Practitioner's Guide to COMPAS Core

April 4, 2019

equivant

Contents

Chapter 1

Introduction

The Practitioner's Guide provides an overview of the COMPAS Core Module in the Northpointe Suite. The Northpointe Suite is an integrated web-based assessment and case management system for criminal justice practitioners. The Northpointe Suite has modules designed for pretrial, jail, probation, prison, parole and community corrections applications. COMPAS Core is designed for both male and female offenders recently removed from the community or currently in the community. The Practitioner's Guide to COMPAS Core covers case interpretation, validity and reliability, and treatment implications. Most of the information provided is specific to COMPAS Core. Throughout this text we use the term COMPAS Core to distinguish an element (scale, typology, decile type) specific to COMPAS Core from general elements in the Northpointe Suite, such as scales found in both COMPAS Core and COMPAS Reentry.

COMPAS is a fourth generation risk and need assessment instrument. Criminal justice agencies across the nation use COMPAS to inform decisions regarding the placement, supervision and case management of offenders. COMPAS was developed empirically with a focus on predictors known to affect recidivism. It includes dynamic risk factors, and it provides information on a variety of well validated risk and need factors designed to aid in correctional intervention to decrease the likelihood that offenders will reoffend.

COMPAS was first developed in 1998 and has been revised over the years as the knowledge base of criminology has grown and correctional practice has evolved. In many ways changes in the field have followed new developments in risk assessment. We continue to make improvements to COMPAS based on results from norm studies and recidivism studies conducted in jails, probation agencies, and prisons. COMPAS is periodically updated to keep pace with with emerging best practices and technological advances.

COMPAS has two primary risk models: General Recidivism Risk and Violent Recidivism Risk. COMPAS has scales that measure both dynamic risk (criminogenic factors) and static risk (historical factors). Additional risk models include the Recidivism Risk Screen and the Pretrial Release Risk Scale II.

Statistically based risk/need assessments have become accepted as established and valid methods for organizing much of the critical information relevant for managing offenders in correctional settings (Quinsey, Harris, Rice, & Cormier, 1998). Many research studies

have concluded that objective statistical assessments are, in fact, superior to human judgment (Grove, Zald, Lebow, Snitz, & Nelson, 2000; Swets, Dawes, & Monahan, 2000). COMPAS is a statistically based risk assessment developed to assess many of the key risk and need factors in adult correctional populations and to provide information to guide placement decisions. It aims to achieve these goals by providing valid measurement and concise organization of important risk/need dimensions. Northpointe recognizes the importance of case management and supports the use of professional judgment along with actuarial risk/need assessment. Following assessment, a further goal is to help practitioners with case plan development/implementation and overall case management support.

In overloaded and crowded criminal justice systems, brevity, efficiency, ease of administration and clear organization of key risk/need data are critical. COMPAS was designed to optimize these practical factors. We acknowledge the trade-off between comprehensive coverage of key risk and criminogenic factors on the one hand, and brevity and practicality on the other. COMPAS deals with this trade-off in several ways; it provides a comprehensive set of key risk factors that have emerged from the recent criminological literature, and it allows for customization inside the software. Therefore, ease of use, efficient and effective time management, and case management considerations that are critical to best practice in the criminal justice field can be achieved through COMPAS.

1.1 Overview for Practitioners

COMPAS Core is comprised of a total of forty-three scales, including four higher order scales that use items from several domains and seventeen scales from the women's risk/needs assessment (WRNA) developed by Van Voorhis, Wright, Salisbury, and Bauman (2010). This document provides an overview of COMPAS Core. Supplemental materials are available that provide details about the scales not covered in the Practitioner's Guide (see, e.g., "Measurement and Treatment Implications of the COMPAS Core Scales").

The COMPAS Core assessment is designed to be configurable by the user at decision points within the local criminal justice system and with different populations. For example, Pre-trial Services may choose to use only the Pretrial Release Risk Scale II to make recommendations to the court regarding pretrial release. Probation may then use the Violent Recidivism Risk and General Recidivism Risk Scales to "triage" their caseloads by recidivism risk, and choose to complete the full assessment only on the higher risk individuals. The full assessment provides a holistic view of the person to address supervision and treatment needs for rehabilitation.

Chapter 2

Case Interpretation

2.1   Introduction

This chapter gives an introduction to the interpretation of a COMPAS assessment. After completing an assessment in the COMPAS software, the practitioner will generally interpret the bar chart that displays scale scores. The bar chart indicates in what areas the person scores higher or lower – that is, which risks or criminogenic needs may exist. The practitioner will also interpret the type assigned by the typology if enabled by the site. The implications for treatment and intervention are discussed in Chapter 4.

Collecting assessment information is important, yet the information is only helpful when we can make sense of it and understand how it can inform our case planning and interaction with the offender. Interpretation skills and activities include accessing and using:

1. The assessment results

2. The criminological theories used in COMPAS

3. The Typologies

A model that everyone can relate to is the medical model for interpretation of information gathered on a person. Think about the different steps taken in the medical field to find a solution to an illness or a problem. When you don't feel well and you go to the doctor, what is the first thing that the doctor does – Asks about symptoms: When did they start? How severe are they? She asks about your medical history: Are you taking any medications? Have you had this or a similar problem before? And, she runs tests, takes your temperature, takes your blood pressure, takes blood samples, orders MRIs, etc. What does she do with all of this information? She makes a diagnosis and prescribes an effective treatment.

Case interpretation involves connecting the dots to understand the relationship between a person's criminal behavior and her history, beliefs, and skills.

2.2 COMPAS Scores

The COMPAS assessment system consists of predictive risk scales for risk prediction and separate need scales for identifying program needs in the domains of employment, housing, substance abuse, and others. Agencies commonly adhere to the risk principle to target individuals for treatment programs who have high recidivism risk scores and high need for treatment (e.g., high substance abuse scores).

2.3 Levels of Interpretation

Skills and issues to consider when interpreting assessment information:

1. Interpretation is a skill that needs to be honed over time.

2. People are complex and multi-faceted. Interpretation is hard, yet is necessary for understanding behavior and for determining strategies for intervention.

3. From research in the field we have several criminological theories to help us understand the paths to criminal behavior.

There are different levels of interpretation.

1. **Level 1:** "Big bars, bad – little bars, good." Crime-producing issues are viewed largely in isolation, thus disregarding the influence high-scoring needs have on one another. This is a simplistic interpretation that fails to consider a chain of possible precursors and antecedents. It is, however, a good place to start, by identifying the areas of need for further consideration.

2. **Level 2:** Helps strengthen the interpretation process beyond Level 1 by identifying criminogenic factors that are interrelated. In particular, Level 2 begins the process of looking at areas of need that influence one another. Palmer (1994) identified three areas of commonality: environmental issues, skill deficiencies, and cognitive/mental health/psychological areas. This level of interpretation allows practitioners to begin developing interventions that might address clusters of needs, rather than individual needs in isolation of others.

3. **Level 3:** This is a fully integrated interpretation, using criminological theories to explain patterns of criminal behavior and help practitioners begin understanding possible underlying causes or contributors to the person's behavior. This approach enables the practitioner to consider a mix of explanatory theories that help "connect the dots" of need and other influencing factors to paint a picture of the individual's pathway to crime.

