	9.5	7	5.2	2	2.9	13	8.7
Second degree felony	20	15.0	5	5.9	16	19.0	9	6.7	3	4.4	22	14.7
First degree felony	5	3.8	1	1.2	5	6.0	1	0.7	1	1.5	5	3.3

Table 13

Most Severe Sanction: Frequencies and Percentages by Gender, Age Group, and Suspension History

Most severe sanction	Gender				Age group				Suspension history			
	Males (n = 133)		Females (n = 85)		Under 15 (n = 84)		At least 15 (n = 134)		Never suspended (n = 68)		Suspended (n = 150)	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
(0.0) No new charge	84	63.2	62	72.9	43	51.2	103	76.9	59	86.8	87	58.0
(1) Non-file	3	2.3	3	3.5	3	3.6	3	2.2	1	1.5	5	3.3
(2) Nolle pros	2	1.5	4	4.7	4	4.8	2	1.5	0	0	6	4.0
(3) Judicial warning	8	6.0	0	0	4	4.8	4	3.0	1	1.5	7	4.7
(4) Probation	23	17.3	10	11.8	16	19.0	17	12.7	6	8.8	27	18.0
(5) Secure detention	0	0	5	5.9	2	2.4	3	2.2	0	0	5	3.3
(6) Committed to DJJ (L)	1	0.8	0	0	1	1.2	0	0	0	0	1	0.7
(7) Committed to DJJ (M)	3	2.3	0	0	2	2.4	1	0.7	0	0	3	2.0
(8) Committed to DJJ (H)	4	3.0	1	1.2	5	6.0	0	0	0	0	5	3.3
(9) Adult jail	2	1.5	0	0	2	2.4	0	0	0	0	2	1.3
(10) Adult prison	3	2.3	0	0	2	2.4	1	0.7	1	1.5	2	1.3

Note. DJJ = Department of Juvenile Justice; L = Low security; M = Medium security; H = High security.

Table 14

Crime Level Change: Frequencies and Percentages by Gender, Age Group, and Suspension History

Crime level change	Gender				Age group				Suspension history			
	Males (n = 133)		Females (n = 85)		Under 15 (n = 84)		At least 15 (n = 134)		Never suspended (n = 68)		Suspended (n = 150)	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
-5	1	0.8	0	0.0	0	0.0	1	0.7	1	1.5	0	0.0
-4	12	9.0	4	4.7	7	8.3	9	6.7	5	7.4	11	7.3
-3	38	28.6	22	25.9	14	16.7	46	34.3	15	22.1	45	30.0
-2	29	21.8	37	43.5	22	26.2	44	32.8	36	52.9	30	20.0
-1	6	4.5	3	3.5	3	3.6	6	4.5	4	5.9	5	3.3
0	10	7.5	3	3.5	5	6.0	8	6.0	1	1.5	12	8.0
1	13	9.8	9	10.6	10	11.9	12	9.0	4	5.9	18	12.0
2	11	8.3	3	3.5	9	10.7	5	3.7	1	1.5	13	8.7
3	13	9.8	4	4.7	14	16.7	3	2.2	1	1.5	16	10.7

Table 15

Number of New Charges: Frequencies and Percentages by Gender, Age Group, and Suspension History

Number of new charges	Gender				Age group				Suspension history			
	Males (n = 133)		Females (n = 85)		Under 15 (n = 84)		At least 15 (n = 134)		Never suspended (n = 68)		Suspended (n = 150)	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
0	84	63.2	62	72.9	43	51.2	103	76.9	59	86.8	87	58.0
1	11	8.3	10	11.8	9	10.7	12	9.0	1	1.5	20	13.3
2	5	3.8	2	2.4	5	6.0	2	1.5	2	2.9	5	3.3
3	7	5.3	3	3.5	4	4.8	6	4.5	3	4.4	7	4.7
4	2	1.5	4	4.7	5	6.0	1	0.7	0	0.0	6	4.0
5	1	0.8	0	0.0	1	1.2	0	0.0	1	1.5	0	0.0
6	2	1.5	3	3.5	3	3.6	2	1.5	0	0.0	5	3.3
7	4	3.0	0	0.0	0	0.0	4	3.0	2	2.9	2	1.3
8	2	1.5	0	0.0	2	2.4	0	0.0	0	0.0	2	1.3
9	3	2.3	0	0.0	1	1.2	2	1.5	0	0.0	3	2.0
10	0	0.0	1	1.2	1	1.2	0	0.0	0	0.0	1	0.7
12	2	1.5	0	0.0	0	0.0	2	1.5	0	0.0	2	1.3
16	1	0.8	0	0.0	1	1.2	0	0.0	0	0.0	1	0.7
22	1	0.8	0	0.0	1	1.2	0	0.0	0	0.0	1	0.7
31	1	0.8	0	0.0	1	1.2	0	0.0	0	0.0	1	0.7
36	2	1.5	0	0.0	2	2.4	0	0.0	0	0.0	2	1.3
38	1	0.8	0	0.0	1	1.2	0	0.0	0	0.0	1	0.7
39	1	0.8	0	0.0	1	1.2	0	0.0	0	0.0	1	0.7
42	1	0.8	0	0.0	1	1.2	0	0.0	0	0.0	1	0.7
48	1	0.8	0	0.0	1	1.2	0	0.0	0	0.0	1	0.7
125	1	0.8	0	0.0	1	1.2	0	0.0	0	0.0	1	0.7

Table 16

Days Charge Free Statistics: Based on Gender, Age Group, and Suspension History

Statistic	Gender		Age group		Suspension history	
	Males (n = 133)	Females (n = 85)	Under 15 (n = 84)	At least 15 (n = 134)	Never suspend- ed (n = 68)	Suspend- ed (n = 150)
Minimum days	15	35	23	15	35	15
Maximum days	2544	2481	2521	2544	2479	2544
Mean days	1513.7	1665.1	1353.2	1710.4	1908.4	1420.6
Maximum years	7	6.8	7	7	6.8	7
Mean years	4.2	4.6	3.7	4.7	5.2	3.9
SD (days)	878.4	786.8	888	790	645.7	882.3

Table 17

Pearson Correlations Between Gender, Age Group, and Suspension History

	Age group	Gender
Gender	0.015	
Suspended	0.057	-0.01

Note. No correlations were significant at the .01 level (2-tailed)

Table 18

Path Analysis for Predictive Model of Recidivism Based on Age Group, Gender, and Suspension History

Construct/Indicator	Construct /Indicator	Standard-ized regression coefficient	Unstandard-ized regression coefficient	Standard error	Critical value	P
Recidivism	Age	0.34	0.37	0.67	5.52	*
Recidivism	Gender	0.10	0.40	0.24	1.63	
Recidivism	Suspended	0.29	1.21	0.26	4.77	*
Severity of new charge	Recidivism	0.95	1.00			*
Reoffended (y/n)	Recidivism	0.99	0.24	0.01	39.84	*
Days charge free	Recidivism	0.95	416.95	13.35	31.23	*
Most severe sanction	Recidivism	0.90	1.13	0.05	24.04	*
Crime level change	Recidivism	0.89	0.97	0.03	37.28	*

Note. * indicates significance at the .001 level (two-tailed).

Table 19

Goodness of Fit Statistics for Predictive Model of Recidivism Based on Age Group, Gender, and Suspension History

Metric	Ideal model	Predictive model all indicators	Predictive model with gender removed
Indicators		6	6
Chi-square	Low	34.8	18.7
Degrees of freedom		15	11
Relative chi-square	< 5	2.321	1.698
<i>p</i>	> .05	0.003	0.067
GFI	> .9	0.963	0.977
AGFI	> .9	0.911	0.940
CFI	> .9	0.999	0.996
RMSEA	< .05	0.078	0.057
Interpretation	Good fit*	Poor fit	Adequate fit

Note. * Good fit found because all goodness of fit measures indicate a good fit between the model and the data.

Table 20

Chi-Square and Mann-Whitney U Tests for Distributions of Suspension History, Age Group, and Gender and Among the

Indicators of Recidivism

Predictor	Indicator	Test conducted	Significance	Decision	Assumption test cells (%)	Chi- Square Table
Suspended	Crime level change	Chi-Square	0.000	Reject Null	6 (33.3)	2x9
	Days charge free	Mann-Whitney U test	0.001	Reject Null		
	Number of new charges	Mann-Whitney U test	0.000	Reject Null		
	Reoffended (y/n)	Chi-Square	0.000	Reject Null	0 (0.0)	2x2
	Severity of new charge	Chi-Square	0.005	Reject Null	7 (50.0)	2x7
	Most severe sanction	Chi-Square	0.027	Reject Null	18 (81.8)	2x11
Age group	Crime level change	Chi-Square	0.001	Reject Null	3 (16.7)	2x9
	Days charge free	Mann-Whitney U test	0.010	Reject Null		
	Number of new charges	Mann-Whitney U test	0.000	Reject Null		
	Reoffended (y/n)	Chi-Square	0.000	Reject Null	0 (0.0)	2x2
	Severity of new charge	Chi-Square	0.004	Reject Null	6 (42.9)	2x7
	Most severe sanction	Chi-Square	0.004	Reject Null	18 (81.8)	2x11
Gender	Crime level change	Chi-Square	0.094	Retain Null	5 (27.8)	2x9
	Days charge free	Mann-Whitney U test	0.358	Retain Null		
	Number of new charges	Mann-Whitney U test	0.045	Reject Null		
	Reoffended (y/n)	Chi-Square	0.177	Retain Null	0 (0.0)	2x2
	Severity of new charge	Chi-Square	0.032	Reject Null	6 (42.9)	2x7
	Most severe sanction	Chi-Square	0.008	Reject Null	18 (81.8)	2x11

Note. Assumption test, which applied only to the Chi-square tests, indicates the number of cells (and percentages of cells) with expected counts less than 5; for age group < 15 years of age and those \geq 15 years of age; the significance level for these tests was .05.

Table 21

Chi-square and Mann-Whitney U Tests for Distributions of Suspension History, Age Group, and Gender and Among Condensed Indicators of Recidivism

Predictor	Indicator	Test conducted	Sig	Decision	Assumption test cells (%)	Chi-Square table
Suspended	Crime level change sign	Chi-Square	0.001	Reject Null	0 (0.0)	2x2
	Severity of new charge category	Chi-Square	0.000	Reject Null	0 (0.0)	2x3
	Most severe sanction category	Chi-Square	0.003	Reject Null	0 (0.0)	2x2
Age category	Crime level change sign	Chi-Square	0.000	Reject Null	0 (0.0)	2x2
	Severity of new charge category	Chi-Square	0.000	Reject Null	0 (0.0)	2x3
	Most severe sanction category	Chi-Square	0.003	Reject Null	0 (0.0)	2x2
Gender	Crime level change sign	Chi-Square	0.131	Retain Null	0 (0.0)	2x2
	Severity of new charge category	Chi-Square	0.005	Reject Null	0 (0.0)	2x3
	Most severe sanction category	Chi-Square	0.164	Retain Null	0 (0.0)	2x2

Note. Assumption test, which applied only to the Chi-square tests, indicates the number of cells (and percentages of cells) with expected counts less than 5; for age group < 15 years of age and those >= 15 years of age; the significance level for these tests was .05.

Table 22

Severity of New Charge Categories: Frequencies and Percentages Based on Competition Status
in the NRJP

Category	Not completed		Completed	
	No.	%	No.	%
No new charge	7	19.4	139	76.4
Status offense/other	2	5.6	0	0.0
Second degree misdemeanor	1	2.8	2	1.1
First degree misdemeanor	5	13.9	16	8.8
Third degree felony	6	16.7	9	4.9
Second degree felony	12	33.3	13	7.1
First degree felony	3	8.3	3	1.6

Table 23

Most Severe Sanction: Frequencies and Percentages Based on Competition Status in the NRJP

Most severe sanction	Not completed		Completed	
	No.	%	No.	%
(0.0) No new charge	7	19.4	139	76.4
(1) Non-file	0	0.0	6	3.3
(2) Nolle pros	0	0.0	6	3.3
(3) Judicial warning	2	5.6	6	3.3
(4) Probation	15	41.7	18	9.9
(5) Secure detention	4	11.1	1	0.5
(6) Committed to DJJ (L)	1	2.8	0	0.0
(7) Committed to DJJ (M)	2	5.6	1	0.5
(8) Committed to DJJ (H)	5	13.9	0	0.0
(9) Adult jail	0	0.0	2	1.1
(10) Adult prison	0	0.0	3	1.6

Note. DJJ = Department of Juvenile Justice; L = Low security; M = Medium security; H = High security.

Table 24

Crime Level Change: Frequencies and Percentages Based on Competition Status in the NRJP

Crime level change	Not completed		Completed	
	No.	%	No.	%
-5	0	0.0	1	0.5
-4	0	0.0	16	8.8
-3	5	13.9	55	30.2
-2	3	8.3	63	34.6
-1	2	5.6	7	3.8
0	3	8.3	10	5.5
1	9	25.0	13	7.1
2	6	16.7	8	4.4
3	8	22.2	9	4.9

Table 25

Number of New Charges: Frequencies and Percentages Based on Competition Status in the NRJP

Number of new charges	Not completed		Completed	
	No.	%	No.	%
0	7	19.4	139	76.4
1	5	13.9	16	8.8
2	2	5.6	5	2.7
3	4	11.1	6	3.3
4	2	5.6	4	2.2
5	1	2.8	0	0.0
6	3	8.3	2	1.1
7	1	2.8	3	1.6
8	1	2.8	1	0.5
9	1	2.8	2	1.1
10	1	2.8	0	0.0
12	1	2.8	1	0.5
16	1	2.8	0	0.0
22	0	0.0	1	0.5
31	1	2.8	0	0.0
36	2	5.6	0	0.0
38	0	0.0	1	0.5
39	1	2.8	0	0.0
42	0	0.0	1	0.5
48	1	2.8	0	0.0
125	1	2.8	0	0.0

Table 26

Days Charge Free: Statistics Based on Competition Status in the NRJP

Statistic	Not completed	Completed
Minimum days	15.0	35.0
Maximum days	2521.0	2544.0
Mean days	621.8	1760.8
Maximum years	6.9	7.0
Mean years	1.7	4.8
SD (days)	811.9	716.7

Table 27

Path Analysis for Predictive Model of Recidivism Based on Completion of Restorative Justice Program

Construct/Indicator	Construct/Indicator	Standardized regression coefficient	Unstandardized regression coefficient	Standard error	Critical value	P
Recidivism	Completed	0.47	2.44	0.32	7.73	*
Severity of new charge	Gender	0.95	1.00	0.00	0.00	
Reoffended (y/n)	Suspended	0.99	0.24	0.01	39.85	*
Days charge free	Recidivism	0.95	417.10			*
Reoffended (y/n)	Recidivism	0.90	1.12	0.05	24.07	*
Crime level change	Recidivism	0.90	0.97	0.03	37.28	*

* indicates significance at .001 level (two-tailed).

Table 28

Goodness of Fit Statistics for Predictors of Recidivism Based on Completion Status in the NRJP

Metric	Ideal model	Predictive model
Indicators		6
Chi-Square	Low	16.0
Degrees of freedom		7
Relative Chi Square	<5	2.29
<i>p</i>	>.05	0.025
GFI	>.9	0.977
AGFI	>.9	0.932
CFI	>.9	0.995
RMSEA	<.05	0.077
Interpretation	Good fit	Poor fit

Note. * Good fit found because all goodness of fit measures indicate a good fit between the model and the data.

Table 29

Chi-Square and Mann-Whitney Tests for Distributions of JRJP Completion Among the Indicators of Recidivism

Predictor	Indicator	Test conducted	Significance	Decision	Assumption test cells (%)	Chi-Square table
Completed	Crime level change	Chi-Square	0.000	Reject Null	8 (44.4)	2x9
Completed	Crime level change sign*	Chi-Square	0.000	Reject Null	0 (0.0)	2x2
Completed	Days charge free	Mann-Whitney U test	0.000	Reject Null		
Completed	Number of new charges	Mann-Whitney U test	0.000	Reject Null		
Completed	Reoffended (y/n)	Chi-Square	0.000	Reject Null	0 (0.0)	2x2
Completed	Severity of new charge	Chi-Square	0.000	Reject Null	8 (57.1)	2x7
Completed	Severity of new charge category*	Chi-Square	0.000	Reject Null	0 (0.0)	2x3
Completed	Most severe sanction	Chi-Square	0.000	Reject Null	15 (68.32)	2x11
Completed	Most severe sanction category*	Chi-Square	0.000	Reject Null	0 (0.0)	2x2

Note. Assumption test, which applies only to the Chi-square tests, indicates the number of cells (and percentage of cells) with an expected count less than 5; the significance level for these tests was 0.05; * indicates this variable was modified.

Table 30

Severity of New Charge: Frequencies and Percentages Based on the Assignment of Drug Testing for Drug Offenders

Category	Drug testing					
	Drug offenders (n = 54)		Not assigned (n = 8)		Assigned (n = 46)	
	No.	%	No.	%	No.	%
No new charge	38	70.4	4	50.0	34	73.9
Status offense/other	1	1.9	1	12.5	0	0.0
Second degree misdemeanor	0	0.0	0	0.0	0	0.0
First degree misdemeanor	6	11.1	2	25.0	4	8.7
Third degree felony	2	3.7	0	0.0	2	4.3
Second degree felony	5	9.3	1	12.5	4	8.7
First degree felony	2	3.7	0	0.0	2	4.3

Table 31

Most Severe Sanction: Frequencies and Percentages Based on Assignment of Drug Testing to Drug Offenders

Category	Drug testing					
	Drug offenders (n = 54)		Not assigned (n = 8)		Assigned (n = 46)	
	No.	%	No.	%	No.	%
(0.0) No new charge	38	70.4	4	50.0	34	73.9
(1) Non-file	1	1.9	0	0.0	1	2.2
(2) Nolle pros	1	1.9	0	0.0	1	2.2
(3) Judicial warning	1	1.9	1	12.5	0	0.0
(4) Probation	10	18.5	2	25.0	8	17.4
(5) Secure detention	2	3.7	1	12.5	1	2.2
(6) Committed to DJJ (L)	0	0.0	0	0.0	0	0.0
(7) Committed to DJJ (M)	0	0.0	0	0.0	0	0.0
(8) Committed to DJJ (H)	1	1.9	0	0.0	1	2.2
(9) Adult jail	0	0.0	0	0.0	0	0.0
(10) Adult prison	0	0.0	0	0.0	0	0.0

Note. DJJ = Department of Juvenile Justice; L = Low security; M = Medium security; H = High security.