The needs measured by the scales are often interwoven and co-occurring. Accurately interpreting a COMPAS bar chart requires the practitioner to take into account all the high scoring needs. Criminological theories provide a framework to help understand the interrelationship between the different needs.

2.4   Criminological Theories

People are complex creatures. To obtain a holistic picture of an individual, salient life events and influences must be considered. Criminological theories explain how people become involved in criminal behavior and may provide guidance for effective interventions. Several important criminological theories are outlined below.

Social Learning Theory

1. This theory matches the traditional way we think about learning through modeling of behavior.

2. The basic principle of the theory is that behavior is modeled, imitated, and if reinforced, then likely to occur again.

Sub-Culture Theory

1. The theory was developed from the Chicago School on Gangs.

2. The theory was developed to explain delinquency and gang behavior.

3. The theory suggests that norms are transmitted through social interactions.

4. Norms in subcultures are different than those in the main culture.

5. Certain behaviors (crime, substance abuse) become the cultural norm within the subculture.

6. All individuals in society are driven toward economic success. Some subcultures aim to achieve that success through illegitimate means.

7. Fischer (1995) defines subculture as "a large set of people who share a defining trait, associate with one another, are members of institutions associated with their defining trait, adhere to a distinct set of values, share a set of cultural tools and take part in a common way of life" (p. 544).

Control/Restraint Theory

1. This theory suggests there are different types of control. These include internal control (bonding to values, beliefs, etc.), external control (bonds to family, friends, social networks, co-workers), and psychological control (emotional attachments, cognitions, etc.).

2. The lower an individual's level of social bonding (or less pro-social) and self control, the more crime-prone they will become (less to lose).

3. Or, they may be bonded to antisocial social norms values and associations, and their level of status depends on adherence to the restraints of that norm group.

Sociopathic/Socialization Breakdown Theory

1. Within this theory lies the concept of the sociopathic offender, which has more layers than the commonly stated "criminal personality."

2. Sociopathic is a specific personality disorder. Personality disorders can be described as a person's world view. A person with a personality disorder does not usually see themselves as needing help to remedy their behavior and typically blames consequences on other people and events.

3. A sociopath is characterized by selfishness, ruthlessness, and the inability to feel guilt or empathy.

4. This cluster of deviant personality traits and behaviors may not include criminal behavior.

Criminal Opportunity Theory (including Routine Activity)

1. This theory draws on the economic theory of markets to describe and predict criminal behavior.

2. The theory suggests that if you alter the quality of opportunity for crime you will reduce criminal behavior.

3. Both individual and environmental factors across time affect criminal acts.

4. The convergence in time and place of a motivated offender, suitable target, and absence of guardianship are strong predictors of criminal behavior.

5. Crime is most likely to occur in the presence of a suitable target (victim) and a motivated offender, and in the absence of inhibiting factors (law enforcement, neighbor, witnesses).

Social Strain Theory

1. This sometimes is referred to as the "means–end" theory of deviance.

2. Crime breeds in the gap between culturally induced aspirations and structurally distributed possibilities for success.

3. It is the combination of cultural emphasis and social structure which produces intense pressure for deviation-criminal behavior.

4. This is an economic explanation for crime. Crime occurs largely in poverty-stricken areas where opportunities to attain the "American Dream" by legitimate means is blocked, producing frustration and a desire to pursue monetary success by any means necessary.

2.5   AIPIE

Interpretation and the related events around case management can be a complex set of activities for professionals. One model that helps to explain the procedures of evidence-based practice is known as AIPIE. The AIPIE model is sequenced so that information triggers decisions which trigger actions.

A = Assessment (COMPAS or other tool)

I = Interpretation of the results

P = Plan, create an action plan based on the information gathered

I = Implement the plan

E = Evaluate the results of the actions and outcomes

The AIPIE model is linear and cyclic, that is, the steps are sequential and inform ongoing practice.

Risk and need scales have been discussed at length in this document. The other element to consider for supervision is responsivity. An offender's responsivity, or any person who is considering making some kind of change, can be understood as their level of readiness and their skill set to make the changes. Responsivity to intervention includes the person's motivation for change and the type of intervention offered. If the intervention does not fit the need, then responsivity factors are lost. If there is good fit, then there is better chance for success.

2.6   Basic Descriptive Information for the Scales

The scales are divided into two categories:

1. The Need Scales provide measures of relatively simple constructs (e.g., financial problems). These scales are not meant to be predictive but aim simply and accurately to describe the offender.

2. The Risk Scales were developed using methods and strategies for predictive modeling. The purpose of the risk scales is prediction - the ability to discriminate between offenders who will and will not recidivate.

2.7 Conversion of Raw Scale Scores to Decile Scores

The COMPAS scale scores are transformed into decile scores. Deciles are obtained by ranking the scale scores of a normative group in ascending order and then dividing these scores into ten equal sized groups. Deciles range from 1 (lowest) to 10 (highest). These scores thus proceed in roughly 10% steps from lowest to highest (1 through 10). A decile rank of 1 indicates that the scale score is in the lowest 10% of all scores in the normative group. A decile rank of 2 places the scale score above 10% and below 20% of the scores, and so on, up to a decile of 10, which places the scale score in the top 10% of all scores in the normative group.

In general the decile rank has the following interpretation:

- 1 – 4: scale score is low relative to other offenders in norm group.

- 5 – 7: scale score is medium relative to other offenders in norm group.

- 8 – 10: scale score is high relative to other offenders in norm group.

Note, however, that the location of the decile cut-points vary depending on the type of COMPAS scale. Table 2.1 shows the cutting points for each type of COMPAS scale. Table 2.2 lists each COMPAS scale and its type.

Table 2.1: Cutting Points for COMPAS Scale Types.

Type 1	Low (1-4)	Medium (5-7)	High (8-10)
Type 2	Unlikely (1-2)	Probable (3-4)	Highly Probable (5-10)
Type 3	Unlikely (1-5)	Probable (6-7)	Highly Probable (8-10)
Type 4	Unlikely (1-4)	Probable (5-7)	Highly Probable (8-10)

The decile cutting points for the scale scores in the COMPAS Core composite norm group (n=7381) are shown in Table 2.3. The column labeled D1 contains the cut-off for the first decile, D2 the cut-off for the second decile, and so on. Thus, for the Criminal Personality Scale (CrimPers), roughly one-tenth of the offenders scored 23 and lower, another one-tenth scored 24 through 25, and so forth. If a score covers more than one decile, we use the convention of assigning it to the lower decile category. For instance, 30% of the composite sample have a score of 0 on the History of Noncompliance Scale (HistNonC), covering D1 through D3 in the table, but this score is assigned to the lower decile (D1). This characteristic is associated with the granularity of certain COMPAS Core scales, which is discussed in the next section.

Table 2.2: COMPAS Core Scales and Types.