Table 32

Crime Level Change: Frequencies and Percentages Based on Assignment of Drug Testing to Drug Offenders

Crime level change	Drug offenders (n = 54)		Drug testing			
			Not assigned (n = 8)		Assigned (n = 46)	
	No.	%	No.	%	No.	%
-5	1	1.9	1	12.5	0	0.0
-4	5	9.3	0	0.0	5	10.9
-3	24	44.4	2	25.0	22	47.8
-2	8	14.8	2	25.0	6	13.0
-1	1	1.9	0	0.0	1	2.2
0	4	7.4	1	12.5	3	6.5
1	6	11.1	2	25.0	4	8.7
2	2	3.7	0	0.0	2	4.3
3	3	5.6	0	0.0	3	6.5

Table 33

Number of New Charges: Frequencies and Percentages Based on Assignment of Drug Testing to Drug Offenders

Category	Drug offenders (n = 54)		Drug testing			
			Not assigned (n = 8)		Assigned (n = 46)	
	No.	%	No.	%	No.	%
0	38	70.4	4	50.0	34	73.9
1	3	5.6	2	25.0	1	2.2
2	2	3.7	1	12.5	1	2.2
3	2	3.7	0	0.0	2	4.3
6	3	5.6	0	0.0	3	6.5
7	2	3.7	0	0.0	2	4.3
8	1	1.9	0	0.0	1	2.2
12	1	1.9	0	0.0	1	2.2
16	1	1.9	1	12.5	0	0.0
36	1	1.9	0	0.0	1	2.2

Table 34

Days Charge Free: Statistics Based on Assignment of Drug Testing to Drug Offenders

Statistic	Drug offenders (n = 54)	Drug testing	
		Not assigned (n = 8)	Assigned (n = 46)
Minimum days	15.0	329.0	15.0
Maximum days	2451.0	2180.0	2451.0
Mean days	1580.9	1324.0	1625.6
Maximum years	6.7	6.0	6.7
Mean years	4.3	3.6	4.5
SD (days)	806.8	828.0	803.9

Table 35

Chi-Square and Mann-Whitney Tests for the Distributions of Drug Testing Assignment to Drug Offenders Among Indicators of Recidivism

Indicator	Test conducted	Significance	Decision	Assumption test cells (%)	Chi-Square table
Crime level change	Chi-Square	0.197	Retain Null	15 (83.3)	2x9
Crime level change sign*	Chi-Square	1.000	Retain Null	2 (50.0)	2x2
Days charge free	Mann-Whitney U test	0.567	Retain Null		
Number of new charges	Mann-Whitney U test	0.312	Retain Null		
Reoffended (y/n)	Chi-Square	0.343	Retain Null	1 (25.0)	2x2
Severity of new charge	Chi-Square	0.120	Retain Null	8 (57.1)	2x5
Severity of new charge category*	Chi-Square	0.082	Retain Null	2 (33.3)	2x3
Most severe sanction	Chi-Square	0.174	Retain Null	15 (68.32)	2x6
Most severe sanction category*	Chi-Square	0.607	Retain Null	2 (50.0)	2x2

Note. The predictor for each test was the assignment of drug testing; assumption test, which applies only to the Chi-square tests, indicates the number of cells (and percentage of cells) with an expected count less than 5; the significance level for these tests was 0.05; * indicates this variable was modified.

Table 36

Severity of New Charge: Frequencies and Percentages Based on the Assigned Curfew Time for Data Set and for Older

Male Offenders

Category	All offenders (n = 216)				Older males (n = 79)			
	Before 7:38 pm		After 7:38 pm		Before 7:38 pm		After 7:38 pm	
	No.	%	No.	%	No.	%	No.	%
No new charge	66	59.5	78	74.3	21	72.4	37	74.0
Status offense/other	1	0.9	1	1.0	0	0.0	0	0.0
Second degree misdemeanor	1	0.9	2	1.9	0	0.0	0	0.0
First degree misdemeanor	13	11.7	8	7.6	4	13.8	2	4.0
Third degree felony	8	7.2	7	6.7	1	3.4	5	10.0
Second degree felony	19	17.1	6	5.7	3	10.3	5	10.0
First degree felony	3	2.7	3	2.9	0	0.0	1	2.0

Table 37

Most Severe Sanction: Frequencies and Percentages Based on the Assigned Curfew Time for Data Set and for Older

Male Offenders

Category	All offenders (n = 216)				Older males (n = 79)			
	Before 7:38 pm		After 7:38 pm		Before 7:38 pm		After 7:38 pm	
	No.	%	No.	%	No.	%	No.	%
(0.0) No new charge	66	59.5	78	74.3	21	72.4	37	74.0
(1) Non-file	4	3.6	2	1.9	2	6.9	0	0.0
(2) Nolle pros	6	5.4	0	0.0	1	3.4	0	0.0
(3) Judicial warning	5	4.5	3	2.9	2	6.9	2	4.0
(4) Probation	16	14.4	17	16.2	3	10.3	9	18.0
(5) Secure detention	3	2.7	2	1.9	0	0.0	0	0.0
(6) Committed to DJJ (L)	1	0.9	0	0.0	0	0.0	0	0.0
(7) Committed to DJJ (M)	2	1.8	1	1.0	0	0.0	1	2.0
(8) Committed to DJJ (H)	5	4.5	0	0.0	0	0.0	0	0.0
(9) Adult jail	2	1.8	0	0.0	0	0.0	0	0.0
(10) Adult prison	1	0.9	2	1.9	0	0.0	1	2.0

Note. DJJ = Department of Juvenile Justice; L = Low security; M = Medium security; H = High security. Older male offenders were at least 15 years of age.

Table 38

Crime Level Change: Frequencies and Percentages Based on the Assigned Curfew Time for Data Set and for Older

Male Offenders

Crime level change	All offenders (n = 216)				Older males (n = 79)			
	Before 7:38 pm		After 7:38 pm		Before 7:38 pm		After 7:38 pm	
	No.	%	No.	%	No.	%	No.	%
-5	1	0.9	0	0.0	1	3.4	0	0.0
-4	6	5.4	9	8.6	2	6.9	5	10.0
-3	29	26.1	30	28.6	12	41.4	17	34.0
-2	30	27.0	36	34.3	6	20.7	11	22.0
-1	3	2.7	6	5.7	0	0.0	4	8.0
0	7	6.3	6	5.7	4	13.8	3	6.0
1	13	11.7	9	8.6	2	6.9	5	10.0
2	8	7.2	6	5.7	1	3.4	4	8.0
3	14	12.6	3	2.9	1	3.4	1	2.0

Note. Older male offenders were at least 15 years of age.

Table 39

Frequencies and Percentages Based on the Assigned Curfew Time for Data Set and for Older Male Offenders

Number of new charges	All offenders (n = 216)				Older males (n = 79)			
	Before 7:38 pm		After 7:38 pm		Before 7:38 pm		After 7:38 pm	
	No.	%	No.	%	No.	%	No.	%
0	66	59.5	78	74.3	21	72.4	37	74.0
1	12	10.8	9	8.6	3	10.3	3	6.0
2	5	4.5	2	1.9	0	0.0	2	4.0
3	5	4.5	5	4.8	2	6.9	2	4.0
4	4	3.6	2	1.9	0	0.0	0	0.0
5	0	0.0	1	1.0	0	0.0	0	0.0
6	5	4.5	0	0.0	1	3.4	0	0.0
7	1	0.9	3	2.9	1	3.4	3	6.0
8	2	1.8	0	0.0	0	0.0	1	2.0
9	2	1.8	1	1.0	1	3.4	0	0.0
10	1	0.9	0	0.0	0	0.0	0	0.0
12	0	0.0	2	1.9	0	0.0	2	4.0
16	0	0.0	1	1.0	0	0.0	0	0.0
22	0	0.0	1	1.0	0	0.0	0	0.0
31	1	0.9	0	0.0	0	0.0	0	0.0
36	2	1.8	0	0.0	0	0.0	0	0.0
38	1	0.9	0	0.0	0	0.0	0	0.0
39	1	0.9	0	0.0	0	0.0	0	0.0
42	1	0.9	0	0.0	0	0.0	0	0.0
48	1	0.9	0	0.0	0	0.0	0	0.0
125	1	0.9	0	0.0	0	0.0	0	0.0

Note. Older male offenders were at least 15 years of age.

Table 40

Days Charge Free: Statistics Based on the Assigned Curfew Time for Data Set and for Older Male Offenders

Statistic	All offenders (n = 216)		Older males (n = 79)	
	Before 7:38 pm	After 7:38 pm	Before 7:38 pm	After 7:38 pm
Minimum days	23.0	15.0	103.0	15.0
Maximum days	2544.0	2521.0	2544.0	2521.0
Mean days	1452.3	1688.1	1589.0	1667.3
Maximum years	7.0	6.9	7.0	6.9
Mean years	4.0	4.6	4.4	4.6
SD (days)	849.7	831.1	784.8	875.9

Note. Older male offenders were at least 15 years of age.

Table 41

Path Analysis for Predictive Model of Recidivism Based on the Assigned Curfew Time for Entire Data Set

Construct/Indicator	Construct/ Indicator	Standardized regression coefficient	Unstandardized regression coefficient	Standard error	Critical value	P
Recidivism	Curfew Time	0.21	0.00	0.00	3.16	0.002
Severity of new charge	Recidivism	0.95	1.00	0.00	0.00	
Reoffended (y/n)	Recidivism	0.99	0.24	0.01	39.56	*
Days charge free	Recidivism	0.95	416.77			*
Reoffended (y/n)	Recidivism	0.90	1.13	0.05	23.74	*
Days charge free	Recidivism	0.89	0.97	0.03	37.24	*

Note. * indicates significance at .001 level (two-tailed).

Table 42

Goodness of Fit Statistics for the Predictive Model of Recidivism Based on the Assigned Curfew
Time for Entire Data Set and for Older Male Offenders

Metric	Ideal model	All data (n = 216)	Older males (n = 79)
Indicators		6	6
Chi-Square	Low	2.4	9.7
Degrees of freedom		7	7
Relative Chi Square	<5	0.349	1.39
<i>p</i>	>.05	0.931	0.204
GFI	>.9	0.996	0.963
AGFI	>.9	0.989	0.89
CFI	>.9	1.000	0.996
RMSEA	<.05	0.000	0.071
Interpretation	Good fit*	Good fit*	Poor fit

Note. * Good fit found because all goodness of fit measures indicate a good fit between the model and the data. Older was defined as at least 15 years of age when charged.

Table 43

Severity of New Charge: Frequencies and Percentages Based on the Assignment of No-Contact
Order for Entire Data Set

Category	Not assigned (n = 176)		Assigned (n = 42)	
	No.	%	No.	%
No new charge	122	69.3	24	57.1
Status offense/other	2	1.1	0	0.0
Second degree misdemeanor	2	1.1	1	2.4
First degree misdemeanor	18	10.2	3	7.1
Third degree felony	12	6.8	3	7.1
Second degree felony	16	9.1	9	21.4
First degree felony	4	2.3	2	4.8

Table 44

Most Severe Sanction: Frequencies and Percentages Based on the Assignment of No-Contact

Order for Entire Data Set

Most Severe Sanction	Not assigned (n = 176)		Assigned (n = 42)	
	No.	%	No.	%
(0.0) No new charge	122	69.3	24	57.1
(1) Non-file	6	3.4	0	0.0
(2) Nolle pros	3	1.7	3	7.1
(3) Judicial warning	7	4.0	1	2.4
(4) Probation	25	14.2	8	19.0
(5) Secure detention	3	1.7	2	4.8
(6) Committed to DJJ (L)	1	0.6	0	0.0
(7) Committed to DJJ (M)	2	1.1	1	2.4
(8) Committed to DJJ (H)	3	1.7	2	4.8
(9) Adult jail	2	1.1	0	0.0
(10) Adult prison	2	1.1	1	2.4

Note. DJJ = Department of Juvenile Justice; L = Low security; M = Medium security; H = High security.

Table 45

Crime Level Change: Frequencies and Percentages Based on the Assignment of No-Contact

Order for Entire Data Set

Crime level change	Not assigned (n = 176)		Assigned (n = 42)	
	No.	%	No.	%
-5	1	0.6	0	0.0
-4	12	6.8	4	9.5
-3	46	26.1	14	33.3
-2	61	34.7	5	11.9
-1	7	4.0	2	4.8
0	13	7.4	0	0.0
1	17	9.7	5	11.9
2	9	5.1	5	11.9
3	10	5.7	7	16.7

Table 46

Number of New Charges: Frequencies and Percentages Based on the Assignment of No-Contact Order for Entire Data Set

Number of new charges	Not assigned (n = 176)		Assigned (n = 42)	
	No.	%	No.	%
0	122	69.3	24	57.1
1	18	10.2	3	7.1
2	6	3.4	1	2.4
3	5	2.8	5	11.9
4	4	2.3	2	4.8
5	1	0.6	0	0.0
6	3	1.7	2	4.8
7	3	1.7	1	2.4
8	1	0.6	1	2.4
9	3	1.7	1	2.4
10	0	0.0	0	0.0
12	2	1.1	0	0.0
16	1	0.6	0	0.0
22	1	0.6	0	0.0
31	1	0.6	0	0.0
36	2	1.1	0	0.0
38	1	0.6	0	0.0
39	0	0.0	1	2.4
42	1	0.6	0	0.0
48	1	0.6	0	0.0
125	0	0.0	1	2.4

Table 47

Days Charge Free: Statistics Based on the Assignment of No-Contact Order for Entire Data Set

Statistic	Not assigned	Assigned
Minimum days	15.0	35.0
Maximum days	2544.0	2467.0
Mean days	1619.2	1378.2
Maximum years	7.0	6.8
Mean years	4.4	3.8
SD (days)	824.4	912.5

Table 48

Goodness of Fit Statistics for Predictive Model of Recidivism Based on No-Contact Order for Entire Data Set and Gender, Age, and Drug Offender Subgroups

Metric	Ideal model	All data (n = 218)	Males (n = 133)	Females (n = 85)	>= 15 (n = 134)	< 15 (n = 84)	Males >=15 (n = 81)	Females >15 (n = 53)	Drug offenders (n = 54)
Chi-square	Low	3.7	5.8	22.0	5.7	3.6	12.9	28.8	1.1
Degrees of freedom		7	7	7	7	7	7	7	7
Relative Chi-square	<5	0.521	0.827	3.142	0.818	0.519	1.839	4.109	0.16
<i>p</i>	>.05	0.819	0.564	0.003	0.572	0.821	0.075	0.000	0.993
GFI	>.9	0.994	0.986	0.928	0.986	0.986	0.951	0.859	0.993
AGFI	>.9	0.983	0.957	0.784	0.959	0.957	0.854	0.576	0.979
CFI	>.9	1.000	1.000	0.978	1.000	1.000	0.992	0.952	1.000
RMSEA	<.05	0.000	0.000	0.160	0.000	0.000	0.102	0.245	0.000
Regression weight **	High	0.108	0.015	0.281	0.102	0.135	0.026	0.244	-0.110
P **	<.05	0.109	0.826	0.008	0.241	0.217	0.817	0.072	0.420
Interpretation	Good fit*	Good fit*	Good fit*	Poor fit	Good fit*	Good fit*	Poor fit	Poor fit	Good fit*

* Good fit found because all goodness of fit measures indicate a good fit between the model and the data.

** Regression weight (standardized) and *p* are reported between the predictor No-contact Order and the construct of Recidivism.

Table 49

Chi-Square and Mann-Whitney Tests for Distributions of Indicators of Recidivism Based on the Assignment of No-Contact

Order for Entire Data Set

Indicator	Test conducted	Significance	Decision	Assumption test cells (%)	Chi-Square table
Crime level change	Chi-Square	0.021	Reject Null	8 (44.4)	2x9
Crime level change sign*	Chi-Square	0.012	Reject Null	0 (0.0)	2x2
Days charge free	Mann-Whitney U test	0.047	Reject Null		
Number of new charges	Mann-Whitney U test	0.099	Retain Null		
Reoffended (y/n)	Chi-Square	0.132	Retain Null	0 (0.0)	2x2
Severity of new charge	Chi-Square	0.302	Retain Null	9 (64.3)	2x7
Severity of new charge category*	Chi-Square	0.096	Retain Null	0 (0.0)	2x3
Most severe sanction	Chi-Square	0.371	Retain Null	17 (77.3)	2x11
Most severe sanction category*	Chi-Square	0.161	Retain Null	0 (0.0)	2x2

Note. The predictor variable for each test was the assignment of no-contact orders; assumption test, which applied only to the Chi-square tests, indicates the number of cells and percentage of cells with expected count less than 5; the significance level for these tests was .05; * indicated this variable was modified.

Table 50

Frequency and Percentage of Assignment of Essays and Apology Letters for the Entire Data Set

Sanction	Not assigned		Assigned	
	No.	%	No.	%
Apology letters	189	86.7	29	13.3
Essays	166	76.1	52	23.9

Table 51

Severity of New Charge: Frequencies and Percentages Based on Assignment of Apology Letters and Essays for the Entire Data

Set

Most severe sanction	Letters				Essay			
	Not assigned (n = 29)		Assigned (n = 189)		Not assigned (n = 52)		Assigned (n = 166)	
	No.	%	No.	%	No.	%	No.	%
(0.0) No new charge	19	65.5	127	67.2	34	65.4	112	67.5
(1) Non-file	1	3.4	5	2.6	1	1.9	5	3.0
(2) Nolle pros	1	3.4	5	2.6	1	1.9	5	3.0
(3) Judicial warning	1	3.4	7	3.7	1	1.9	7	4.2
(4) Probation	4	13.8	29	15.3	8	15.4	25	15.1
(5) Secure detention	2	6.9	3	1.6	2	3.8	3	1.8
(6) Committed to DJJ (L)	0	0.0	1	0.5	2	3.8	1	0.6
(7) Committed to DJJ (M)	1	3.4	2	1.1	2	3.8	1	0.6
(8) Committed to DJJ (H)	0	0.0	5	2.6	1	1.9	3	1.8
(9) Adult jail	0	0.0	2	1.1	0	0.0	1	0.6
(10) Adult prison	0	0.0	3	1.6	0	0.0	3	1.8

Note. DJJ = Department of Juvenile Justice; L = Low security; M = Medium security; H = High security.

Table 52

Most Severe Sanction: Frequencies and Percentages Based on Assignment of Apology Letters and Essays for the Entire Data

Set

Sanction	Letters				Essay			
	Not assigned		Assigned		Not assigned		Assigned	
	No.	%	No.	%	No.	%	No.	%
(0.0) No new charge	19	65.5	127	67.2	34	65.4	112	67.5
(1) Non-file	1	3.4	5	2.6	1	1.9	5	3.0
(2) Nolle pros	1	3.4	5	2.6	1	1.9	5	3.0
(3) Judicial warning	1	3.4	7	3.7	1	1.9	7	4.2
(4) Probation	4	13.8	29	15.3	8	15.4	25	15.1
(5) Secure detention	2	6.9	3	1.6	2	3.8	3	1.8
(6) Committed to DJJ (L)	0	0.0	1	0.5	2	3.8	1	0.6
(7) Committed to DJJ (M)	1	3.4	2	1.1	2	3.8	1	0.6
(8) Committed to DJJ (H)	0	0.0	5	2.6	1	1.9	3	1.8
(9) Adult jail	0	0.0	2	1.1	0	0.0	1	0.6
(10) Adult prison	0	0.0	3	1.6	0	0.0	3	1.8

Note. DJJ = Department of Juvenile Justice; L = Low security; M = Medium security; H = High security.

Table 53

Crime Level Change: Frequencies and Percentages Based on Assignment of Apology Letters and Essays for the Entire Data Set

Set

Crime level change	Letters				Essay			
	Not assigned (n = 29)		Assigned (n = 189)		Not assigned (n = 52)		Assigned (n = 166)	
	No.	%	No.	%	No.	%	No.	%
-5	0	0.0	1	0.5	0	0.0	1	0.6
-4	2	6.9	14	7.4	4	7.7	12	7.2
-3	12	41.4	48	25.4	12	23.1	48	28.9
-2	4	13.8	62	32.8	17	32.7	49	29.5
-1	3	10.3	6	3.2	1	1.9	8	4.8
0	5	17.2	8	4.2	4	7.7	9	5.4
1	3	10.3	19	10.1	7	13.5	15	9.0
2	0	0.0	14	7.4	3	5.8	11	6.6
3	0	0.0	17	9.0	4	7.7	13	7.8

Table 54

Number of New Charges: Frequencies and Percentages Based on Assignment of Apology Letters and Essays for the Entire

Data Set

Number of new charges	Letters				Essay			
	Not assigned (n = 29)		Assigned (n = 189)		Not assigned (n = 52)		Assigned (n = 166)	
	No.	%	No.	%	No.	%	No.	%
0	19	65.5	127	67.2	34	67.5	112	67.5
1	3	10.3	18	9.5	4	10.2	17	10.2
2	2	6.9	5	2.6	3	2.4	4	2.4
3	2	6.9	8	4.2	1	5.4	9	5.4
4	1	3.4	5	2.6	1	3.0	5	3.0
5	0	0.0	1	0.5	0	0.6	1	0.6
6	1	3.4	4	2.1	1	2.4	4	2.4
7	0	0.0	4	2.1	0	2.4	4	2.4
8	0	0.0	2	1.1	0	1.2	2	1.2
9	0	0.0	3	1.6	2	0.6	1	0.6
10	0	0.0	1	0.5	1	0.0	0	0.0
12	1	3.4	1	0.5	2	0.0	0	0.0
16	0	0.0	1	0.5	1	0.0	0	0.0
22	0	0.0	1	0.5	0	0.6	1	0.6
31	0	0.0	1	0.5	0	0.6	1	0.6
36	0	0.0	2	1.1	0	1.2	2	1.2
38	0	0.0	1	0.5	0	0.6	1	0.6
39	0	0.0	1	0.5	1	0.6	1	0.6
42	0	0.0	1	0.5	0	0.6	1	0.6
48	0	0.0	1	0.5	0	0.0	0	0.0
125	0	0.0	1	0.5	1	0.0	0	0.0

Table 55

Days Charge Free Statistics Based on Assignment of Apology Letters and Essays for the Entire

Data Set

Statistic	Letters		Essays	
	Not assigned (n = 29)	Assigned (n = 189)	Not assigned (n = 52)	Assigned (n = 166)
Minimum days	35.0	15.0	54.0	15.0
Maximum days	2481.0	2544.0	2544.0	2521.0
Mean days	1561.9	1574.4	1586.3	1568.5
Maximum years	6.8	7.0	7.0	6.9
Mean years	4.3	4.3	4.4	4.3
SD (days)	948.7	831.0	913.6	825.6

Table 56

Goodness of Fit Statistics for Predictive Model of Recidivism Based on Assignment Apology Letters and Essays for the Entire Data Set, All Females, and Older Female

Metric	Ideal model	Data set (n = 218)	All females (n = 85)	Older females (n = 53)
Indicators		6	6	6
Chi-Square	Low	16.3	23.5	24.6
Degrees of freedom		11	11	11
Relative Chi Square	<5	1.477	2.138	2.233
<i>p</i>	>.05	0.132	0.015	0.011
GFI	>.9	0.979	0.938	0.898
AGFI	>.9	0.948	0.842	0.74
CFI	>.9	0.997	0.982	0.969
RMSEA	<.05	0.047	0.116	0.154
Interpretation	Good fit	Good fit *	Poor fit	Poor fit

Note. Older is defined as at least 15 years of age; * Good fit determined because all goodness of fit measures indicated a good fit between the model and the data.

Table 57

Chi-Square and Mann-Whitney Tests for Distributions of Indicators of Recidivism Based on Assignment Apology Letters and Essays for the Entire Data Set

Predictor	Indicator	Test Conducted	Significance	Decision	Assumption test cells (%)	Chi-Square Table
Apology letters	Crime level change	Chi-Square	0.009	Reject Null	8 (44.4)	2x9
	Crime level change sign*	Chi-Square	0.099	Retain Null	0 (0.0)	2x2
	Days charge free	Mann-Whitney U test	0.373	Retain Null		
	Number of new charges	Mann-Whitney U test	0.946	Retain Null		
	Reoffended (y/n)	Chi-Square	0.858	Retain Null	0 (0.0)	2x2
	Severity of new charge	Chi-Square	0.210	Retain Null	8 (57.1)	2x7
	Severity of new charge category*	Chi-Square	0.049	Reject Null	1 (16.7)	2x3
	Most severe sanction	Chi-Square	0.816	Retain Null	16 (72.7)	2x11
	Most severe sanction category*	Chi-Square	1.000	Retain Null	0 (0.0)	2x2
Essays	Crime level change	Chi-Square	0.940	Retain Null	7 (38.9)	2x9
	Crime level change sign*	Chi-Square	0.751	Retain Null	0 (0.0)	2x2
	Days charge free	Mann-Whitney U test	0.298	Retain Null		
	Number of new charges	Mann-Whitney U test	0.592	Retain Null		
	Reoffended (y/n)	Chi-Square	0.912	Retain Null	0 (0.0)	2x2
	Severity of new charge	Chi-Square	0.635	Retain Null	7 (50.0)	2x7
	Severity of new charge category*	Chi-Square	0.341	Retain Null	0 (0.0)	2x3
	Most severe sanction	Chi-Square	0.687	Retain Null	17 (77.3)	2x11
	Most severe sanction category*	Chi-Square	0.333	Retain Null	0 (0.0)	2x2

Note. Assumption test, which applied only to the Chi-square tests, indicates the number of cells (and percentage of cells) with expected count less than 5; the significance level for these tests was .05; * indicates this variable was modified; (n = 218).

Table 58

Chi-Square and Mann-Whitney Tests for Distributions of Indicators of Recidivism Based on Assignment Apology Letters and Essays for Females at Least 15 Years of Age

Predictor	Indicator	Test conducted	Significance	Decision	Assumption test cells (%)	Chi-Square Table
Apology letters	Crime level change	Chi-Square	0.211	Retain Null	12 (85.7)	2x7
	Crime Level change sign*	Chi-Square	0.055	Retain Null	3 (50.0)	2x3
	Days charge free	Mann-Whitney U test	0.274	Retain Null		
	Number of new charges	Mann-Whitney U test	0.710	Retain Null		
	Reoffended (y/n)	Chi-Square	0.331	Retain Null	1 (25.0)	2x2
	Severity of new charge	Chi-Square	0.385	Retain Null	10 (83.3)	2x6
	Severity of new charge category*	Chi-Square	0.143	Retain Null	3 (50.0)	2x3
	Most severe sanction	Chi-Square	0.132	Retain Null	8 (80.0)	2x5
	Most severe sanction category*	Chi-Square	0.055	Retain Null	1 (25.0)	2x2
Essays	Crime level change	Chi-Square	0.230	Retain Null	12 (87.5)	2x7
	Crime level change sign*	Chi-Square	0.032	Retain Null	3 (50.0)	2x3
	Days charge free	Mann-Whitney U test	0.733	Retain Null		
	Number of new charges	Mann-Whitney U test	0.316	Retain Null		
	Reoffended (y/n)	Chi-Square	0.481	Retain Null	1 (25.0)	2x2
	Severity of new charge	Chi-Square	0.552	Retain Null	10 (83.3)	2x6
	Severity of new charge category*	Chi-Square	0.504	Retain Null	3 (50.0)	2x3
	Most severe sanction	Chi-Square	0.071	Retain Null	8 (80.0)	2x5
	Most severe sanction category*	Chi-Square	0.285	Retain Null	1 (25.0)	2x2

Note. Assumption test, which applied only to the Chi-square tests, indicates the number of cells (and percentage of cells) with expected count less than 5; the significance level for these tests was .05; * indicates this variable was modified; (n=53).

Table 59

Pearson Correlations Between Indicators of Offender Risk Propensity

	Charge type	Drug offender	Race	Age	Gender
Drug offender	-.371**				
Race	.104	-.162*			
Age	-.003	.204**	-.089		
Gender	-.136*	-.100	.096	.034	
Suspended	.026	.157*	.176*	.024	-.014

* indicates the correlation is significant at the 0.05 level (2-tailed).

** indicates the correlation is significant at the 0.01 level (2-tailed).

Table 60

Pearson Correlations Between Indicators of Multiplicity of Sanctions

	No-contact	Anger management	Drug testing	Restitution	Apology letters	Community service
Anger management	-.029					
Drug testing	.154*	-.118				
Restitution	.044	-.081	-.069			
Apology letters	.123	-.075	-.037	.028		
Community service	.016	.100	.173*	-.215**	.080	
Boot camp	.034	.181**	.129	.051	-.062	-.010

* indicates the correlation is significant at the 0.05 level (2-tailed).

** indicates the correlation is significant at the 0.01 level (2-tailed).

Table 61

Pearson Correlation Between Indicators of the Interaction Construct, Offender Risk Propensity, and Multiplicity of Sanctions

	No-contact & Drug Offender	No-contact & Charge Type	No-contact & Age	No-contact & Suspended
No-contact & Charge Type	.402**			
No-contact & Age	.579**	.908**		
No-contact & Suspended	.525**	.763**	.836**	
Drug Test & Charge Type	.242**	.215**	.184**	.126
Drug Test & Drug Offender	.505**	.075	.170**	.163*
Drug Test & Age	.331**	.124	.197**	.154*
Drug Test & Suspended	.314**	.095	.174**	.258**
Community Service & Charge Type	-.040	.125	.053	.016
Community Service & Drug Offender	.190**	.001	.004	.004
Community Service & Age	-.048	.036	.010	-.003
Community Service & Suspended	-.016	.043	.042	.140*
Letters & Charge Type	.007	.202**	.094	.094
Letters & Drug Offender	.492**	.063	.156**	.151*
Letters & Age	.151**	.069	.131	.108
Letters & Suspended	.146**	.121	.148**	.326**

* indicates the correlation is significant at the 0.05 level (2-tailed).

** indicates the correlation is significant at the 0.01 level (2-tailed).

Bold indicates correlation between indicators that do not share a common indicator.

	Drug Test & Charge Type	Drug Test & Drug Offender	Drug Test & Age	Drug Test & Suspended
No-contact & Charge Type				
No-contact & Age				
No-contact & Suspended				
Drug Test & Charge Type				
Drug Test & Drug Offender	.377**			
Drug Test & Age	.861**	.639**		
Drug Test & Suspended	.630**	.586**	.785**	
Community Service & Charge Type	.307**	-.019	.154**	.022
Community Service & Drug Offender	.277**	.606**	.364**	.308**
Community Service & Age	.231**	.063	.202**	.084
Community Service & Suspended	.109	.087	.120	.289**
Letters & Charge Type	.104	-.157*	-.131**	-.159**
Letters & Drug Offender	.303**	.864**	.518**	.465**
Letters & Age	.008	.139*	.048	.007
Letters & Suspended	.014	.162*	.069	.356**

* indicates the correlation is significant at the 0.05 level (2-tailed).

** indicates the correlation is significant at the 0.01 level (2-tailed).

Bold indicates correlation between indicators that do not share a common indicator.

	Community Service & Charge Type	Community Service & Drug Offender	Community Service & Age	Community Service & Suspended
No-contact & Charge Type				
No-contact & Age				
No-contact & Suspended				
Drug Test & Charge Type				
Drug Test & Drug Offender				
Drug Test & Age				
Drug Test & Suspended				
Community Service & Charge Type				
Community Service & Drug Offender	.277**			
Community Service & Age	.889**	.426**		
Community Service & Suspended	.662**	.365**	.718**	
Letters & Charge Type	.214**	.013	.046	.012
Letters & Drug Offender	.012	.691**	.108	.119
Letters & Age	.050	.189**	.126	.027
Letters & Suspended	.007	.138*	.018	.404**

* indicates the correlation is significant at the 0.05 level (2-tailed).

** indicates the correlation is significant at the 0.01 level (2-tailed).

Bold indicates correlation between indicators that do not share a common indicator.

	Letters & Charge Type	Letters & Drug Offender	Letters & Age
No-contact & Charge Type			
No-contact & Age			
No-contact & Suspended			
Drug Test & Charge Type			
Drug Test & Drug Offender			
Drug Test & Age			
Drug Test & Suspended			
Community Service & Charge Type			
Community Service & Drug Offender			
Community Service & Age			
Community Service & Suspended			
Letters & Charge Type			
Letters & Drug Offender	-.080		
Letters & Age	.654**	.277**	
Letters & Suspended	.356**	.224**	.445**

* indicates the correlation is significant at the 0.05 level (2-tailed).

** indicates the correlation is significant at the 0.01 level (2-tailed).

Bold indicates correlation between indicators that do not share a common indicator.

Table 62

Goodness of Fit Statistics for Measurement and Predictive Models

Metric	Ideal model	Measurement model of offender risk propensity	Predictive model of recidivism based on offender risk propensity	Measurement model of multiplicity of sanctions	Predictive model of recidivism based on multiplicity of sanctions	Measurement model of interaction construct	Predictive model of recidivism based on interaction construct
Chi-Square	Low	11.8	101.4	17.3	91.8	1.12	23.1
Degrees of freedom		8	31	16	58	1	22
Relative Chi-square	<5	1.472	3.27	1.082	1.582	1.117	1.050
<i>p</i>	>.05	0.161	0.000	0.366	0.003	0.290	0.400
GFI	>.9	0.983	0.924	0.980	0.993	0.997	0.978
AGFI	>.9	0.956	0.866	0.955	0.903	0.974	0.955
CFI	>.9	0.940	0.962	0.959	0.981	1.000	1.000
RMSEA	<.05	0.047	0.102	0.020	0.052	0.024	0.015
Interpretation	Good fit *	Good fit *	Poor fit	Good fit *	Poor fit	Good fit *	Good fit *

Note. * Good fit found because all goodness of fit measures indicate a good fit between the model and the data; n = 218

Table 63

Path Analysis for Offender Risk Propensity Measurement Model

Indicator	Standardized regression coefficient	Unstandardized regression coefficient	Standard error	Critical value
Gender	0.068	1.252	0.88	1.422
Suspended	0.081	1.398	0.994	-1.405
Race	0.060	1		
Age	0.099	6.475	4.025	-1.609
Drug offender	2.025	33.038	67.705	-0.488
Charge type	0.189	6.539	3.672	1.781

Note. The construct examined is Offender Risk Propensity; none of the regression weights is significant at the 0.05 level (two-tailed).

Table 64

Path Analysis for Multiplicity of Sanctions Measurement Model

Indicator	Standardized regression coefficient	Unstandardized regression coefficient	Standard error	Critical value
No-contact	0.479	1		
Anger management	0.046	0.104	0.667	0.155
Drug testing	-2.486	-6.477	9.305	-0.696
Restitution	0.040	0.046	0.295	0.157
Curfew time	0.041	45.237	289.439	-0.156
Boot camp	0.104	0.214	1.384	0.154
Community service	-0.073	-0.190	1.207	-0.157
Apology letters	0.034	0.059	0.381	0.156

Note. The construct examined is Multiplicity of Sanctions; none of the regression weights is significant at the 0.05 level (two-tailed).

Table 65

Goodness of Fit Statistics for Predictive Model of Recidivism Based on Assignment of Apology Letters and Essays for Entire Data Set

Metric	Ideal model	Main effect model	Mediating effect model	Moderating effect model
Chi-Square	Low	108.0	108.0	342.6
Degrees of freedom		47	47	93
Relative Chi-square	<5	2.298	2.298	3.684
<i>p</i>	>.05	0.000	0.000	0.000
GFI	>.9	0.927	0.927	0.844
AGFI	>.9	0.878	0.878	0.771
CFI	>.9	0.967	0.967	0.931
RMSEA	<.05	0.087	0.087	0.113
Interpretation	Good fit *	Poor fit	Poor fit	Poor fit

Note. * Good fit found because all goodness of fit measures indicate a good fit between the model and the data; n = 218.

Table 66

Chi-Square Tests for the Distribution of Specific Sanction Assignment Across the Categories of Reoffended for the Entire Data Set

Predictor	Significance	Decision	Assumption test cells (%)
Jail tours	0.471	Retain Null	0 (0.0)
Community service	0.961	Retain Null	0 (0.0)
Home chores	0.136	Retain Null	0 (0.0)
Anger management	0.059	Retain Null	0 (0.0)
School progress	0.654	Retain Null	0 (0.0)
Boot camp	0.724	Retain Null	0 (0.0)

Note: Results presented are from a 2x2 Chi-square test; assumption test indicates the number of cells and percentage of cells with an expected count less than 5; the significance level for these tests was .05; n = 218.

Table 67

Chi-Square Tests for the Distribution of Specific Sanction Assignment Across the Categories of Reoffended for Older Offenders

Predictor	Significance	Decision	Assumption test cells (%)
Jail tours	0.809	Retain Null	0 (0.0)
Community service	1.000	Retain Null	0 (0.0)
Home chores	0.802	Retain Null	0 (0.0)
Anger management	0.097	Retain Null	0 (0.0)
School progress	0.323	Retain Null	0 (0.0)
Boot camp	0.613	Retain Null	0 (0.0)

Note: Results presented are from a 2x2 Chi-square test; assumption test indicates the number of cells and percentage of cells with an expected count less than 5; the significance level for these tests was .05; older defined as at least 15 years of age; n = 134.

Table 68

Chi-Square Tests for the Distribution of Specific Sanction Assignment Across the Categories of Reoffended for Older Female Offenders

Predictor	Significance	Decision	Assumption test cells (%)
Jail tours	0.939	Retain Null	1 (25.0)
Community service	0.879	Retain Null	1 (25.0)
Home chores	0.713	Retain Null	1 (25.0)
Anger management	0.184	Retain Null	1 (25.0)
School progress	0.618	Retain Null	1 (25.0)
Boot camp	1.000	Retain Null	1 (25.0)

Note: Results presented are from a 2x2 Chi-square test; assumption test indicates the number of cells and percentage of cells with an expected count less than 5; the significance level for these tests was .05; older defined as at least 15 years of age; n = 53.

Table 69

Chi-Square Tests for the Distribution of Specific Sanction Assignment Across the Categories of Reoffended for Older Male Offenders

Predictor	Significance	Decision	Assumption test cells (%)
Jail tours	0.992	Retain Null	0 (0.0)
Community service	0.647	Retain Null	0 (0.0)
Home chores	1.000	Retain Null	1 (25.0)
Anger management	0.496	Retain Null	1 (25.0)
School progress	0.663	Retain Null	1 (25.0)
Boot camp	0.496	Retain Null	1 (25.0)

Note: Results presented are from a 2x2 Chi-square test; assumption test indicates the number of cells and percentage of cells with an expected count less than 5; the significance level for these tests was .05; older defined as at least 15 years of age; n = 81.

Table 70

Chi-Square Tests for the Distribution of Specific Sanction Assignment Across the Categories of Reoffended for Drug Offenders

Predictor	Significance	Decision	Assumption test cells (%)
Jail tours	1.000	Retain Null	0 (0.0)
Community service	0.439	Retain Null	0 (0.0)
Home chores	0.981	Retain Null	1 (25.0)
Anger management	0.860	Retain Null	1 (25.0)
School progress	0.806	Retain Null	1 (25.0)
Boot camp	0.251	Retain Null	1 (25.0)

Note: Results presented are from a 2x2 Chi-square test; assumption test indicates the number of cells and percentage of cells with an expected count less than 5; the significance level for these tests was .05; drug offender defined as having had a drug-related charge or admitted drug use; n = 54.

APPENDIX B: FIGURES

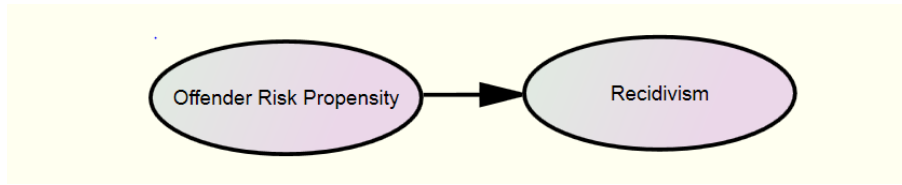


Figure 1: Recidivism Based on Offender Risk Propensity

This model illustrates the predictive relationship the latent construct of Offender Risk Propensity has on the latent construct of Recidivism.