Scale	Scale Type
Violent Recidivism Risk	1
General Recidivism Risk	1
Pretrial Release Risk	1
Criminal Involvement	1
History of Noncompliance	1
History of Violence	1
Current Violence	1
Criminal Associates/Peers	4
Substance Abuse	2
Financial Problems/Poverty	3
Vocational/Education Problems	3
Criminal Thinking	3
Family Criminality	3
Social Environment Problems	3
Leisure and Recreation	3
Residential Instability	3
Social Adjustment Problems	3
Socialization Failure	3
Criminal Opportunity	3
Criminal Personality	3
Social Isolation	3

Table 2.3: Decile Cut-Points for COMPAS Core Scales in the Composite Norm Group.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
CrimInv	1.0	3.0	6.0	7.0	9.0	11.0	12.0	13.0	15.0	19.0
HistNonC	0.0	0.0	0.0	1.0	2.0	2.0	3.0	5.0	7.0	21.0
HistViol	1.0	1.0	1.0	1.0	1.0	2.0	3.0	4.0	6.0	20.0
CurrViol	8.0	8.0	8.0	8.0	8.0	8.0	8.0	9.0	11.0	14.0
CassPeer	7.0	8.0	8.0	9.0	10.0	10.0	11.0	12.0	15.0	22.0
SubAbuse	10.0	10.0	11.0	12.0	13.0	14.0	14.0	15.0	16.0	20.0
Financ	5.0	6.0	7.0	8.0	9.0	9.0	10.0	11.0	12.0	15.0
VocEd	14.0	15.0	16.0	17.0	18.0	19.0	21.0	22.0	24.0	30.0
FamCrim	6.0	6.0	6.0	7.0	7.0	8.0	8.0	9.0	10.0	12.0
SocEnv	6.0	6.0	6.0	6.0	6.0	7.0	7.0	8.0	10.0	12.0
Leisure	5.0	5.0	6.0	7.0	8.0	9.0	10.0	12.0	14.0	17.0
ResInst	9.0	11.0	11.0	12.0	14.0	15.0	16.0	18.0	20.0	30.0
SocAdj	16.0	17.0	18.0	19.0	20.0	21.0	22.0	24.0	25.0	35.0
EJuvSoc	10.0	11.0	12.0	13.0	13.0	14.0	15.0	17.0	19.0	30.0
CrimOpp	15.0	17.0	18.0	20.0	21.0	22.0	24.0	25.0	28.0	40.0
Soc.Isol	11.0	13.0	15.0	16.0	18.0	19.0	21.0	23.0	26.0	40.0
CrimAttC	13.0	15.0	18.0	20.0	21.0	22.0	23.0	25.0	28.0	50.0
CrimPers	23.0	25.0	27.0	29.0	30.0	32.0	34.0	36.0	40.0	59.0
PretrialRisk	2.89	3.08	3.24	3.39	3.54	3.69	3.86	4.08	4.38	8.01
ViolRecidRisk	−2.90	−2.50	−2.20	−2.00	−1.70	−1.50	−1.20	−1.00	−0.60	1.90
GenRecidRisk	−1.30	−0.90	−0.70	−0.40	−0.20	0.00	0.20	0.40	0.60	1.90

2.8   Interpreting Decile Scores

It is important to note that decile scores can only be interpreted in a relative sense, and are always linked to the norm group. If, for example, the norm group that is referenced for decile scoring of the Violent Recidivism Risk Scale happens to consist of offenders at high risk of violent recidivism, then low decile scores would not necessarily indicate low risk of violent recidivism. Similarly, if the norm group happens to consist mainly of offenders with low risk of violent recidivism, the decile scores for Violent Recidivism Risk would be biased in the other direction – high scores could be associated with individuals who are actually not high risk for violent recidivism.

It is also important to note that for some scales, it is not always possible to break the sample into ten groups of exactly equal size. Hence, for some scales it was necessary to skip over some decile scores.

When it was not possible to divide the sample into ten groups, an algorithm was used to identify cutting points that divided the offenders into as many roughly equal-sized groups as possible and that used the full range of decile values (i.e., 1-10).

The issue of clumping affects a limited number of scales. Overall, the use of decile ranks has clear advantages over the use of raw scale scores in terms of interpretability. Low scores (e.g., 1 thru 4) directly reflect the lowest ends of the distribution, and high scores (e.g., 8 thru 10) reflect the highest ends of the distribution.

2.9   Norm Groups

The COMPAS Core normative data were sampled from over 30,000 COMPAS Core assessments conducted between January 2004 and November 2005 at prison, parole, jail and probation sites across the United States. The Core Norm Group was compiled to obtain proportions of prison, parole, jail, and probation assessment data that reflect proportions of adult correctional populations in the criminal justice system. Based on recent criminal justice statistics, 21.6% of persons under adult correctional supervision during 2011 were in prison, 12.2% were on parole, 10.5% were in jail, and 56.9% were on probation (Bureau of Justice Statistics, 2012). The Composite Norm Group consists of assessments from state prisons and parole agencies (33.8%); jails (13.6%); and probation agencies (52.6%). The Core Norm includes 7,381 offenders. Men represent 76.9% of the Core Norm Group (n=5,681), and women represent 23.1% of the Core Norm Group (n=1,700). The median age at assessment is 31.0 ($M = 32.6$) in the Core Norm Group. The racial composition of the Core Norm Group is 61.6% Caucasian, 24.9% Black, 10.3% Latino and 3.2% other racial groups.

In the current version of COMPAS Core, scale scores can be referenced to the scale distributions of eight normative subgroups: (1) male prison/parole, (2) male jail, (3) male probation, (4) male composite, (5) female prison/parole, (6) female jail, (7) female probation and (8) female composite.

COMPAS Core norm data are evaluated through client norm studies. Agency-specific norm groups are developed for some clients.

Chapter 3

COMPAS Validity and Reliability

In this section we summarize research findings from multiple studies that demonstrate COMPAS Core is reliable (test-retest and internal consistency), that its scales measuring needs have construct validity and behave consistently and that its risk scales have predictive validity.[1] An overall conclusion is that COMPAS Core was found to be reliable and has good predictive and construct validity.

Northpointe has an established history of working in partnership with our clients to advance knowledge and practice. From our early work in jail classification to our recent partnerships with the California Department of Corrections and Rehabilitation (CDCR) and the University of Cincinnati, Northpointe leverages the opportunity of public and private partnership to test and advance knowledge. Our research and evaluation findings are publicly shared through conference papers, technical reports, peer-reviewed articles and book chapters to advance the availability of current information for use in practice.

3.1   Predictive Validity of the COMPAS Risk Scales

COMPAS distinguishes between risk scales (designed to predict recidivism) and need scales (designed to measure needs, inform case plans and identify intervention targets). This approach of separating risk and needs aligns with current best practices in risk assessment (C. Baird, 2009; S. D. Gottfredson & Moriarty, 2006). The risk scales are used for classification and forecasting. The risk scales should have good discriminative ability and predictive ability (e.g., Singh, 2013). COMPAS has two main risk scales: General Recidivism Risk Scale (GRRS) and Violent Recidivism Risk Scale (VRRS). Some researchers believe risk scales should be dynamic (composed of dynamic, criminogenic needs) so that one can measure change in risk of recidivism over time. Others have argued that risk scales should be composed of static criminal history factors available in criminal justice management information systems, arguing that static risk scales are more objective, reliable, and efficient (Barnoski & Drake, 2007). Our risk scales make limited use of dynamic variables.

---

[1]The General Recidivism Risk and Violent Recidivism Risk scales are used in both COMPAS Core and COMPAS Reentry. Identical linear equations are used to calculate the risk scales in the two applications.

There are different statistical approaches to predictive modeling. Machine learning methods are highly flexible and are usually preferred in applications where there is a complex relationship between predictors and outcomes (Berk, 2012; Berk & Bleich, 2013). Some methods are less flexible but the models they generate are easier to interpret. There are tradeoffs of predictive performance and interpretability between methods (James, Witten, Hastie, & Tishirani, 2013; Brennan & Oliver, 2013). Our methods for developing and validating the GRRS were strongly influenced by the research of John Copas and colleagues who have developed an outcomes-based recidivism scale for England and Wales (Copas & Marshall, 1998). The methods used to develop both risk scales are described in various books on regression modeling and machine learning (e.g., Harrell, 2001; Hastie, Tibshirani, & Friedman, 2008; Kuhn & Johnson, 2013).

Northpointe is committed to testing, evaluating, and improving our risk scales. The General Recidivism Risk and Violent Recidivism Risk scales have been validated with prospective outcomes in new samples in several different studies since they were first developed.

When possible we include an outcomes component in the pilot test of COMPAS in new jurisdictions. This component is designed to evaluate the predictive validity of the risk scales. In 2006 we conducted pilot tests in the New York Office of Probation and Correctional Alternatives (OPCA), the New York State Division of Parole (NYSDP), and the Michigan Department of Corrections (MDOC). These three pilot studies all had outcomes studies built into them. In 2008 we conducted outcomes studies at all three sites using their pilot data. We also conducted separate studies in the California Department of Corrections and Rehabilitation (CDCR) and the OPCA. This latter study was published in the Journal of Criminal Justice and Behavior (Brennan, Dieterich, & Ehret, 2009).

Table 3.1 below shows the results of subsequent tests of the predictive validity of the COMPAS risk scales. These outcomes studies were conducted on large samples in the Michigan Department of Corrections (Brennan & Dieterich, 2008; Dieterich, Oliver, & Brennan, 2011; Dieterich, Brennan, & Oliver, 2011); New York State Office of Probation and Correctional Alternatives (Brennan & Dieterich, 2009; Brennan et al., 2009; Lansing, 2012); California Department of Corrections and Rehabilitation (Farabee, Zhang, Roberts, & Yang, 2010); Broward County Sheriff's Office (Flores, Bechtel, & Lowenkamp, 2016; Blomberg, Bates, Mann, Meldrum, & Nedelec, 2010); Santa Barbara County Probation Department (Dieterich, Mendoza, Hubbard, Ferro, & Brennan, 2017); and Riverside County Probation Department (Dieterich, Mendoza, Hubbard, Ferro, & Brennan, 2018).

Table 3.1 shows the area under the receiver operating characteristic curve (AUC) for the GRRS and VRRS from several outcomes studies. "The receiver operating characteristic (ROC) curve is currently the best-developed statistical tool for describing the performance" of a risk scale (Pepe, 2003, p. 66). The AUC is the most widely used measure of discrimination ability in criminal justice, psychology, medicine, and related fields. There are various conventions for interpreting the magnitude of the AUC. One of the more liberal grading scales is provided by Desmarais and Singh (2013): AUC values of 0.50 to 0.54 are poor, 0.55 to 0.63 are fair, 0.64 to 0.70 are good, and 0.71 to 1.00 are excellent. The consensus in the field of recidivism research seems to be that AUC values below 0.65 are poor, 0.65 to 0.69 are fair, 0.70 to 0.75 are good, and 0.76 and above are excellent. The sizes of AUC that

constitute poor, fair, good, or excellent discrimination depend on the area of study and the outcome (e.g., Swets, 1988). Note that for arrest, felony arrest, noncompliance and return to prison outcomes, the GRRS is tested. For person offense arrests the VRRS is tested.

The results of these studies indicate that the COMPAS risk scales generally fall into the moderate to good range of discrimination ability. They also indicate that COMPAS generally meets or exceeds the AUC values produced by competitive instruments such as the LSI-R (see p.).

Table 3.1: Summary of AUC results for the General Recidivism Risk Scale (GRRS) and Violent Recidivism Risk Scale (VRRS) in several outcomes studies.

Study	N	Year	Any Arrest	Felony	Person	Supervision Failure
NY Probation[a]	(n=2,328)	2009	0.680	0.700	0.710	
NY Probation[b]	(n=13,993)	2012	0.710			
MDOC Reentry[c]	(n=25,347)	2011		0.710	0.700	0.690
MDOC Probation[d]	(n=21,101)	2011		0.670	0.740	0.710
CDCR Reentry[e]	(n=25,009)	2010	0.700		0.650	
Broward Jail[f]	(n=6,172)	2016	0.710		0.710	
Mental Health Court[g]	(n=242)	2016	0.730			
Santa Barbara Probation[h]	(n=5,363)	2017	0.722		0.672	0.702
Riverside Probation[i]	(n=4,435)	2018	0.694		0.636	0.692

[a] (Brennan et al., 2009).
[b] (Lansing, 2012).
[c] (Dieterich, Brennan, & Oliver, 2011).
[d] (Dieterich, Oliver, & Brennan, 2011).
[e] (Farabee et al., 2010).
[f] (Flores et al., 2016).
[g] (Reich et al., 2016).
[h] (Dieterich et al., 2017).
[i] (Dieterich et al., 2018).

Differential Validity

A few independent outcomes studies have examined the predictive validity of the COMPAS risk scales for gender and racial groups. Brennan et al. (2009) found that the COMPAS risk scales performed equally well for African American and Anglo men at discriminating recidivists in a probation sample. A prior study examined the discriminative ability of the GRRS for different ethnic groups, and that study reported much weaker results for African American men (Fass, Heilbrun, DeMatteo, & Fretz, 2008). In predicting rearrest within 1 year of release, Fass et al. (2008) reported AUCs for the GRRS of 0.81 for Whites, 0.67 for Hispanics, 0.48 for African Americans, and 0.53 for the total sample assessed with COMPAS (N = 276). However, their study has at least one critical weakness that renders its findings unreliable. Their small overall sample size and base rates resulted in extremely small effective

sample sizes for the ethnic groups (African American = 36, Latino = 4, Anglo = 1). These effective sample sizes are too small for ROC analysis and unreliable results were obtained.

Angwin, Larson, Mattu, and Kirchner (2016) claimed to have found evidence that the COMPAS risk scales were biased against African Americans in a sample of defendants in Broward County, Florida. Results of subsequent secondary analyses of the Broward County study data conducted by Northpointe researchers (Dieterich, Mendoza, & Brennan, 2016) and independent researchers (Flores et al., 2016) completely refuted the findings of Angwin et al. (2016). Both of the secondary studies found that the GRRS and VRRS performed equally well for African American and Anglo defendants. Flores et al. (2016) found that the GRRS and VRRS were good predictors for arrests and violent arrests, respectively. They also determined that the risk scales predicted equally well for African American and Anglo defendants. Flores et al. obtained the following AUCs for the GRRS decile score predicting any arrest: 0.70 for African American, 0.69 for Anglo and 0.71 in the sample overall. Flores et al. obtained the following AUCs for the VRRS decile score predicting any violent arrest: 0.70 for African Americans, 0.68 for Anglos and 0.71 in the sample overall. Table 3.2 shows the results obtained by Dieterich et al. in the same data analyzed by Flores et al. and Angwin et al.

Table 3.2: AUC results for the General Recidivism Risk Scale (GRRS) decile scores and Violent Recidivism Risk Scale (VRRS) decile scores in the data analysis samples used by Angwin and colleagues.