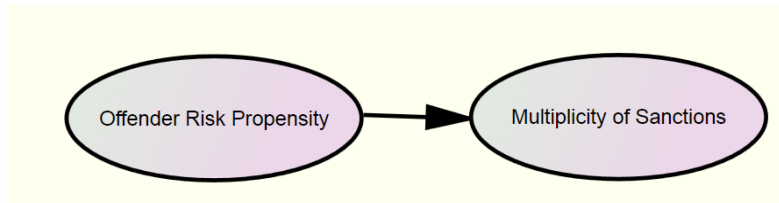


Figure 2: Multiplicity of Sanctions Based on Offender Risk Propensity

This model illustrates predictive relationship the latent construct of Offender Risk Propensity has on the latent construct of Multiplicity of Sanctions.

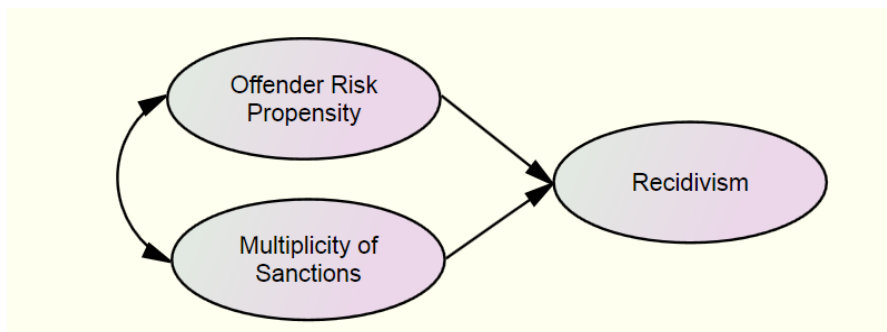


Figure 3: Simplified Main Effect Model

This model illustrates the main effect predictive relationship that the latent constructs Offender Risk Propensity and Multiplicity of Sanctions have on the latent construct of Recidivism.

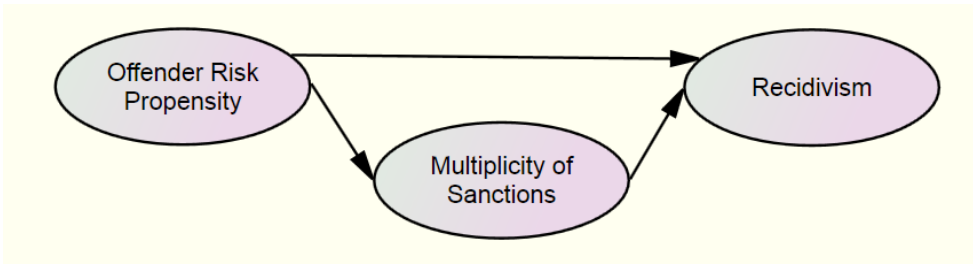


Figure 4: Simplified Mediating Effect Model

This model illustrates the mediating effect predictive relationship that the latent constructs Offender Risk Propensity and Multiplicity of Sanctions have on the latent construct of Recidivism.

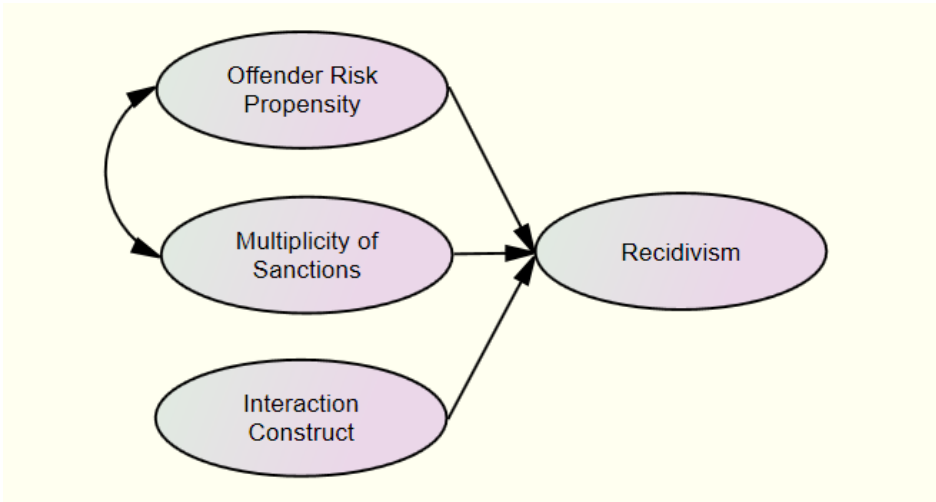


Figure 5: Simplified Moderating Effect Model

This model illustrates the moderating effect predictive relationship that the latent constructs Offender Risk Propensity and Multiplicity of Sanctions have on the latent construct of Recidivism.

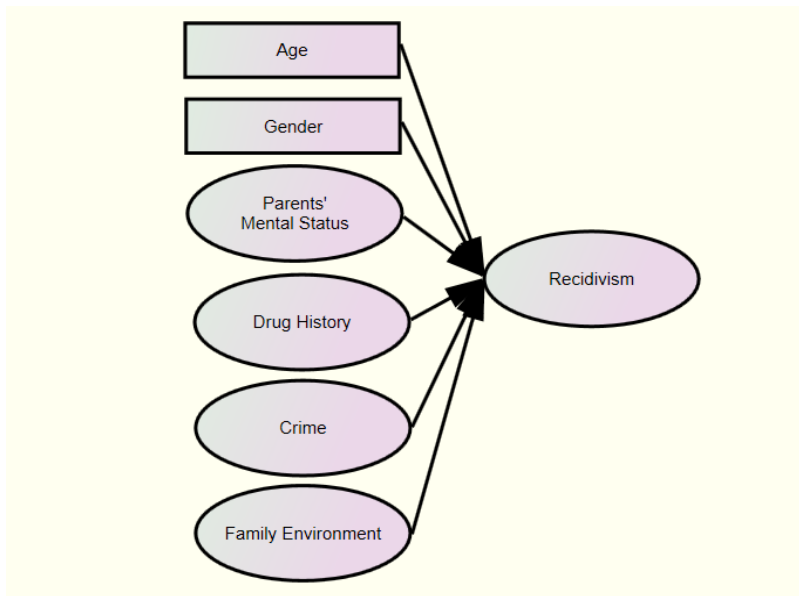


Figure 6: Offender Characteristics and Recidivism

This model illustrates the relationships between six offender characteristics and recidivism discussed in criminal justice literature.

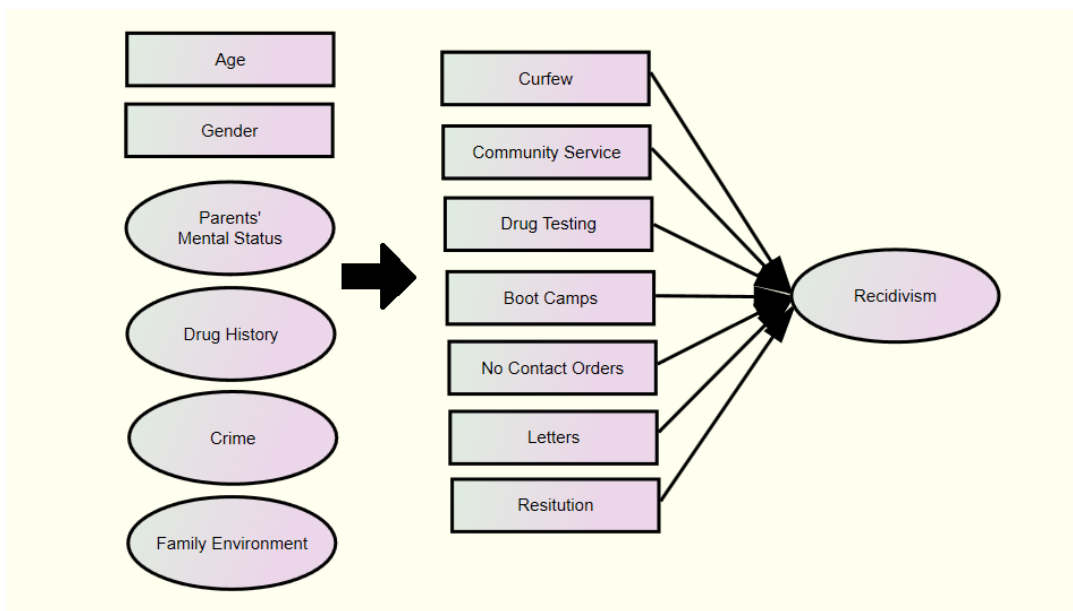


Figure 7: Recidivism as a Function of Sanctions Based on Offender Characteristics

This model illustrates the relationships among offender characteristics, sanctions, and recidivism not examined in the criminal justice literature.

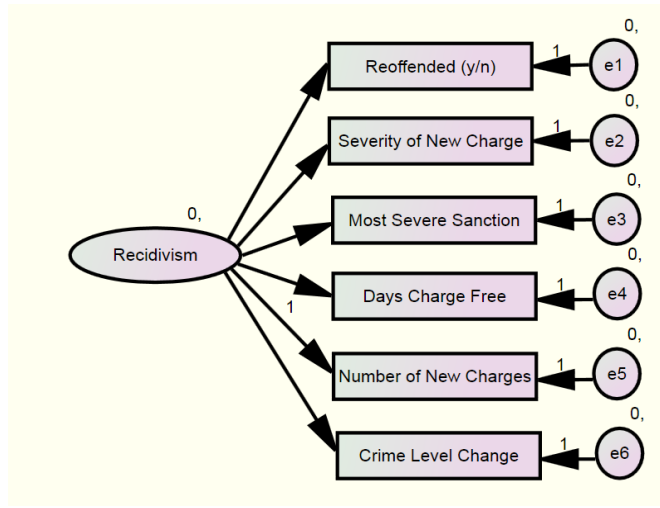


Figure 8: Measurement Model of the Latent Construct of Recidivism

This model illustrates the latent construct of Recidivism as a function of six indicators.

This model is tested in Hypothesis 1.

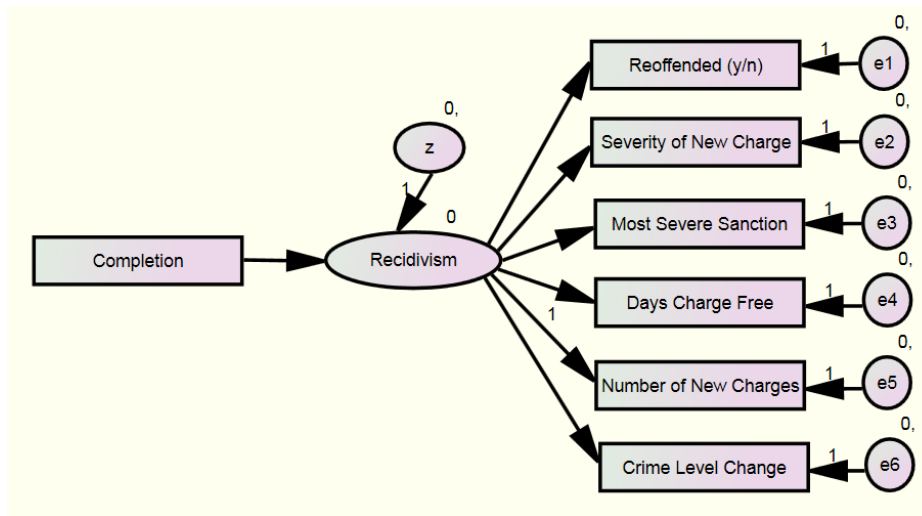


Figure 9: Predictive Model of Recidivism Based on NRJP Completion Status

This model illustrates the relationship between completion status of the NRJP and the latent construct of Recidivism as a function of six indicators. This model was proposed to test Hypothesis 2.

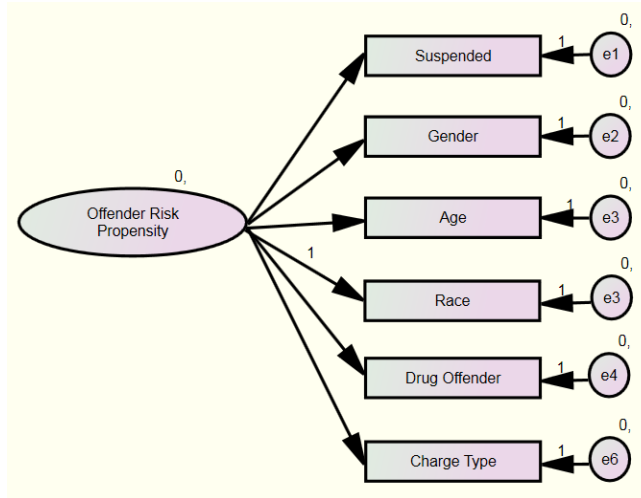


Figure 10: Measurement Model of Offender Risk Propensity

This model illustrates the construct of Offender Risk Propensity as a function of the indicators.

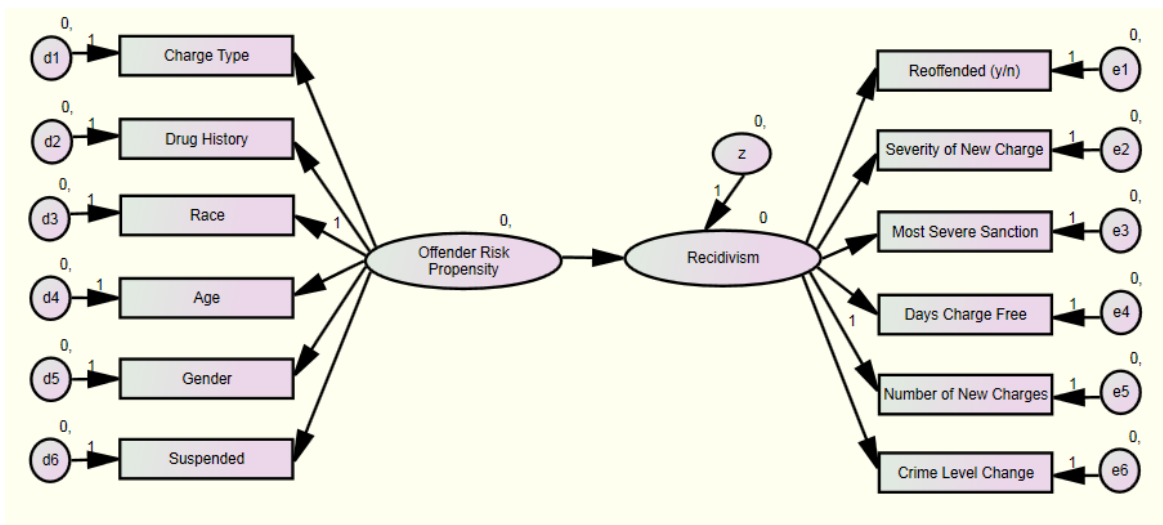


Figure 11: Predictive Model of Recidivism Based on Offender Risk Propensity

This model illustrates the relationship among the constructs of Offender Risk Propensity and Recidivism and their indicators.

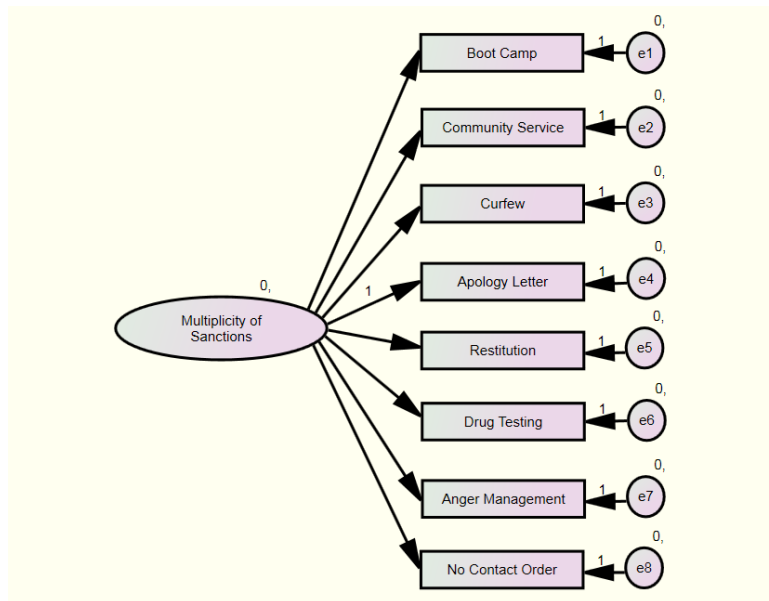


Figure 12: Measurement Model of Multiplicity of Sanctions

This model illustrates the construct of Multiplicity of Sanctions as a function of eight sanctions simultaneously. This model is an example only and is not tested through structural equation modeling in this study.

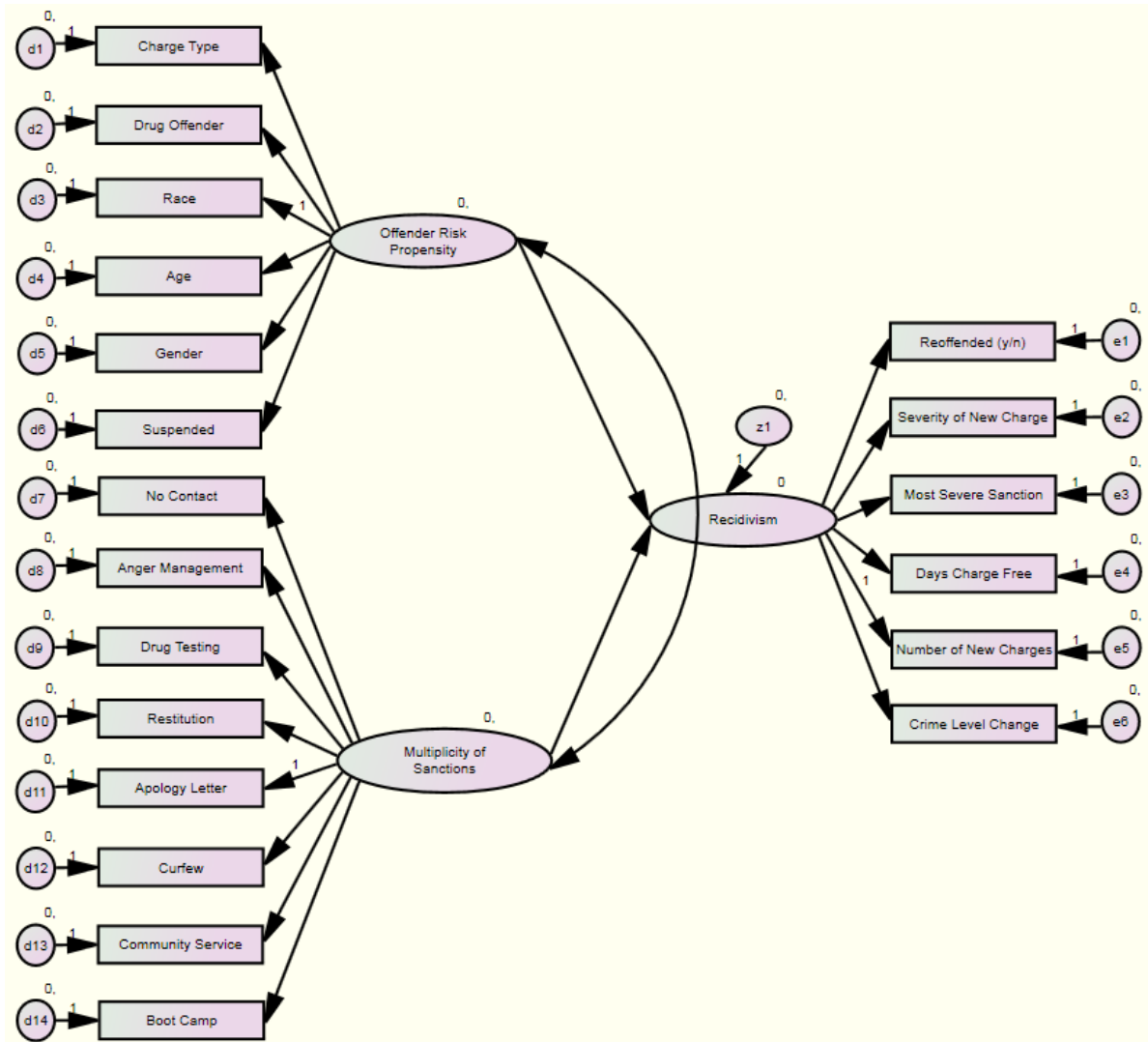


Figure 13: Expanded Main Effect Predictive Model

This figure illustrates the main effect relationships among constructs of Offender Risk Propensity, Multiplicity of Sanctions, and Recidivism shown with their indicators.