	Sample	n	events	base rate	AUC	Lower 95% CI	Upper 95% CI
GRRS							
	Anglo	2103	822	0.39	0.69	0.67	0.72
	African American	3175	1661	0.52	0.70	0.69	0.72
	All	6172	2809	0.46	0.71	0.70	0.72
VRRS							
	Anglo	1459	174	0.12	0.68	0.64	0.73
	African American	1918	404	0.21	0.71	0.68	0.74
	ALL	4020	652	0.16	0.72	0.70	0.74

Note. GRRS outcome is a misdemeanor or felony arrest. VRRS outcome is a violent misdemeanor or felony arrest.

Farabee et al. (2010) report separate findings for men and women in a California Department of Corrections and Rehabilitation sample of persons released onto parole with two years of follow-up. They present a matrix with bivariate correlation coefficients for the GRRS and any arrest separately for men and women. The Pearson product moment correlation between the General Recidivism Risk Score and any arrest is 0.32 for men and 0.32 for women, thus providing evidence that the risk scale has similar predictive validity for men and women.

Table 3.3 displays AUCs for the any arrest outcome for the data set used by Farabee et al. (2010). The AUCs in the table give an indication of how well the GRRS discriminates the offenders who were rearrested from those who were not rearrested. The results are for the

entire sample (All) and for Men, Women, Anglo, African American, and Latino groups. The values for the AUCs in the different groups are very nearly the same.

Table 3.3: AUCs for the General Recidivism Risk Scale (GRRS) for a California prison sample. The AUCs are calculated separately for the different subgroups defined by gender and ethnicity/race. The lower (Low) and upper bounds (High) of the 95 percent confidence interval are displayed along with the number of failures (Nfail) and the number of offenders in the sample (N).

	AUC	Low	High	Nfail	N
Men	0.71	0.70	0.71	14819	21015
Women	0.69	0.67	0.71	1595	2638
Anglo	0.70	0.69	0.71	4683	7268
African American	0.69	0.67	0.70	4813	6447
Latino	0.71	0.70	0.72	5980	8514
All	0.70	0.70	0.71	16414	23653

Table 3.4 displays AUCs for a large reentry sample from the Michigan Department of Corrections. The outcome was any arrest within three years following release from prison into the community. Offenders who did not have opportunity to fail in a three year period were excluded from the sample. As in the previous analysis, the results are for the entire sample (All) and for Men, Women, Anglo, African American, and Latino groups.

The AUCs in Table 3.4 vary from 0.71 (African American) to 0.78 (Latino). The effective sample size for the Latino group is relatively small, which results in a broad 95% confidence interval. The AUCs for Men (0.73) and Women (0.74) are nearly the same. The AUCs for Anglo (0.75) and African Americans (0.71) do noticeably differ but both values are reasonably high.

These results taken together are encouraging. They suggest that the predictive validity of the GRRS is good overall and nearly equivalent for men and women, and for Anglo, African American and Latino offenders.

Table 3.4: AUCs for the General Recidivism Risk Scale (GRRS) and any arrest outcome for a Michigan reentry sample. The AUCs are calculated separately for the different subgroups defined by gender and ethnicity/race. The lower (Low) and upper bounds (High) of the 95 percent confidence interval are displayed along with the number of failures (Nfail) and the number of offenders in the sample (N).

	AUC	Low	High	Nfail	N
Men	0.73	0.72	0.74	5427	13439
Women	0.74	0.71	0.77	341	961
Anglo	0.75	0.74	0.76	2807	7177
African American	0.71	0.69	0.72	2720	6571
Latino	0.78	0.73	0.84	89	289
All	0.73	0.72	0.74	5768	14400

Table 3.5 displays AUCs for the GRRS from a validation study conducted for the Santa Barbara County Probation Department (Dieterich et al., 2017). The outcome was any arrest within 3 years of the probation intake assessment. The results are for the entire sample (All) and for Men, Women, Anglos, and Latinos.

AUC results for the GRRS show good discriminative ability for women, men, Latinos, and Anglos. The AUCs for Men (0.72) and Women (0.72) are not statistically different. There are moderate and significant differences between the AUC for Latinos (0.74) and the AUC for Anglos (0.70).

Table 3.5: AUCs for the General Recidivism Risk Scale (GRRS) and any arrest outcome in the Santa Barbara County Probation outcomes study sample. The AUCs are calculated separately for the different subgroups defined by gender and ethnicity. The lower (Low) and upper bounds (High) of the 95 percent confidence interval are displayed along with the number of failures (Nfail) and the number of offenders in the sample (N).

	AUC	Low	High	Nfail	N
Men	0.72	0.71	0.74	2823	4277
Women	0.72	0.69	0.75	717	1086
Anglo	0.70	0.68	0.73	1440	2149
Latino	0.74	0.72	0.76	1759	2706
All	0.72	0.71	0.74	3543	5363

The validity of the GRRS has also been demonstrated in a specialty court setting. Reich et al. (2016) conducted a COMPAS validation study in a racially diverse sample of 242 Mental Health Court participants at three sites in New York City: Brooklyn Mental Health Court, Bronx Mental Health Court, and Queens Felony Mental Health Court. The study examined the ability of the COMPAS GRRS to accurately discriminate recidivists from non-recidivists at one-year and two-years. The GRRS Decile Score was found to have good discriminative ability at one-year (AUC = 0.70) and two-years (AUC=0.73). Serin and Lowenkamp (2015) identified the risk instruments that are best suited for use by Drug Courts on the basis of widely accepted validity criteria. The General Recidivism Risk Scale met all validity criteria and was one of only three instruments recommended for use by Drug Courts.

Differential Validity and Fairness Criteria

Several different fairness schemes based on classification statistics (true positive rate, false positive rate, positive predictive value, negative predictive value, and selection ratio) have been proposed in the criminal justice and computer science literature.

The true positive rate (tpr) is the percentage of persons that recidivated that have a high risk score. The false positive rate (fpr) is the percentage of persons that did not recidivate that have a high risk score. The selection ratio is the percentage of cases that fall into the High Level. The tpr and fpr are measures of diagnostic accuracy that assess discriminative ability.

The positive predictive value (PV+) is the probability that a person with high risk score will recidivate. The negative predictive value (PV-) is the probability that a person with a low risk score will not recidivate. The PV+ and PV- assess predictive ability.

Sensitivity (true positive rate) and Specificity (true negative rate) quantify the diagnostic accuracy of the risk scale and the predictive values quantify its clinical value (Pepe, 2003). A useful prediction will have a PV+ that is greater than the base rate and a PV- that is greater than 1 minus the base rate. A perfect test will predict the outcome perfectly with PV+ = 1 and PV- = 1. The predictive values depend on the accuracy of the test and the base rate of failure.

Practitioners in criminal justice settings are most interested in the probability that an individual with a high risk score will be arrested in the future (PV+), as opposed to the probability that an individual who has already been arrested has a high risk score (tpr). The purpose of administering a risk scale is to use the results to assess an individual's risk of re-offending at the time of assessment. The tpr and fpr are of no practical use to a practitioner in a criminal justice agency who is assessing an individual's probability of re-offending. The practitioner does not know at the time of the assessment if the individual is a recidivist or not. The tpr and fpr cannot be directly applied to an individual at the time of assessment (see Linn, 2004, for example).

Berk (2016) presented a primer that organized the different types of fairness that arise when risk scales are used in samples with different gender or ethnic groups. Three of those types are salient for most applications. The risk scale demonstrates *predictive fairness* if the selection ratio is the same for both groups, *use fairness* if the complements of the positive predictive values are the same in both groups, and *model fairness* if the false positive and false negative rates are the same for both groups.[2]

We can use the GRRS results from the Santa Barbara County Probation outcomes study for men and women and for Anglos and Latinos to demonstrate risk assessment fairness. The results for the GRRS show that the false positive rate at the High cut point is similar for women (0.16) and men (0.13) and for Anglos (0.14) and Latinos (0.14), providing evidence of *model error fairness.* The positive predictive value at the High cut point is similar for women (0.83) and men (0.86) and for Anglos (0.85) and Latinos (0.85), providing evidence of the salient fairness criterion in probation practice and termed variously as *use fairness* (Berk, 2016), *predictive parity* (Dieterich et al., 2016), or *calibration* (Kleinberg, Mullainathan, & Raghavan, 2016). The selection ratio at the High cut point is similar for women (0.32) and men (0.32) and for Anglos (0.31) and Latinos (0.33), providing evidence of *predictive fairness* (Berk, 2016). As reported in the previous section on ROC results, the AUC is similar for women (0.72) and men (0.72) and somewhat higher for Latinos (0.74) compared with Anglos (0.70). The combined results for the GRRS provide good evidence of fairness for women and men and for Anglos and Latinos.

---

[2]The false negative rate (fnr) is just the complement of the true positive rate (tpr). If the tpr is the same in both groups, then the fnr is the same.

Effect of Group-Specific Base Rates on Fairness Criteria

Dieterich et al. (2016) conducted a simulation analysis to assess the effects of differences in the risk scale distribution and base rate on the false positive and false negative rates. Results of their analysis indicate that larger differences in mean scores for two groups correspond to larger base rate differences as well as higher false positive rates and lower false negative rates for the group with the higher mean score (i.e., the group with the higher base rate). This is the same pattern of results reported by Angwin et al. (2016). Dieterich et al. (2016) stated that this pattern does not show evidence of bias, but rather is a natural consequence of using unbiased scoring rules for groups that happen to have different distributions of scores. These results help to explain the effects of the relatively higher risk scores and higher base rates of African Americans on the false positive and false negative rates in the Angwin et al. (2016) study.

Following the Dieterich et al. (2016) study, Kleinberg et al. (2016) put forward three fairness properties of risk scales: calibration (same predictive value) within groups, balance for the negative class (generalization of the false positive rate at the score level) within groups, and balance for the positive class (generalization of the false negative rate at the score level) within groups. They formally demonstrated that the only way these three properties can be achieved simultaneously is if the risk scale predicts perfectly or the two groups have equal base rates.

Risk scales may exhibit race and gender effects because race and gender are correlated with the outcomes that risk scales are designed to predict (S. D. Gottfredson & Jarjoura, 1996). Disadvantage in the domains of employment, education, and housing stems from structural inequalities in our society. Constructs within these domains correlate with criminal behavior. In some respects many of the widely accepted criminogenic needs are indirect measures of disadvantage. Risk scale scores use inputs (prior arrest, age at first arrest) and predict outcomes (arrests) that are impacted by intense police practices in some geographical areas (Committee on Assessing Juvenile Justice Reform, 2013). These effects are at the heart of methodological controversies in criminology related to risk assessment and racial bias that have emerged in different contexts over the years.

Berk, Heidari, Jabbari, Kearns, and Roth (2018) conducted a thorough examination of risk assessment fairness in criminal justice settings and concluded:

> Except in trivial cases, it is impossible to maximize accuracy and fairness at the same time and impossible simultaneously to satisfy all kinds of fairness. In practice, a major complication is different base rates across different legally protected groups. There is a need to consider challenging trade-offs (p. 1).

Mendoza, Dieterich, Oliver, and Brennan (2016) developed such an approach that adjusts for the effects of intense police practices on recidivism risk scores and classification statistics when making decisions across thresholds of a recidivism risk scale. They use an example data set in which African American defendants have a higher base rate and higher risk scores relative to Anglo defendants. Their approach assumes that intense police practices increase arrest rates which shift the risk scores of African American defendants higher relative

to Anglos. With simulation analyses, they show that larger differences in mean scores for two groups correspond to larger base rate differences as well as higher false positive rates and lower false negative rates for the group with the higher mean score (i.e., the group with the higher base rate) when a single threshold is used. They demonstrate a principled approach using decision analysis that weighs in these prior effects of intense police practices to adjust the risk scale threshold for African Americans.

Examples of Validity Results for Different Tools and Outcomes

Here we provide examples of the AUCs obtained with other risk tools to help contextualize the findings of our studies. Perhaps the best known instruments are the Violence Risk Appraisal Guide [VRAG] (Quinsey et al., 1998); the Level of Services Inventory-Revised [LSI-R] (Andrews, Bonta, & Wormith, 2006); the Ohio Risk Assessment System [ORAS]) (Latessa, Lemke, Makarios, Smith, & Lowenkamp, 2010); the Static Risk Offender Need Guide for Recidivism [STRONG-R])(Hamilton et al., 2017); and the Psychopathy Checklist-Revised [PCL-R] (Hare, 1991). The AUC values for these instruments in recent studies are quite varied depending on the populations, outcome periods, and dependent variables used in specific studies.

VRAG: Quinsey et al. (1998) found an AUC of 0.76 in a large scale, multiyear recidivism study. Barbaree, Seto, Langton, and Peacock (2001) reported AUCs of 0.69 in predicting serious reoffending and 0.77 when predicting any re-offense for sex offenders. Kroner, Stadtland, Eidt, and Nedopil (2007) obtained an AUC of 0.70 in a study of re-offending among mentally ill offenders.

LSI-R: The LSI-R has been tested more than any other risk assessment tool used in criminal justice settings. Results from several meta-analyses indicate that the LSI-R has good discriminative ability for general and violent recidivism. In their review, Andrews et al. (2006) reported an average r_{pb} value of 0.36 for general recidivism from 74 effect sizes. For violent recidivism, Andrews et al. reported an average r_{pb} value of 0.25 from 26 effect sizes. Note that a point-biserial correlation of 0.35 approximately corresponds to an AUC of 0.70 for base rates near 50% (Rice & Harris, 2005). Other meta-analyses have found similarly good results for the LSI and its variants (Gendreau, Goggin, & Little, 1996; Vose, Cullen, & Smith, 2008; Smith, Cullen, & Latessa, 2009). Barnoski and Aos (2003) found AUCs of 0.64 - 0.66 for the LSI-R in predicting felony and violent recidivism among Washington State prisoners. Flores, Lowenkamp, Smith, and Latessa (2006) reported an AUC of 0.69 using the LSI-R to predict re-incarceration among federal probationers. Dahle (2006) reported an AUC of 0.65 using the LSI-R to predict violent recidivism. Barnoski and Drake (2007) reported an AUC of 0.65 using the LSI-R to predict felony sex recidivism.