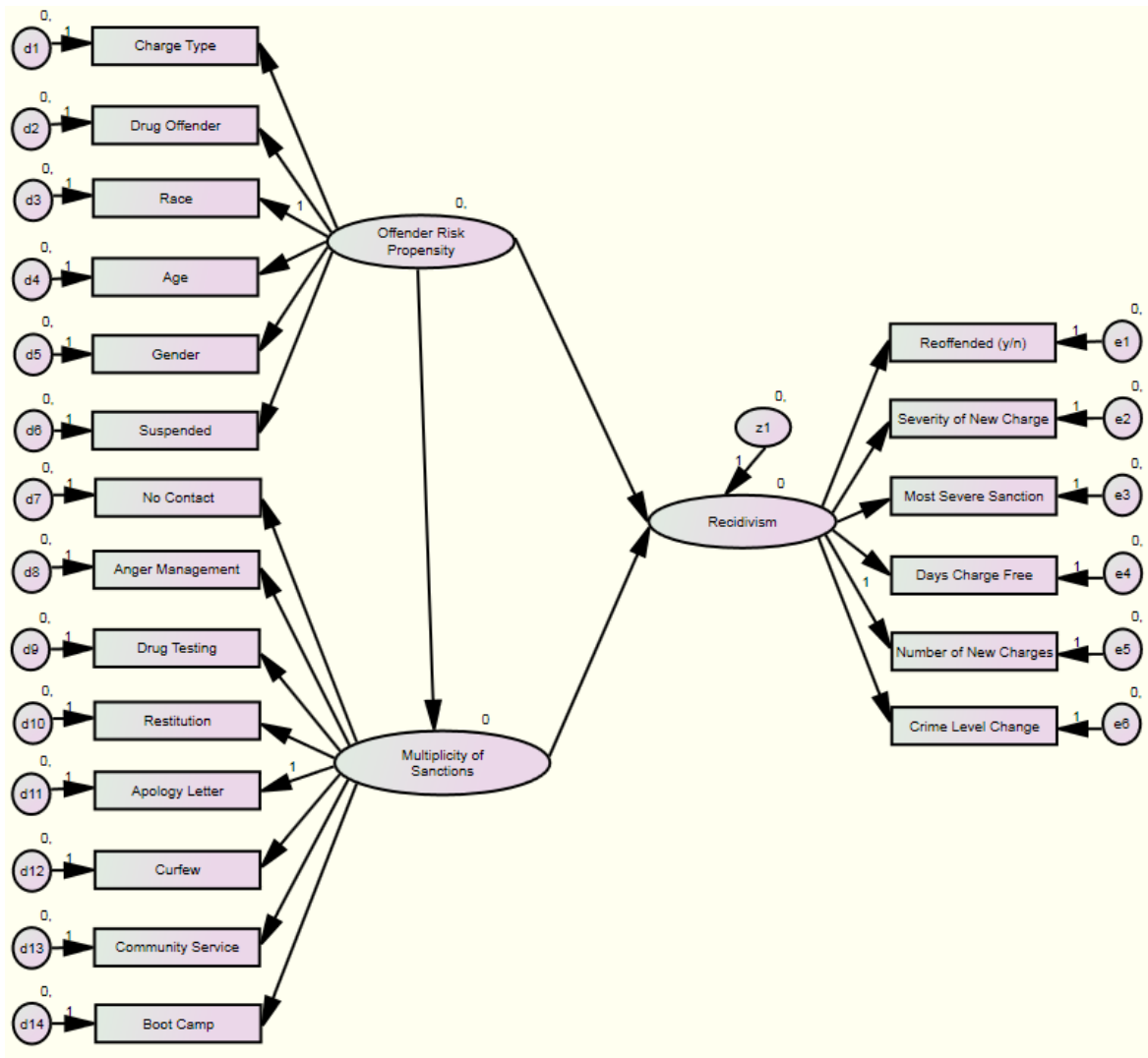


Figure 14: Expanded Mediating Effect Predictive Model

This figure illustrates the mediating effect relationships among constructs of Offender Risk Propensity, Multiplicity of Sanctions, and Recidivism shown with their indicators.

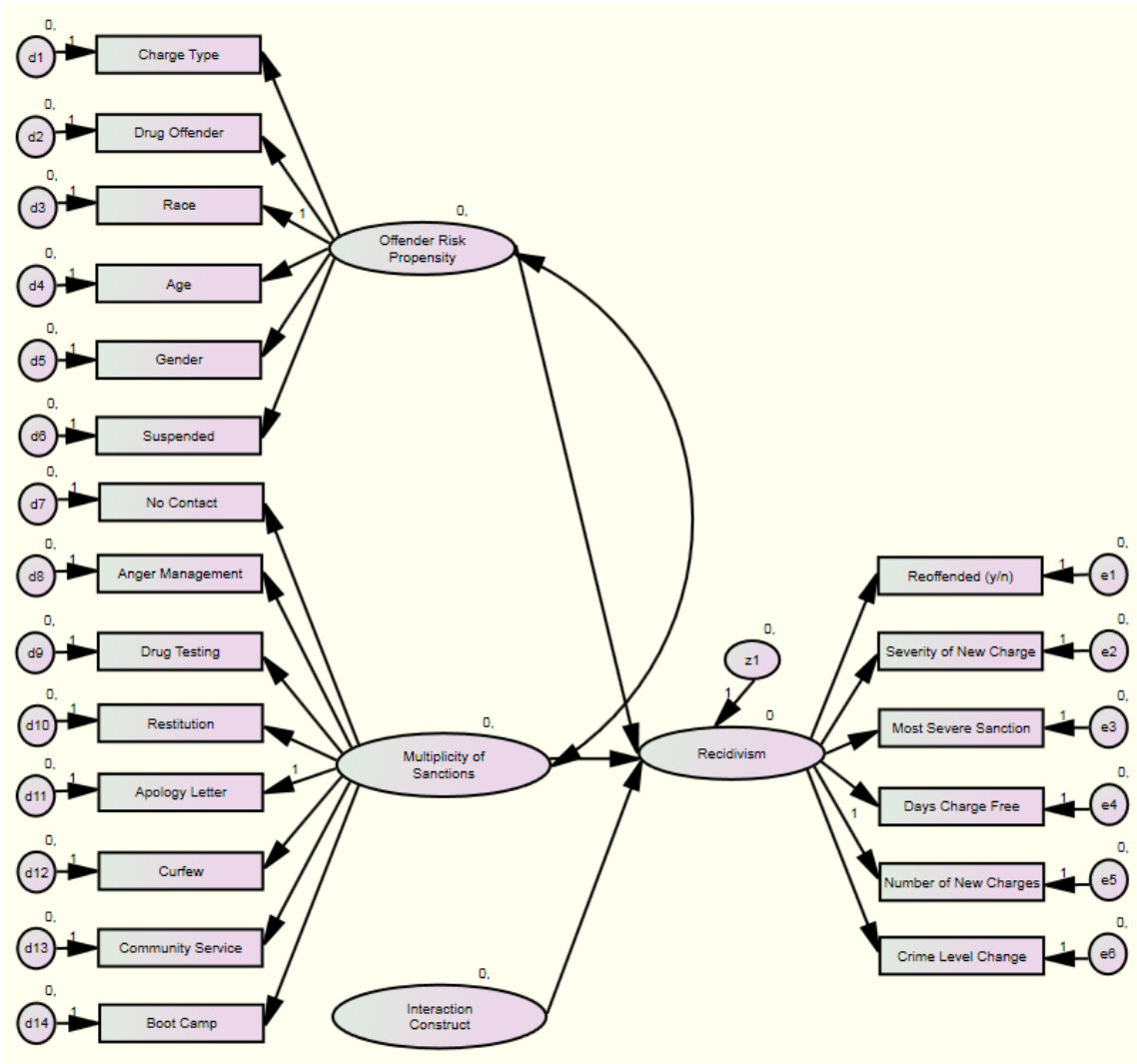


Figure 15: Expanded Moderating Effect Predictive Model

This figure illustrates the moderating effect relationships among constructs of Offender Risk Propensity, Multiplicity of Sanctions, and Recidivism shown with their indicators.

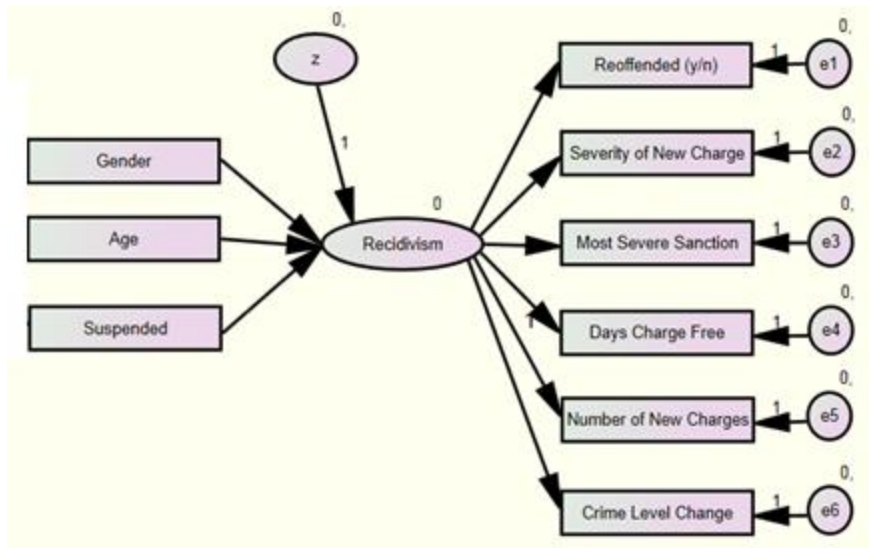


Figure 16: Predictive Model of Recidivism Based on Gender, Age, and Suspension History

This model illustrates the conventional view in criminal justice literature that age of offense, gender and prior history of deviance predict recidivism. This model is tested in Hypothesis 2.

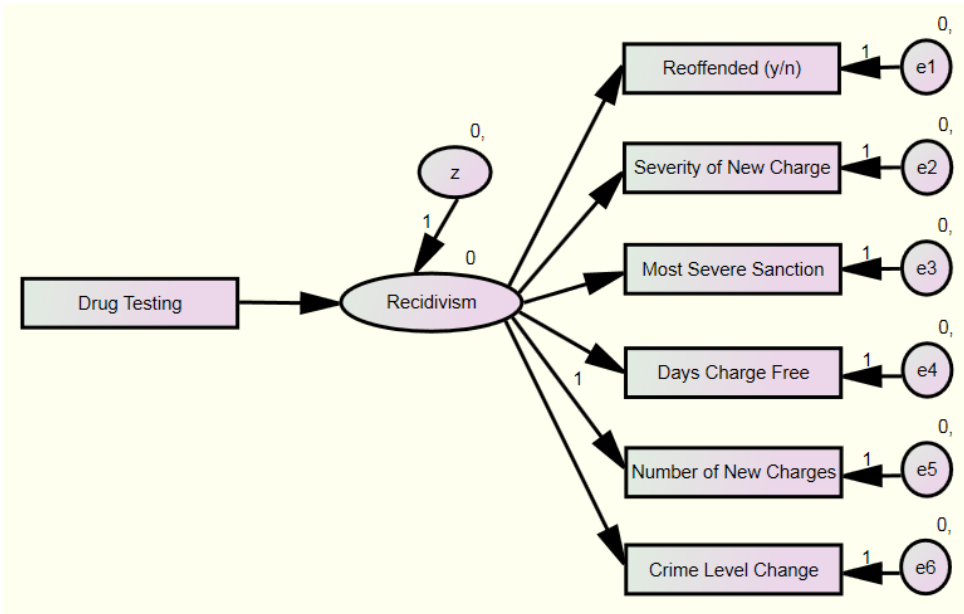


Figure 17: Predictive Model of Recidivism Based on Drug Testing Assignment

This predictive model illustrates the relationship between the assignment of drug testing and Recidivism. This model is tested against data for drug- and non-drug-related offenders.

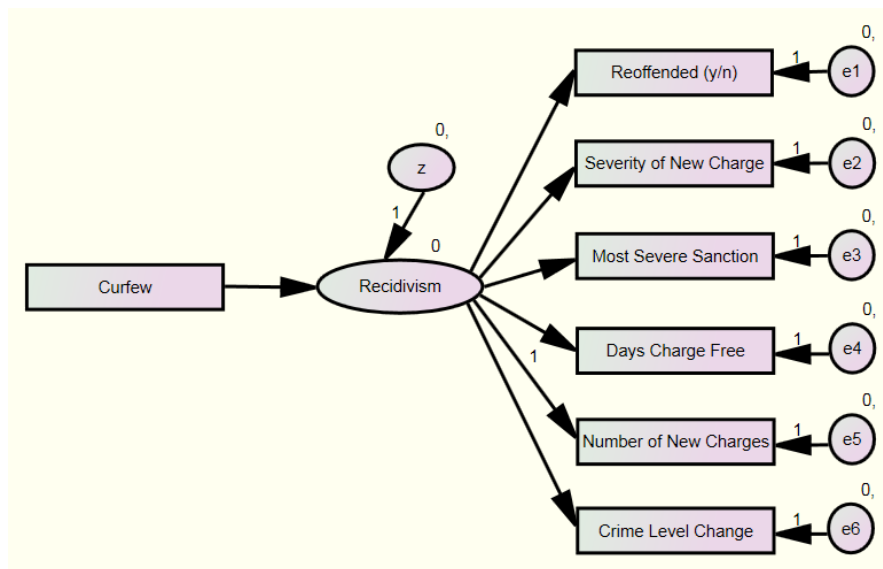


Figure 18: Predictive Model of Recidivism Based on Curfew Assignment

This model illustrates the relationship between the assignment of a curfew and Recidivism. It was the intention to test this model; however, in the data set only one offender

was not assigned a curfew. For that reason, the test was conducted later with the assigned curfew time as the predictor.

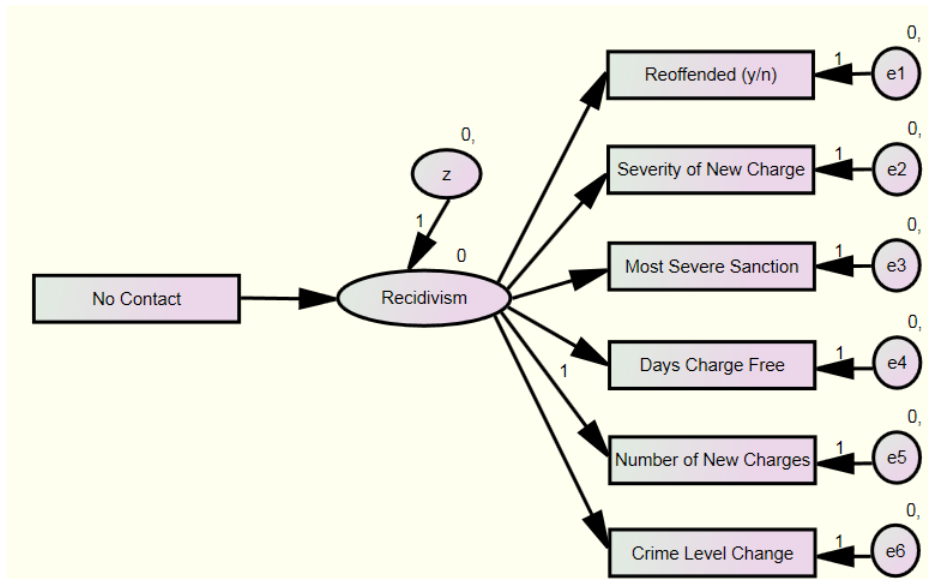


Figure 19: Predictive Model of Recidivism Based on the Assignment of a No Contact Order

This model illustrates the relationship between the assignment of no contact orders and Recidivism. This model is tested against all offenders, males at least 15 years of age, and drug offenders.

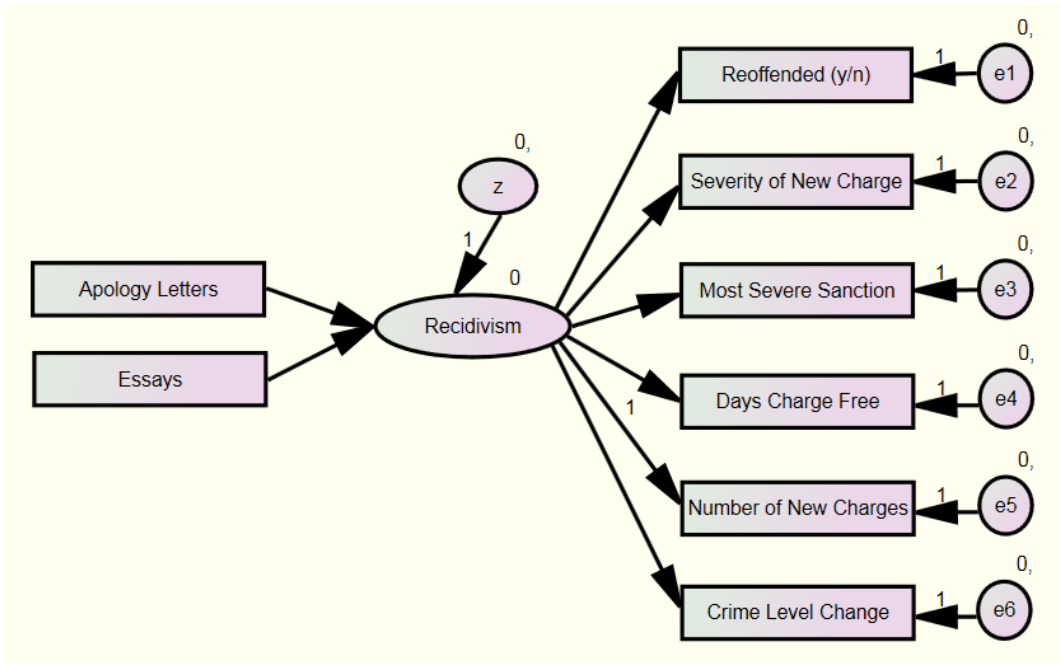


Figure 20: Predictive Model of Recidivism Based on Essay and Apology Letter Assignment

This figure illustrates the relationship between the assignment of apology letters and essays and Recidivism. This model is tested against the data for female offenders.