ORAS: (Latessa et al., 2017) conducted a re-validation of the ORAS tools in Ohio. From that study, we review the arrest outcomes for the Community Supervision Tool (CST), Prison Intake Tool (PIT), and Reentry Tool (RT). For the CST they obtained AUCs of 0.62 for men (n=501) and 0.65 for women (n=492). For the PIT they obtained AUCs of 0.61 for men (n=246) and 0.67 for women (n=246). For the RT they obtained AUCs of 0.60 for men (n=255) and 0.66 for women (n=257). In Indiana, Latessa, Lovins, and Makarios

(2013) conducted a validation of the CST in a community supervision sample (n=626) and a validation of the RT in a prison reentry sample (n=362). The outcome was arrest within two years. For the CST, Latessa et al. found moderate effects with men ($r_{pb} = 0.29$) and weak effects with women ($r_{pb} = 0.12$). Lovins, Latessa, May, and Lux (2017) conducted a validation study of the CST in a community corrections sample in Texas (n=5,481). They obtained AUCs of 0.67 for men and 0.68 for women. Latessa, Lux, Lugo, and Long (2016) conducted a validation study of the ORAS CST in a large Massachusetts Probation Service sample [MPS] (n=10,548). The outcome was a new arraignment. In the MPS study they obtained AUCs of 0.65 for men (n=1,382) and 0.63 for women (n=9,076). We could find no independent validation studies for the Ohio Risk Assessment System (ORAS).

STRONG-R: Hamilton et al. (2017) report AUC results for three offense-specific risk scales developed and internally validated using bootstrap re-sampling in a large Washington State Department of Correction sample. The recidivism follow-up was two years. For men, the AUCs for the Violent, Property, and Drug risk scales were 0.74, 0.78, and 0.76. For women the AUCs for the Violent, Property, and Drug risk scales were 0.74, 0.74, 0.73. We could find no independent validation studies of the STRONG-R.

PCL-R: Predictive accuracy varied across studies. For example, a Swedish study of mentally ill violent offenders (Grann, Belfrage, & Tengstrom, 2000) found AUC levels of 0.64 - 0.75 based on various follow-up time frames. Barbaree et al. (2001) reported AUCs of 0.61, 0.65, and 0.71 for the PCL-R in predicting various recidivism outcomes among sex offenders.

3.2 Validity of COMPAS Core Need Scales

3.2.1 Criterion Validity

In contrast to the COMPAS risk scales, the COMPAS need scales have a separate purpose and were developed using different methods. The risk scales were developed using methods and strategies for predictive modeling.

The need scales are not meant to be predictive but aim simply and accurately to describe the offender along dimensions relevant for correctional practice. Research findings indicate that individuals involved in the criminal justice system often have problems and deficits in the domains of education, housing, employment, substance abuse, relationships, and cognition. The need scales should be valid and reliable measures of constructs in these domains and other aspects of the person-in-environment that represent potential targets for interventions. The need scales guide individualized decisions for case planning, including identifying targets and choosing interventions. Within some theoretical frameworks, needs are expected to be criminogenic, suggesting that they cause recidivism and that recidivism can be reduced if the criminogenic need is effectively addressed. But research results indicate many constructs in these domains are only modestly correlated with recidivism, and evidence of a causal link between needs, treatment, and recidivism is lacking (e.g., Monahan & Skeem, 2014). Within the risk, need, and responsivity framework, high risk and high need individuals are targeted for the most intensive interventions (Andrews, Bonta, & Hoge, 1990). Here we focus only

on correlations to demonstrate that the COMPAS Core need scales are relevant and useful measures for correctional practice.

The following tables show measures of association between the COMPAS Core scales and recidivism in large samples from two COMPAS outcomes studies. The results obtained in the respective COMPAS outcomes studies provide evidence of the criterion validity of the COMPAS Core scales. The results demonstrate that in general the COMPAS Core need scales measure factors associated with recidivism, and hence, they are useful measures of potential intervention targets. The results can be compared with the results from published studies. For example Barnoski and Aos (2003) conducted an outcomes study in a sample of 22,533 offenders and provide a table with similar measures of association between the LSI-R subscales and recidivism.

Table 3.6 shows measures of association between the COMPAS Core scales and any arrest within two years in the study sample used by Farabee et al. (2010). The sample consists of 23,635 soon-to-be-released inmates assessed with COMPAS Core who were followed for two years after release from prison. The first column shows the correlation between each COMPAS Core scale and recidivism. For correlations between a continuous variable (e.g. Voced, Subabuse, etc.) and a dichotomous variable (recidivism), we estimate the point biserial correlation (r_{pb}).[3] The point biserial correlation is mathematically equivalent to the Pearson product moment correlation (r). J. Cohen and Cohen (1983) provide the following conventions for interpreting r when both variables are continuous: 0.10 = small; 0.30 = medium; 0.50 = large. But r is sensitive to the base rate when one of the variables is dichotomous, and the conventions for small, medium, and large should be adjusted lower depending on the deviation from a 50% base rate (e.g., Rice & Harris, 2005). The base rate for arrest in the Farabee et al. study sample is 0.69, so the following interpretation for r_{pb} adjusted using the formulae in Rice and Harris can be used: 0.09 = small; 0.23 = medium; 0.35 = large. The next column shows the area under the receiver operating characteristic curve (AUC). The AUC is a rank measure indicating how well the respective scales discriminate recidivists from nonrecidivists. The AUC is more resistent to the base rate (e.g., Babchishin & Helmus, 2016). An AUC equal to 1 indicates that the scale discriminates perfectly. An AUC equal to 0.50 indicates that the scale does not discriminate any better than chance. By convention an AUC of 0.70 is regarded as good in criminal justice settings. The AUC is 0.60 for the Criminal Associates Scale - a modest result if this were a standalone risk scale, but for a need scale, the result indicates good criterion validity. The last column shows the odds ratio. The odds ratio indicates how much the odds of recidivating change for every one-unit increase in the respective COMPAS Core scale. The odds ratio for Criminal Associates is 1.09, which indicates that for every one-unit increase in the Criminal Associates raw score the odds of recidivism increases by 9%. There is solid evidence of criterion validity in this study sample for most of the COMPAS Core scales.

---

[3]In a previous version of the Practitioner's Guide, the biserial correlation was reported. The biserial coefficients are inferred estimates of what the Pearson correlation would be if both variables were continuous and normally distributed. We now use the Pearson product moment correlation (r), which is usually called the point biserial correlation (r_{pb}) when one of the variables is dichotomous.

Table 3.6: Measures of Association Between COMPAS Core Scales and Any Arrest Within Two Years in Farabee et al. Study Data.

COMPAS Scale	Point-Biserial Correlation	AUC	Odds Ratio
General Recidivism Risk	0.34	0.70	3.31
Criminal Involvement	0.20	0.61	1.10
Noncompliance History	0.16	0.61	1.11
Violence History	0.11	0.58	1.06
Current Violence	−0.05	0.52	0.92
Criminal Associates	0.14	0.60	1.09
Substance Abuse	0.02	0.51	1.02
Financial Problems	0.09	0.55	1.08
Voced Problems	0.17	0.61	1.11
Family Crime	0.07	0.54	1.11
Social Environment	0.10	0.56	1.12
Leisure	0.11	0.57	1.07
Residential Instability	0.08	0.55	1.04
Social Adjustment	0.15	0.60	1.10
Socialization Failure	0.18	0.62	1.13
Criminal Opportunity	0.19	0.62	1.10
Social Isolation	0.04	0.52	1.02
Criminal Thinking	0.11	0.57	1.04
Criminal Personality	0.13	0.58	1.05
Cognitive Behavioral	0.23	0.63	1.05

With n=23,635, a correlation of .013 is significant at $p < .05$ (2-tailed).