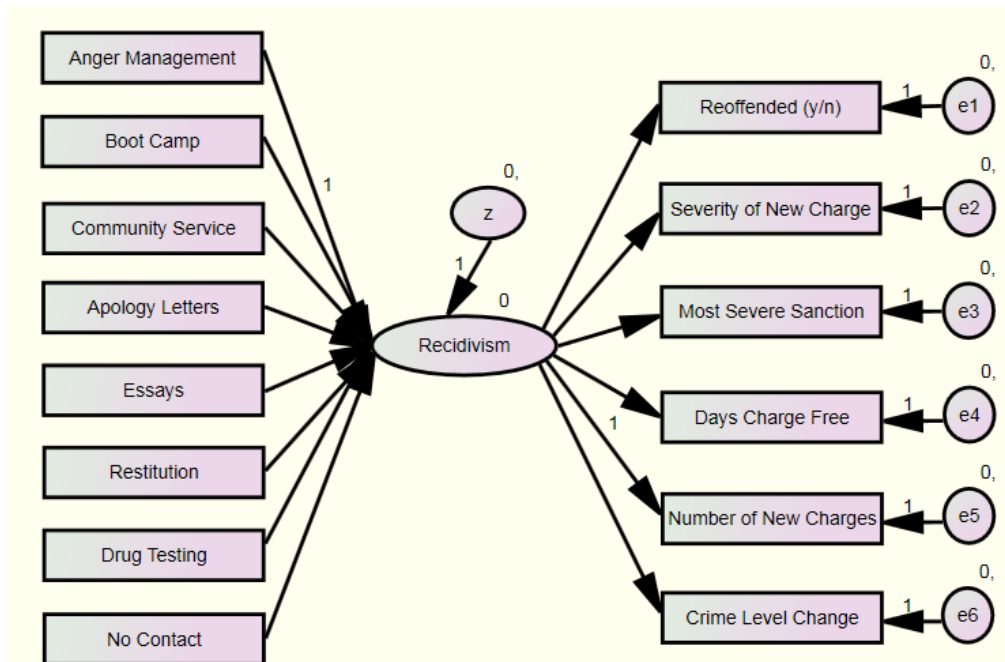


Figure 21: Predictive Model of Recidivism Based on the Assignment of Eight Commonly Assigned Sanctions

This figure illustrates the relationship between the assignment of sanctions and Recidivism. This model is an example only and is not tested through structural equation modeling in this study.

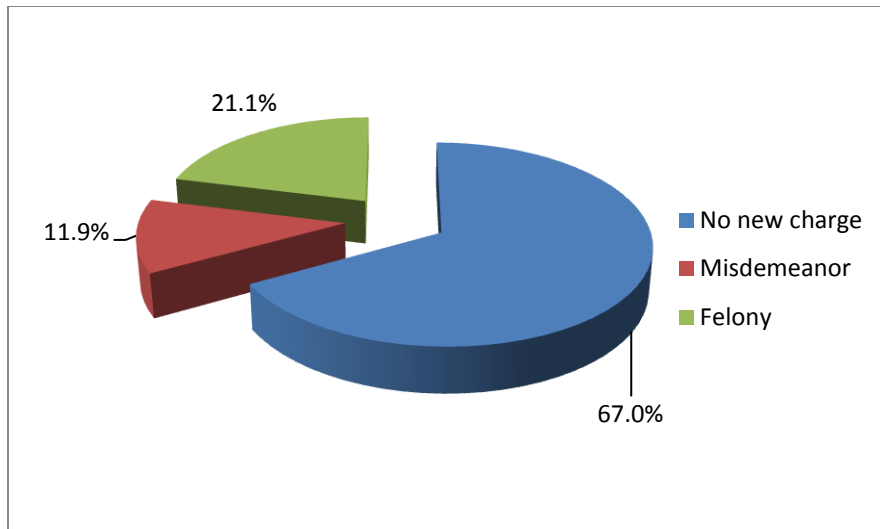


Figure 22: Pie Chart of Severity of New Charge for Entire Sample

This figure illustrates that although a majority of the sample did not receive additional charges, those who did were charged with felonies at almost twice the rate of those charged with only misdemeanors.

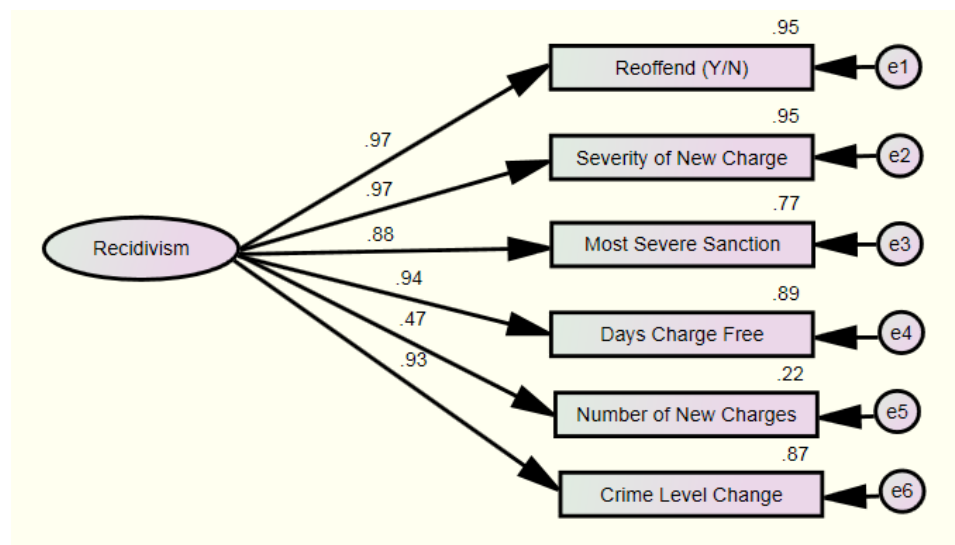


Figure 23. Results of Testing the Measurement Model of Recidivism With All Six Indicators

This model is shown with the factor weights of the indicators of the latent construct of Recidivism when tested against the entire data set (n = 218).

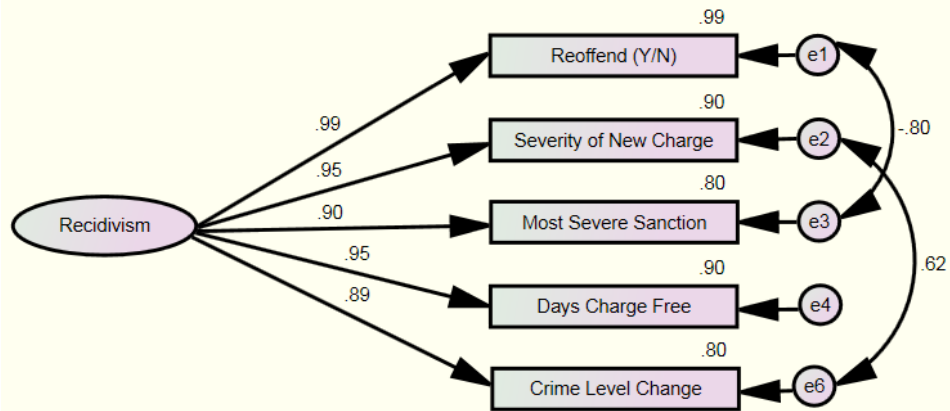


Figure 24: Results of Testing the Revised Measurement Model of Recidivism With Five Indicators

This model is shown with the factor weights of the indicators of the latent construct of Recidivism when tested against the entire data set (n = 218).

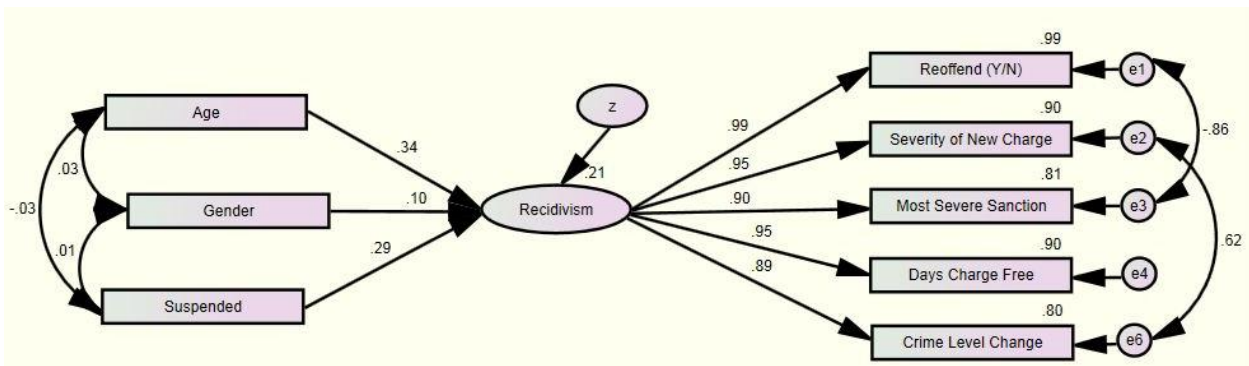


Figure 25: Predictive Model of Recidivism Based on Age Group, Gender, and Suspension History

This model is shown with the factor weights when tested against the entire data set (n = 218). Age groups are defined as those under 15 years of age and those at least 15 years of age at the time of their initial charge (see Table 19 for goodness of fit statistics).

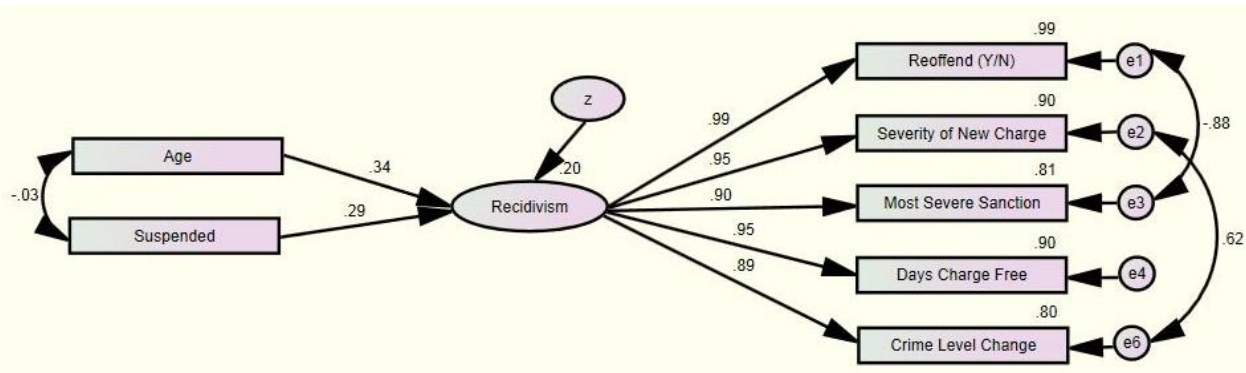


Figure 26: Revised Predictive Model of Recidivism Based on Age Group and Suspension History

This model is shown with the factor weights when tested against the entire data set ($n = 218$). Age groups are defined as those less than 15 years of age and those at least 15 years of age at the time of their initial charge (see Table 19 for goodness of fit statistics).

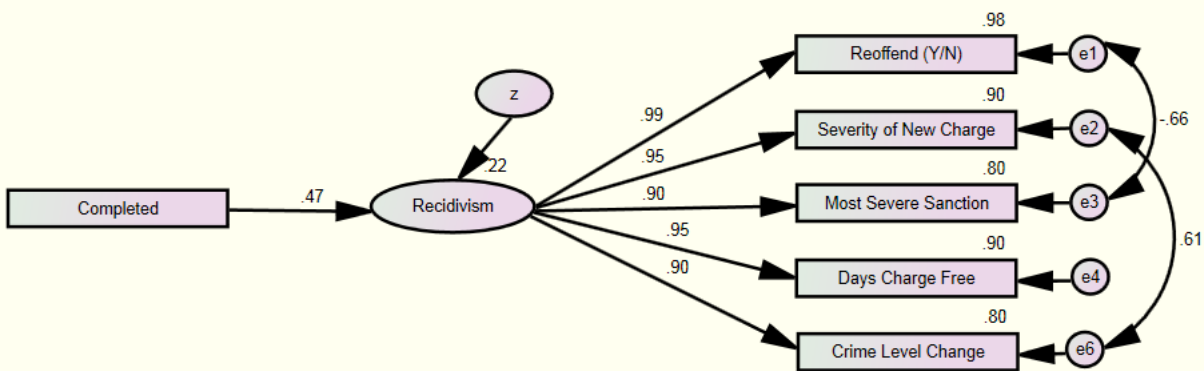


Figure 27: Predictive Model of Recidivism Based on Completion of the Restorative Justice Program

This model is shown with the factor weights when tested against the entire data set ($n = 218$). This predictive model has a chi-square of 16.0 (with 7 degrees of freedom); a relative chi-square of 2.29; and GFI, AGFI and CFI values of 0.997, 0.932, and 0.995, respectively. These statistics show a good fit between the model and the data. The p -value is below 0.05

(0.025), and the RMSEA is above 0.05 (0.077). These statistics indicate an adequate fit between the data and the model (see Table 28 in Appendix A).

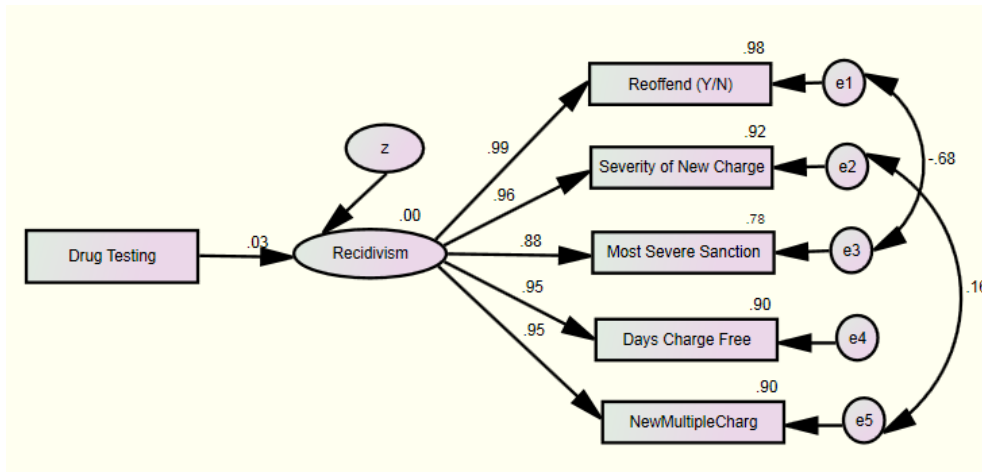


Figure 28: Predictive Model of Recidivism Based on the Assignment of Drug Testing

This model has a good fit with the data (chi-square of 7.7 with 7 degrees of freedom; p -value of 0.358; GFI, AGFI and CFI of 0.956, 0.867, and 0.998, respectively; and an RMSEA of 0.044). However, the p -value for the regression weight between drug testing and Recidivism is 0.133, which indicates that the regression weight of drug testing in the prediction of Recidivism is not significantly different from zero at the 0.05 level (two-tailed).

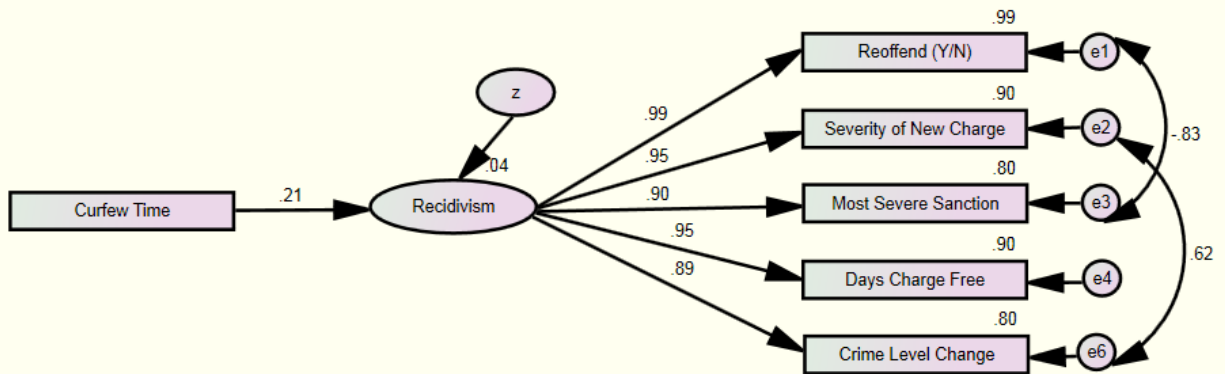


Figure 29: Predictive Model of Recidivism Based on the Assigned Curfew Time

The figure is shown with results from the data set of 216 offenders. All the goodness of fit measures indicate a good fit between the data and the model (chi-square of 2.4 with 7 degrees of freedom; *p*-value of 0.931; GFI, AGFI, and CFI of 0.996, 0.989, and 1.000, respectively; and RMSEA of 0.000).

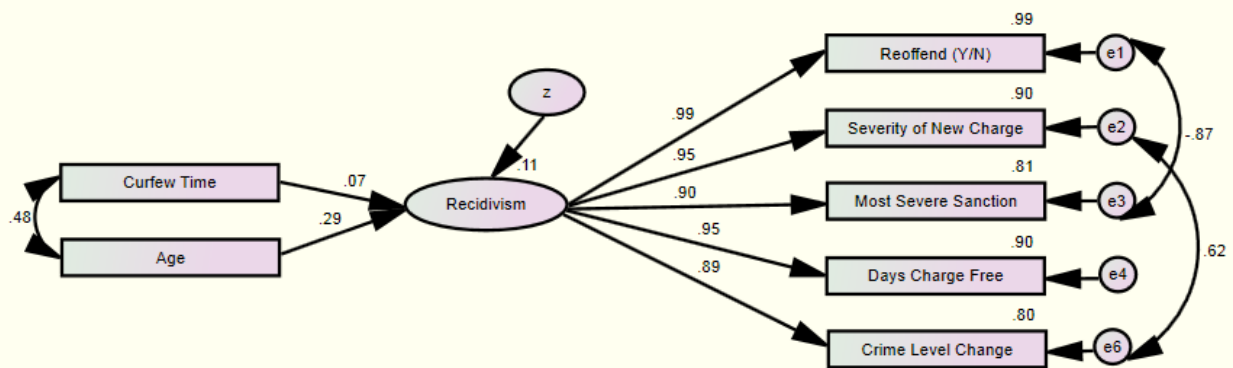


Figure 30: Predictive Model of Recidivism Based on the Assigned Curfew Time Controlled for Age Group

This figure is shown with results from the data set of 216 offenders. All the goodness of fit measures indicate a good fit between the data and the model (chi-square of 19.0 with 11 degrees of freedom; *p*-value of 0.061; GFI, AGFI, and CFI of 0.976, 0.939, and .0996, respectively; and RMSEA of 0.058). However, the regression weight for the predictor curfew

time in the prediction of the latent construct Recidivism is not statistically different from zero at the 0.05 level (two-tailed).

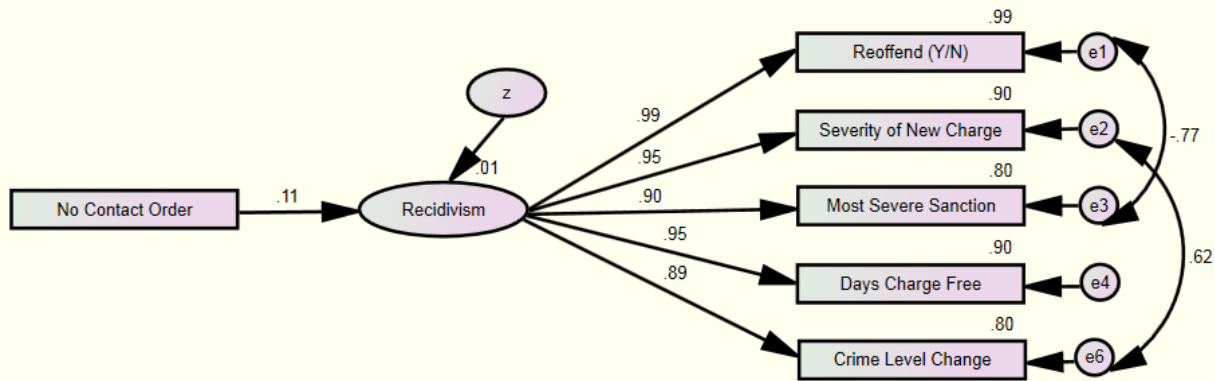


Figure 31: Predictive Model of Recidivism Based on the Assignment of No-Contact Order

When tested against the entire data set ($n = 218$), the model is a very good fit with the data. All the goodness of fit measures indicate a good fit between the data and the model (chi-square of 3.7 with 7 degrees of freedom; p -value of 0.819; GFI, AGFI, and CFI of 0.994, 0.983, and 1.000, respectively; and RMSEA of 0.000). However, the regression weight for the predictor no-contact order in the prediction of the latent construct Recidivism is not statistically different from zero at the 0.05 level (two-tailed).

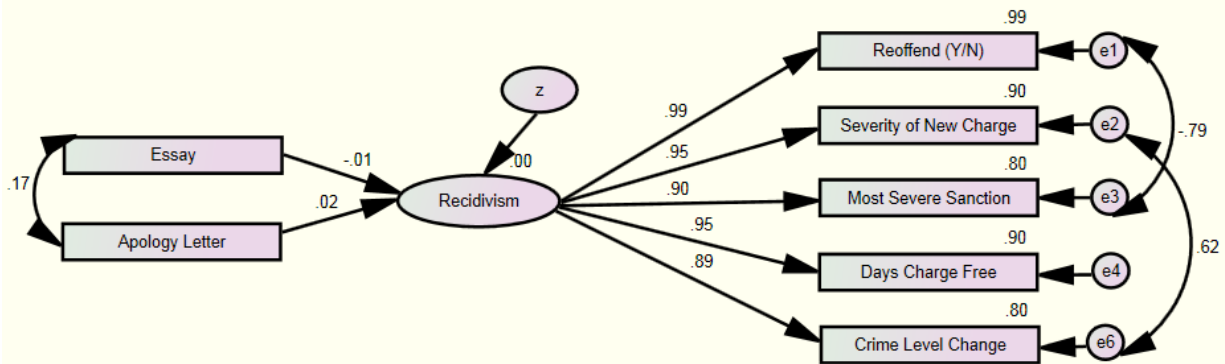


Figure 32: Predictive Model of Recidivism Based on the Assignment of Apology Letters and Essays

Figure is shown with results from the data set of 218 offenders. All the goodness of fit measures indicate a good fit between the data and the model (chi-square of 16.3 with 11 degrees of freedom; p -value of 0.132; GFI, AGFI, and CFI of 0.979, 0.948, and 0.997, respectively; and RMSEA of 0.047) (see Table 56 in Appendix A). However, the regression weight for the predictors, apology letters and essays, in the prediction of the latent construct of Recidivism are not statistically different from zero at the 0.05 level (two-tailed).

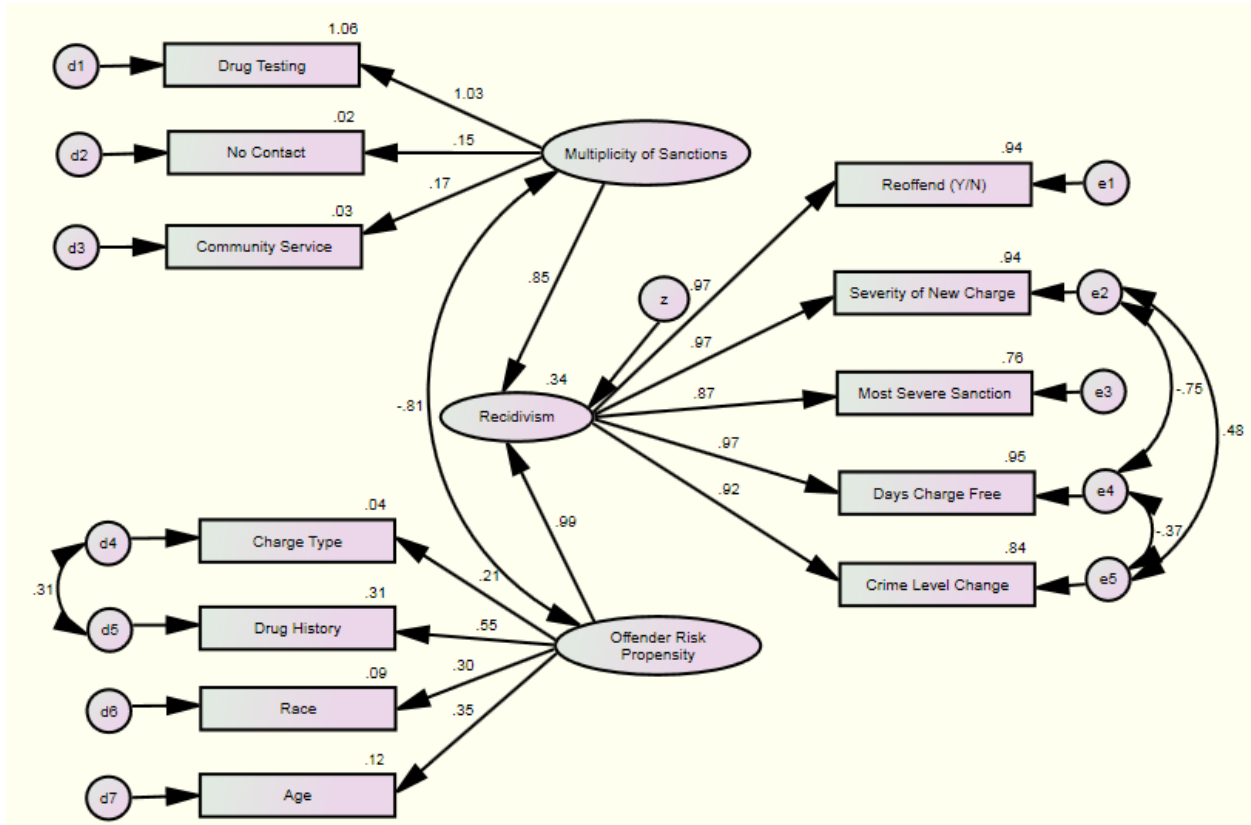


Figure 33: Results of Testing the Revised Main Effect Predictive Model of Recidivism Based on Offender Risk Propensity and Multiplicity of Sanctions

Figure is shown with results from the data set of 213 offenders. The model is a poor fit with the data (chi-square of 108.0 with 47 degrees of freedom; *p*-value of 0.000; GFI, AGFI, and CFI of 0.927, 0.878, and 0.967, respectively; and RMSEA of 0.078) (see Table 64 in Appendix A). The regression weights of Offender Risk Propensity and Multiplicity of Sanctions in the prediction of Recidivism are not significantly different from zero at the 0.05 level (two-tailed).

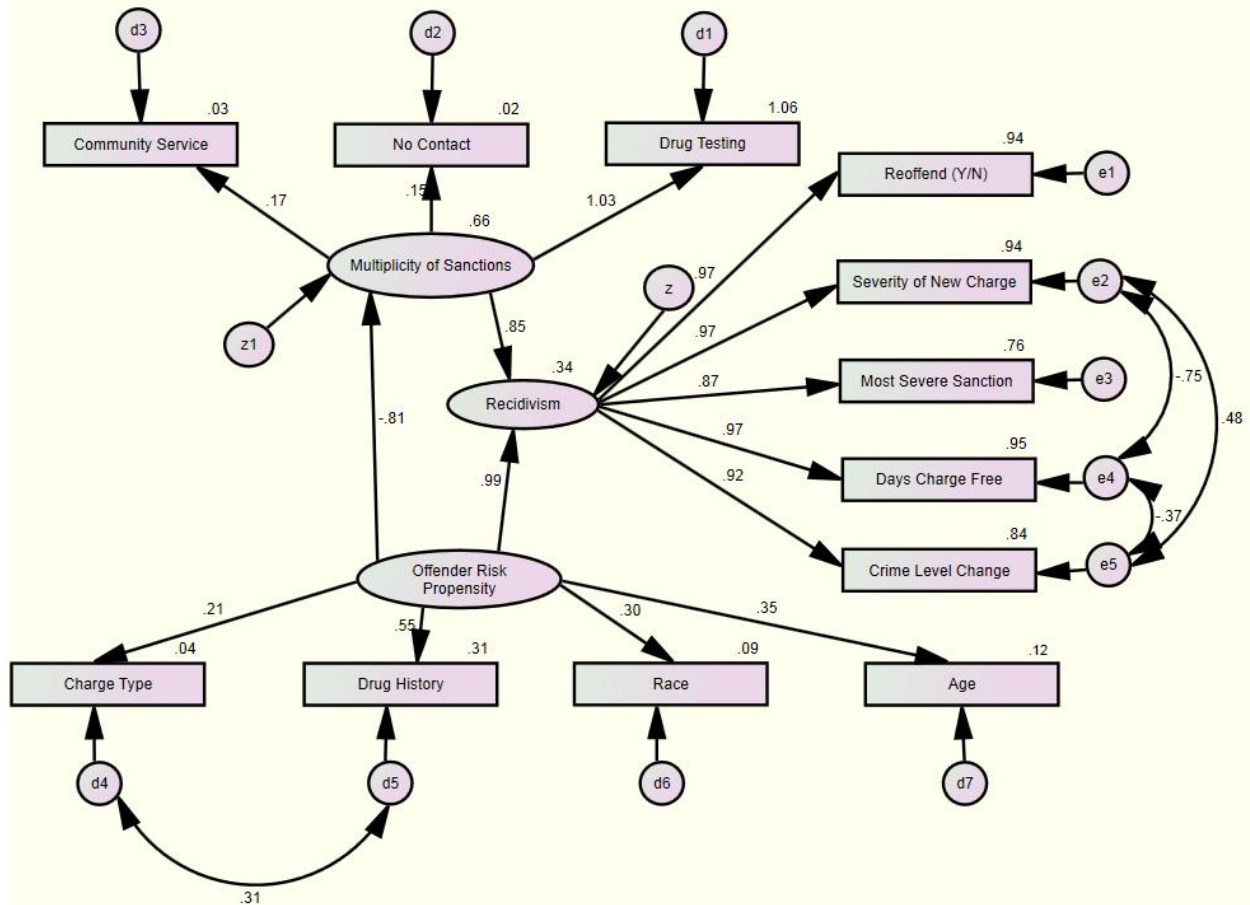


Figure 34: Results of Testing the Revised Mediating Effect Predictive Model of Recidivism Based on Offender Risk Propensity and Multiplicity of Sanctions

Figure is shown with results from the data set of 213 offenders offered participation in restorative justice program. The model is a poor fit with the data (chi-square of 108.0 with 47 degrees of freedom; p -value of 0.000; GFI, AGFI, and CFI of 0.927, 0.878, and 0.967, respectively; and RMSEA of 0.078) (see Table 64 in Appendix A). The regression weights of Offender Risk Propensity and Multiplicity of Sanctions in the prediction of Recidivism are not significantly different from zero at the 0.05 level (two-tailed).

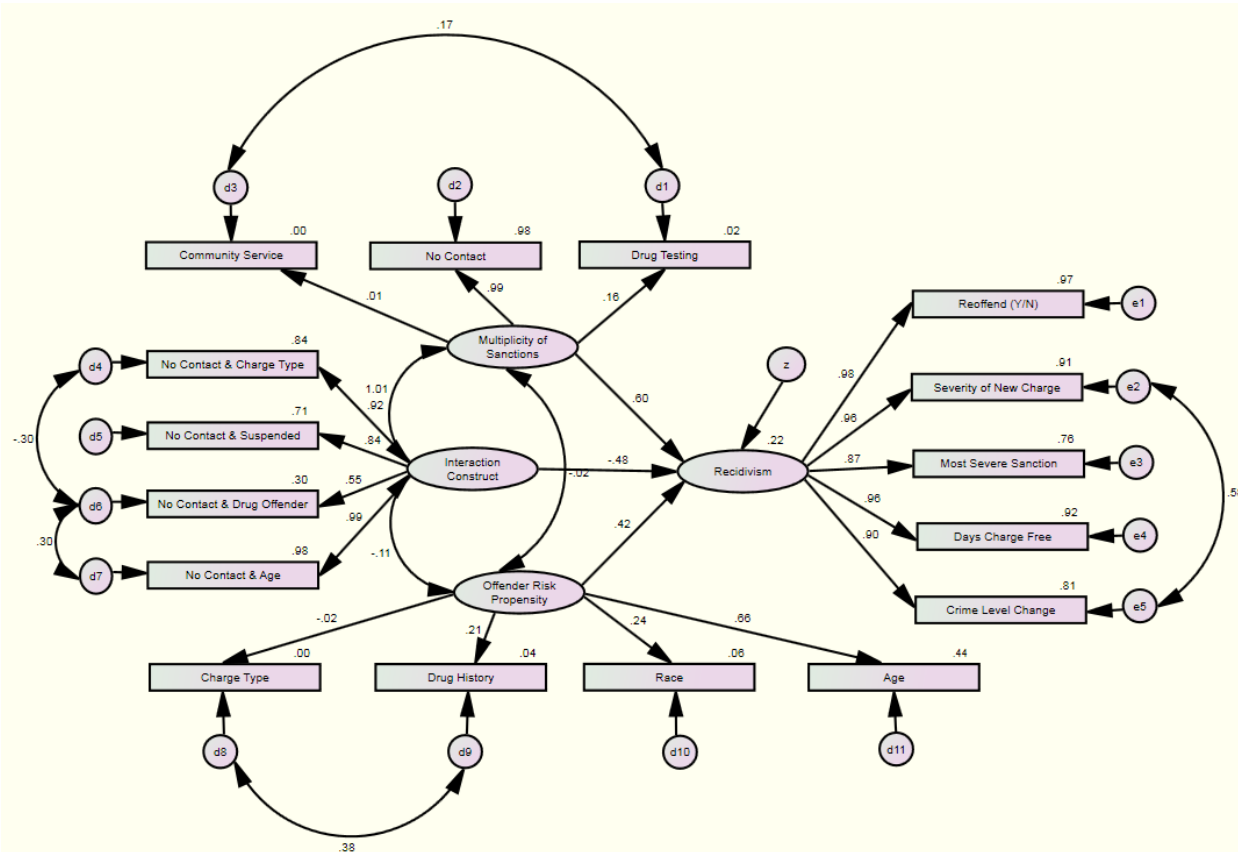


Figure 35: Results of Testing the Revised Moderating Effect Predictive Model of Recidivism Based on Offender Risk Propensity, Multiplicity of Sanctions, and the Interaction Construct.

Figure is shown with results from the data set of 213 offenders and it is a poor fit with the data (chi-square of 342.6 with 93 degrees of freedom; p -value of 0.000; GFI, AGFI, and CFI of 0.844, 0.771, and 0.931, respectively; and RMSEA of 0.113) (see Table 65 in Appendix A). The regression weights of Offender Risk Propensity and Multiplicity of Sanctions in the prediction of Recidivism are not significantly different from zero at the 0.05 level (two-tailed).

APPENDIX C: INTERNAL REVIEW BOARD FINDINGS



University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901, 407-882-2012 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

From : **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To : **Thomas T. Maroney**

Date : **November 04, 2011**

Dear Researcher:

On 11/04/2011 the IRB determined that the following proposed activity is not human research as defined by DHHS regulations at 45 CFR 46 or FDA regulations at 21 CFR 50/56:

Type of Review:	Not Human Research Determination
Project Title:	Sanction Effectiveness in Youth Diversion Programs
Investigator:	Thomas T. Maroney
IRB ID:	SBE-11-07993
Funding Agency:	
Grant Title:	
Research ID:	N/A

University of Central Florida IRB review and approval is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are to be made and there are questions about whether these activities are research involving human subjects, please contact the IRB office to discuss the proposed changes.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 11/04/2011 11:47:47 AM EST

A handwritten signature in cursive script that reads "Joanne Muratori".

IRB Coordinator

REFERENCES

- Ahlin, E. M., Zandor, P. L., Rauch, W. J., Howard, J. M., & Duncan, G. D. (2011). First time DWI offenders at risk of recidivating regardless of sanctions imposed. *Journal of Criminal Justice, 39*(3), 137–142.
- Allison, P. D. (1982). Discrete-time methods for the analysis of event history. *Sociology Methodology, 13*, 61–98.
- Applegate, B., Turner, M., Sanborn, J., Latessa, E., & Moon, M. (2000). Individualization, criminalization, or problem resolution: A factorial survey of juvenile court judges' decisions to incarcerate youthful felony offenders. *Justice Quarterly, 17*(2), 309–331.
- Ashford, J. B., & LeCroy, C. W. (1990). Juvenile recidivism: A comparison of three predictive instruments. *Adolescence, 25*(98), 441–450.
- Baird, S. C. (1982). *Classification of juveniles in corrections: A model system approach*. Madison, WI: National Council on Crime and Delinquency.
- Bank, L., Marlowe, J. H., Reid, J. B., Patterson, G. R., & Weinrott, M. R. (1991). A comparative evaluation of parent-training interventions for families of chronic offenders. *Journal of Abnormal Child Psychology, 19*(1), 15–33.
- Bardach, E. (2000). *A practical guide for policy analysis: The eight fold path to more effective problem solving*. New York, NY: Seven Bridges Press.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.
- Bazemore, G., & Walgrave, L. (1999). *Restorative juvenile justice: Repairing the harm of youth crime*. Monsey, NY: Criminal Justice Press.
- Bishop, D. M., & Frazier, C. E. (1992). Gender bias in juvenile justice processing: Implications of the JJDP act. *The Journal of Criminal Law and Criminology, 82*(4), 1162–1186.
- Bishop, D. M., & Frazier, C. E. (1996). Race effects in juvenile justice decision-making. *Journal of Criminal Law and Criminology, 86*(2), 392–414.
- Black, D. J., & Reiss, A. J., Jr. (1970, February). Police control of juveniles. *American Sociological Review, 35*, 63–77.
- Blair, J. (2000, January 2). Boot camps: An idea whose time came and went. *The New York Times*, pp. 3, 4.

- Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. J. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, 34(2), 105–113.
- Borduin, C. M., Mann, B. J., Cone, L. T., Henggeler, S. W., Fucci, B. R., Blaske, D. M., & Williams, A. W. (1995). Multisystemic treatment of serious juvenile offender: Long term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology*, 63(4), 569–578.
- Bottcher, J., & Ezell, M. E. (2005). Examining the effects of boot camps: A randomized experiment with a long term follow up. *Journal of Research in Crime and Delinquency*, 42(3), 309–552.
- Bouffard, J. A., & Bergseth, K. J. (2008). The impact of reentry services on juvenile offenders' recidivism. *Youth Violence and Juvenile Justice*, 6(3), 295–318.
- Bouffard, J. A., & Muftic, M. (2007). The effectiveness of community service compared to traditional fines for low-level offenders. *The Prison Journal*, 87(2), 171–194.
- Boulden, W. T. (Spring 2010). The Behavior Intervention Support Team (BIST) Program: Underlying theories. *Reclaiming Children and Youth*, 19(1), 17–21.
- Braithwaite, J. (1989). *Crime, shame and reintegration*, Cambridge, MA: Cambridge University Press.
- Britt, C. L., Gottfredson, M. R., & Goldkamp, J. S. (1992). Drug testing and pretrial trial misconduct: An experiment on the specific deterrent effects of drug monitoring defendants on pretrial release. *Journal of Research in Crime and Delinquency*, 29(1), 62–78.
- Carter, C. (1989). *Analysis for public decisions*. Salem, MA: North Holland Press.
- Crowell, A. (1996). Minor restrictions: The challenge of juvenile curfews. *Public Management*, 78(8), 4–9.
- Cullen, F. T., & Gendreau, P. (2000). Assessing correctional rehabilitation: Policy, practice, and prospects. In J. Horney (Ed.), *NIJ criminal justice 2000: Changes in decision making and discretion in the criminal justice system* (pp. 109–175). Washington, DC: U.S. Department of Justice, National Institute of Justice.
- De Beus, K., & Rodriguez, N. (2007). Restorative justice practice: An examination of program completion and recidivism. *Journal of Criminal Justice*, 35(3), 337–347.
- Dick, A., Pence, D., Jones, R., & Geetresen, H. (2004). The need for theory in assessing peer courts. *The American Behavioral Scientist*, 47(11), 1448–1452.

- Doolin, K. (2007). But what does it mean? Seeking definitional clarity in restorative Justice. *Journal of Criminal Law*, 71(5), 427–440.
- Dowden, C., & Serin, R. (1999). Anger management programming for federal male inmates: An effective intervention (Research Report R82e). Retrieved from http://www.csc-scc.gc.ca/text/rsrch/reports/r106/r106_e.pdf
- East Point, Georgia, City of. (2011). City facts and demographics. Retrieved from <http://www.eastpointcity.org/index.aspx?NID=17>
- Elis, L. (2005). Restorative justice programs, gender, and recidivism. *Public Organization Review*, 5(4), 375–389.
- Fagan, J. A. (1990). Treatment and reintegration of violent juvenile offenders: Experimental results. *Justice Quarterly*, 7(2), 233–263.
- Florida Department of Corrections. (2003). County detention facilities daily averages for October 2003. Retrieved from: <http://www.dc.state.fl.us/pub/jails/2003/10/table1.html>
- Florida Department of Juvenile Justice. (2003). Neighborhood accountability boards: A guide to learn more about neighborhood accountability boards. Retrieved from <http://www.djj.state.fl.us/dJJServices/prevention/prevention/NAB%20GUIDE%20MANUAL03-2003.DOC>
- Florida Department of Juvenile Justice. (2009). 2009 Legislative and General Budget Report. Retrieved from: http://www.djj.state.fl.us/Legislative_affairs/documents/2009-Legislative-and-General-Budget-Report.pdf
- Florida Department of Juvenile Justice. (2010). Intensive delinquency diversion services. Retrieved from: <http://www.djj.state.fl.us/Research/statsnresearch/factsheets/idds.html>
- Florida Senate. (2011). Crimes: Definitions; General Penalties; Registration of criminals. Retrieved from <http://www.flsenate.gov/Laws/Statutes>
- Fredrick, B., & Roy, D. (2003). *Recidivism from youth released from the Youth Leadership Academy to the City Challenge Intensive Aftercare Program*. Retrieved from http://criminaljustice.state.ny.us/crimnet/ojsa/yla/yla_report.pdf
- Georgia ‘juvenile diversions’ cuts crime, reports zero recidivism. (2001). *Juvenile Justice Digest*, 29(4), 1–2.
- Gilligan, C. (1982). *In a different voice*. Cambridge, MA: Harvard University Press.
- Giordano, P., Cernkovich, S., & Pugh, M. D. (1986). Friendships and delinquency. *The American Journal of Sociology*, 21, 1170–1202.

- Glass, G., McGaw, B., & Smith, M. (1981). *Meta analysis in social research*. Newbury Park, CA: Sage.
- Gliner, J. A., & Morgan, G. A. (2000). *Research methods in applied settings: An integrated approach to design and analysis*. Morgan, GA: Lawrence Erlbaum.
- Gottfredson, D. C., & Barton, W. H. (1993, November). Deinstitutionalization of juvenile offenders. *Criminology*, *31*(4), 591–611.
- Guerra, N. G., & Slaby, R. G. (1990). Cognitive mediators of aggression in adolescent offenders: II. Intervention. *Developmental Psychology*, *26*(2), 269–277.
- Haapanen, R., & Britton, L. (2002, March). Drug testing for youthful offenders on parole: An experimental evolution. *Criminology & Public Policy*, *1*(2), 217–244.
- Hagan, J., Simpson, J., & Gillis, A.R. (1988, September). Feminist scholarship, relational and instrumental control, and power control theory of gender and delinquency. *The British Journal of Sociology*, *39*(3), 301–336.
- Hagan, M. P., & Cho, M. E. (1996). A comparison of treatment outcomes between adolescent rapists and child sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, *40*(2), 113–122.
- Harris, P. W., Lockwood, B. & Mengers, L. (2009). *Defining and measuring recidivism* [White Paper]. Retrieved from http://ojj.la.gov/ojj/files/CJCA_Recidivism_WhitePaper.pdf
- Hayes, H. (2005). Assessing re-offending in restorative justice conferences. *Australian and New Zealand Journal of Criminology*, *38*(1), 77–101.
- Hayes, H., & Daly, K. (2003, December). Youth justice conferencing and reoffending. *Justice Quarterly*, *20*(4), 725–764.
- Herzfeld, S. M., Shady, T. A., Mayo, J., Marsh, K., Dorabawila, V., & Rees, F. (2008). Effects of multisystemic therapy (MST) on recidivism among juvenile delinquents in New York State. Retrieved from <http://www.ocfs.state.ny.us/main/reports/FINAL%20MST%20report%206-24-08.pdf>
- Hipple, N. K., & McGarrell, E. F. (2008). Comparing police and civilian run family group conferences. *Policing: An International Journal of Police Strategies & Management*, *31*(4), 553–557.
- Horowitz, R., & Pottieger, A. E. (1991). Gender bias in juvenile justice handling of serious crime involving youths. *The Journal of Research in Crime and Delinquency*, *28*(1), 75–100.

- Josi, D. A., & Sechrest, D. K. (1999). A pragmatic approach to parole aftercare: Evaluation of a community reintegration program for high-risk youthful offenders. *Justice Quarterly*, *16*(1), 51–80.
- Kahn, T. J., & Chambers, H. J. (1991). Assessing reoffense risk with juvenile sexual offenders. *Child Welfare*, *70*(3), 333–345.
- Kaplan, D. (2011). Statistical power in structural equation models. Retrieved from <http://www2.gsu.edu/~mkteer/power.html>
- Kelly, T. M., Kennedy, D. B., & Homart, R. J. (2003, Winter). Evaluation of an individualized treatment program for adolescent shoplifters. *Adolescence*, *38*(152), 725–734.
- Kerlinger, F. (1986). *Foundations of behavioral research* (3rd ed.). New York, NY: Holt, Rinehart and Winston.
- Kim, I. J., Benson, B. L., Rasmussen, D. W., & Zuehlke, T. W. (1993, July). An economic analysis of recidivism among drug offender. *Southern Economic Journal*, *60*(1), 169–84.
- Kingston. (2011). Data Traveler Vault- Privacy Edition. Retrieved from <http://www.kingston.com/flash/DTVVaultPrivacy.asp>
- Kraska, P. B. (2004). *Theorizing criminal justice*. Groveland, IL: Waveland Press.
- Kubak, F. A., & Salekin, R. (2009). Psychopathy and anxiety in children and adolescents: New insights on developmental pathways to offending. *Psychopathy Behavior Assessment*, *31*(4), 271–284.
- Lab, S. P., Williams, M., Holcomb, J. E., King, W., & Buerger, M. (2003). Explaining criminal justice. Los Angeles, CA: Roxbury.
- Latimer, J., Dowden, C., & Muise, D. (2005, June). The effects of restorative justice practices: A meta-analysis. *The Prison Journal*, *85*(2), 127–144.
- Lawrence, P. R., & Lorsch, J. W. (1967): *Organization and environment: Managing differentiation and integration*. Boston, MA: Harvard University.
- Leeman, L. W., Gibbs, J. C., & Fuller, D. (1993). Evaluation of a multi-component group treatment program for juvenile delinquents. *Aggressive Behavior*, *19*, 281–292.
- Levine, S. Z. (2009). Examining the incidence of and time to recidivism within the risk contingency framework: A 20-year follow up study. *Law and Human Behavior*, *33*(2), 167–174.
- Levrant, S., Cullen, F., Fulton, B. & Wozniak, J. (1999). Reconsidering restorative justice: The corruption of benevolence revisited. *Crime and Delinquency*, *45*(1), 3-27.

- Lidz, C. W., Mulvey, E. P., & Gardner, W. (1993). The accuracy of predictors of violence to others. *Journal of the American Medical Association*, 269(8), 1007–1011.
- Lundman, R., Sykes, R., & Clark, J. (1978). Police control of juveniles: A replication. *Journal of Research in Crime and Delinquency*, 15(1), 74–91.
- Males, M. A., & Macallair, D. (1999). An analysis of curfew enforcement and juvenile crime in California. *Western Criminology Review*. Retrieved from <http://wcr.sonoma.edu/v1n2/males.html>
- Marshall, T. F. (1999). *Restorative justice: An overview*. London, England: Home Office.
- Martin, G. (2003, February). The effectiveness of community-based sanctions in reducing recidivism. *Corrections Today*, 61(1), 26–29.
- McCold, P., & Wachtel, D. (1998). *Restorative policing experiment: The Bethlehem Pennsylvania Police Family Group Conferencing Project*. Retrieved from <http://www.iirp.edu/pdf/BPD.pdf>
- McDonald, D. C. (1989). *Punishment without walls: Community service sentences in New York City*. Rutgers, NJ: Rutgers University Press.
- McDowall, D., Loftin, C., & Wiersema, B. (2000, January). The impact of youth curfew laws on juvenile crime rates. *Crime & Delinquency*, 46(1), 76–91.
- McIntosh, C. N. (2007, May). Rethinking fit assessment in structural equation modeling: A contemporary and elaboration on Barrett. *Personality and Individual Differences*, 2(5), 859–867.
- Mears, D., Ploeger, M., & Warr, M. (1989). Explaining the gender gap in delinquency. *The Journal of Research in Crime and Delinquency*, 35(3), 251–276.
- Microsoft. (2011). Create strong passwords. Retrieved from <http://www.microsoft.com/security/online-privacy/passwords-create.aspx>
- Morash, M. (1986). Gender, peer group experiences, and social delinquency. *The Journal of Research on Crime and Delinquency*, 23, 43–67.
- Muikuvuori, M. (2001). Recidivism among people sentenced to community service in Finland. *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 4(1), 72–82.
- Mulder, E., Brand, E., Bullens, R., & Van, M. (2010). A classification of risk factors in serious juvenile offenders and the relation between patterns of risk factors and recidivism. *Criminal Behavior and Mental Health*, 20(1), 23–38.

- Narveson, J. (2002). Libertarianism. In H. LaFollette (Eds.), *The Blackwell guide to ethical theory* (165–182). Somerset, United Kingdom: Blackwell.
- Nugent, W. R., & Paddock, J. (1996). Evaluating the effects of victim-offender reconciliation program on reoffense. *Research on Social Work Practice*, 6(2), 155–178.
- Nugent, W. R., Williams, M., & Umbreit, M. S. (2004). Participation in victim-offender mediation and the prevalence of subsequent delinquent behavior: A meta analysis. *Research on Social Work Practice*, 16(6), 408–416.
- Orange County Public Schools SAFE Program (2004). *News flash*. Retrieved from http://www.safe.ocps.net/SAFEweb2/2_District_Dix.html
- Pallant, J. (2005). *SPSS Survival Manual 2nd ed.* New York, NY: McGraw-Hill Education.
- Peters, B. (2004). *American public policy*. Washington, DC: Congressional Quarterly.
- Peters, M., Thomas, D., & Zamberlan, C. (1997). *Boot camps for juvenile offenders: Program summary*. Retrieved from <https://www.ncjrs.gov/pdffiles/164258.pdf>
- Prison Fellowship International. (2000). Restorative justice briefing paper. Retrieved from <http://www.pfi.org/cjr/downloads/focus-on-justice/focus-on-justice/what-is-restorative-justice/view>
- Putnam, R. D. (2000). *Bowling alone*. New York, NY: Simon & Schuster.
- Puzzanchera, C. (2009). *Juvenile arrests 2008. Juvenile Justice Bulletin*. Retrieved from www.ncjrs.gov/pdffiles1/ojjdp/228479.pdf
- Rankin, J. (1980). School factors and delinquency: Interactions by age and sex. *Sociology and Social Research*, 64(3), 420–434.
- Reynolds, K. M., Myers, S., & Dziegielewski, S. (2002) Orange County Sheriff's Office J.A.M. Project Evaluation 2001–2002. Orange County, FL: University of Central Florida.
- Roberg, R., Novak, K., & Cordner, G. (2005). *Police and society* (3rd ed.). Los Angeles, CA: Roxbury.
- Robinson, P., & Darley, J. (2004). Does criminal law deter? A behavior science investigation. *Oxford Journal of Legal Studies*, 24(2), 173–216.
- Rodriguez, N. (2005). Restorative justice, communities, and delinquency: Who do we reintegrate? *Criminology & Public Policy*, 4(1), 601–629.
- Rodriguez, N. (2007, July). Restorative justice at work: Examining the impact of restorative justice resolutions on juvenile recidivism. *Crime and Delinquency*, 53(3), 355–379.

- Rosenthal, R. (1991). *Meta-analysis procedures for social research*. Newbury, CA: Sage.
- Roskey, J. W., Pasini-Hill, D., Lowden, K., Harrison, L., & English, K. (2004). Evaluation of the Youthful Offender System (YOS) in Colorado. Retrieved from: http://dcj.state.co.us/ors/pdf/docs/yos_report_2004.pdf
- Ruefle, W., & Reynolds, K. M. (1995). Curfews and delinquency in major American cities. *Crime and Delinquency*, 41(3), 347–363.
- Ruefle, W., & Reynolds, K. M. (1996). Keep them at home: Juvenile curfew ordinances in 200 American cities. *American Journal of Police*, 15(1), 63–84.
- Salamon, L. M. (2002). *The tools of government: A guide to the new governance*. New York, NY: Oxford Press.
- Schaeffer, C. M., & Borduin, C. M. (2005). Long-term follow-up to a randomized clinical trial of Multisystemic Therapy with serious and violent juvenile offenders. *Journal of Consulting and Clinical Psychology*, 73(3), 445–453.
- Schneider, A. L. (1990). *Deterrence and juvenile crime: Results from a national policy experiment*. New York, NY: Springer-Verlag.
- Scott, W. R. (2003). *Organizations: Rational, natural, and open systems* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- Sherman, L. W., Strang, H., & Woods, D. J. (2000). Recidivism patterns in the Canberra Reintegrative Shaming Experiments (RISE). Retrieved from http://www.aic.gov.au/criminal_justice_system/rjustice/rise/recidivism.aspx
- Simpson, S., & Elis, L. (1995). Doing gender: Sorting out the caste and crime continuum. *Criminology*, 33(1), 47–81.
- Smith, W. R., & Monastersky, C. (1986). Assessing juvenile sex offenders' risk of reoffending. *Criminal Justice and Behavior*, 13(2), 115–140.
- Sontheimer, H., & Goodstein, L. (1993). Evaluation of juvenile intensive aftercare probation: Aftercare versus system response effects. *Justice Quarterly*, 10(2), 197–227.
- Spatz, C. (2005). *Basic Statistics: Tales of distributions*. Belmont, CA: Wadsworth.
- Steadman, H. J., Silver, E., Monahan, J., Appelbaum, P. S., Robbins, P. C., Mulvey, E.P., ... Banks, S. (2000). A classification tree approach to the development of actuarial violence risk assessment tools. *Law and Human Behavior*, 24(1), 83–100.
- Steiner, B., & Giacomazzi, A. (2007). Juvenile waiver, boot camp, and recidivism in a northwest state. *The Prison Journal*, 87(2), 227–240.

- Stoodley, B. (2010, August). To measure or not to measure: The “Recidivism Dilemma.” *Corrections Today*, 72(4), 86–87.
- Tarte, J. M., Mackin, J. R., Cox, A., & Furrer, C. J. (2007). Juvenile crime prevention program 2005–2007 evaluation report. Portland, OR: NPC Research. Retrieved from http://www.oregon.gov/OCCF/Documents/JCP/JCP_2005-07_Final.pdf?ga=t
- Tobin, T. J., & Sugai, G. (1999, Spring). Using sixth-grade school records to predict school violence, chronic discipline problems, and high school outcomes. *Journal of Emotional and Behavioral Disorders*, 7(1), 40–53.
- Umbreit, M. S., & Coates, R. B. (1993). Cross-site analysis of Victim Offender Mediation in four states. *Crime & Delinquency*, 39(4), 565–585.
- Umbreit, M. S., Coates, R. B., & Vos, B. (2001). The impact of victim offender mediation: Two decades of research. *Federal Probation*, 65(3), 99–112.
- Umbreit, M. S., Coates, R. B., & Vos, B. (2007, March). Restorative justice dialogue: A multi-dimensional, evidenced-based practice theory. *Contemporary Justice Review*, 10(1), 23–41.
- Wan, T. T. H. (2002). *Evidence-based health care management: Multivariate modeling approached*. Boston, MA: Kluwer Academic.
- Von Hirsch, A. (1976). *Doing justice: The choice of punishments: Report of the Committee for the Study of Incarceration*. New York, NY: Hill & Wang.
- Wang, B. B., Wan, T. T. H., Clement, J., & Begun, J. (2001, September). Managed care, vertical integration strategies and hospital performance. *Healthcare Management Science*, 4(3), 181–191.
- Wheeldon, J. (2009, March). Finding common ground: Restorative justice and its theoretical constructions(s). *Contemporary Justice Review*, 12(1), 91–100.
- Wiebush, R. G. (1993). Juvenile intensive supervision: The impact on felony offenders diverted from institutional placement. *Crime and Delinquency*, 39(1), 68–89.
- Worden, R. E., & Myers, S. M. (1999). Police encounters with juvenile suspects. National Research Council’s Panel on Juvenile Crime: Prevention, Treatment and Control, National Institute of Justice (pursuant to Grant 95-IJ-CX-0071).
- Wordes, M., Bynum, T. S., & Corley, C. J. (1994). Locking up youth: The impact of race on detention decisions. *Journal of Research in Crime and Delinquency*, 31(2), 149–165.

- Yin, R. (1997). The abridged version of case study research: Design and method. In L. Brickman & D. J. Rog (Eds.), *Handbook of applied social science methods* (pp. 229–260). London, England: Sage.
- Zehr, H. (2002). *The little book of restorative justice*. Intercourse, PA: Good Books.
- Zhang, N. J., & Wan, T. T. (2007, Spring). Effects of institutional mechanisms on nursing home quality. *Journal of Health and Human Services Administration*, 28(4), 380–407.