Table 3.7 shows the point biserial correlations between the COMPAS Core scales and any arrest within 1 year in the study sample from Brennan et al. (2009). The sample consists of 2,328 probation intakes assessed with COMPAS Core. The results in Table 3.7 can be compared to the results in Table 3 in Brennan et al. The sample and event of interest (any arrest) are identical, but here we fit a logistic regression model with a binary outcome (any arrest within one year), and in Brennan et al. we fit a Cox proportional hazards model in which the outcome is defined as failure over the entire follow-up which ranged out to 1,722 days. The base rate for arrest in the binary outcome sample is 0.17, so the following interpretation for r_{pb} adjusted using the formulae in Rice and Harris can be used: $0.08 =$ small; $0.19 =$ medium; $0.29 =$ large.

Table 3.7: Measures of Association Between the COMPAS Core Scales and Any Arrest Within 1 Year in the 2010 New York Probation Study Data.

COMPAS Scale	Point-Biserial Correlation	AUC	Odds Ratio
General Recidivism Risk	0.27	0.71	2.94
Criminal Involvement	0.08	0.56	1.05
Noncompliance History	0.13	0.59	1.15
Violence History	0.08	0.55	1.09
Current Violence	0.04	0.53	1.12
Criminal Associates	0.13	0.60	1.15
Substance Abuse	−0.07	0.55	0.93
Financial Problems	0.05	0.53	1.06
Voced Problems	0.17	0.63	1.12
Family Crime	0.12	0.57	1.22
Social Environment	0.09	0.57	1.18
Leisure	0.14	0.59	1.11
Residential Instability	0.08	0.56	1.06
Social Adjustment	0.17	0.63	1.12
Socialization Failure	0.18	0.64	1.15
Criminal Opportunity	0.22	0.66	1.14
Social Isolation	0.06	0.55	1.03
Criminal Thinking	0.09	0.57	1.04
Criminal Personality	0.15	0.62	1.06

With n=2,328, a correlation of .041 is significant at $p < .05$ (2-tailed).

3.2.2 Construct Validity

Construct validity refers to the extent to which a scale measures what it is supposed to measure. Construct validity is tested by observing correlations between measures of the same or divergent constructs. Construct validity is relevant only for the COMPAS need scales and refers in part to unidimensionality of the scale and to its factor structure. Construct validity additionally is based on establishing evidence that a scale correlates in an expected manner with similar scales, and to other relevant variables in theoretically expected ways. To demonstrate the construct validity of a measure requires the testing of different types of validity including convergent and divergent validity. Here we only address the convergent validity of the COMPAS Core need scales. A direct approach to convergent validity is to measure the correlation between matched scales of the LSI-R and COMPAS Core. The LSI-R is considered a gold standard because it is the current industry leader. This would be a good indication for how well the COMPAS Core scales are measuring the same concept. Results from a study conducted in the California Department of Corrections and Rehabilitation (Farabee et al., 2010) show a direct and high level of correlation between matching LSI-R and COMPAS Core scales. The findings shown in Table 3.8 offer strong evidence of the convergent validity of the COMPAS Core scales. Farabee et al. (2010) found high Pearson product moment correlations between the LSI-R and COMPAS Core measures of Criminal Involvement (0.64); Vocation/Education (0.51); Criminal Associates (0.48); Substance Abuse (0.53); Financial (0.49); and Residential Stability (0.57).

Table 3.8: Correlations between COMPAS Core and LSI-R scales in Farabee et al., 2010

COMPAS	LSI-R	Correlation
Criminal Involvement	Criminal History	0.64 ($p < .0001$)
Criminal Associates/Peers	Companions	0.48 ($p < .0001$)
Substance Abuse	Alcohol/Drug Problem	0.53 ($p < .0001$)
Financial	Financial	0.49 ($p < .0001$)
Vocation/Education	Education/Employment	0.51 ($p < .0001$)
Family Criminality	Family/Marital	0.16 ($p > .10$)
Leisure	Leisure/Recreation	0.05 ($p > .10$)
Residential Instability	Accommodation	0.57 ($p < .0001$)
Criminal Attitudes	Attitudes/Orientation	0.20 ($p = .08$)

Shifting to more general issues of convergent validity, we consider additional evidence to support the convergent validity of the COMPAS Core need scales. For example, the COMPAS Core Substance Abuse Scale correlates positively ($r = 0.44$) with the Substance Abuse Subtle Screening Inventory (SASSI) in the Michigan Department of Corrections pilot data (n=769). We also find the Core Substance Abuse Scale correlates with the Texas Christian University Drug Screen (TCU Drug Screen) (Knight, Simpson, & Morey, 2002) at 0.51 in a sample of offenders assessed with both scales in the Wyoming Department of Corrections (n=4,874). We find similar correlations between the Core Substance Abuse Scale and the TCU Drug Screen in a sample of 2,029 men assessed with both scales in the Massachusetts Department of Correction ($r = 0.54$).

Convergent validity is also demonstrated if a measure correlates in the predicted manner with other variables with which it theoretically should correlate. For example, research in developmental delinquency (longitudinal research in which anti-social behaviors and attitudes are studied over the life course) consistently finds that youth with early onset of delinquent behavior tend to have more serious delinquency trajectories and more negative emotionality, lower achievement, and problems in social adjustment (Moffitt, 1993). Thus, when we consistently find, over multiple studies, that our Criminal Personality, Criminal Attitudes, Social Adjustment and Vocational Educational scales correlate with age-at-first-arrest, just as developmental delinquency research predicts, we take this as evidence of convergent validity. Note that these correlations with age-at-first-arrest hold up when current age is statistically controlled.

Furthermore, age-at-first-arrest is a good external variable to demonstrate convergent validity of the COMPAS Core need scales. Although age-at-first-arrest is collected inside COMPAS, it comes from official records, while the need scales are scored using a different method (interview and self-report).

We have evidence of convergent validity of this type from psychometric studies in the Michigan Department of Corrections (MDOC), New York Office of Probation and Correctional Alternatives (OPCA), New York State Division of Parole, Virginia Department of Corrections, South Carolina Department of Probation, Parole, and Pardon Services, and many other sites. To illustrate our approach to demonstrating convergent validity, we present results in Table 3.9 from a sample in the Wisconsin Department of Corrections, Division of Community Corrections (DCC). The DCC sample consists of 25,773 Core COMPAS assessments conducted between July 1, 2012 and August 31, 2013. Men comprise 76.7% of the sample.

There are many notable correlation patterns in Table 3.9 that provide evidence of convergent validity for the COMPAS Core scales. For example, we see that age-at-first-arrest correlates negatively with the higher-order personality scales Criminal Thinking ($r = -.13$) and Criminal Personality ($r = -.24$). This comports with findings in developmental research that indicate offenders with early onset are more likely to have high scores on similar types of personality measures and more serious and persistent criminal involvement (Moffitt, 1993). Similarly, we see that offenders with earlier age-at-first-arrest are more likely to have higher scores on scales measuring factors that have been identified as criminogenic in longitudinal developmental studies. These scales include Criminal Associates ($r = -.28$), Family Crime ($r = -.22$), Vocational/Educational Problems ($r = -.24$), and Social Environment ($r = -.16$) (Farrington, Jolliffe, Loeber, Stouthamer-Loeber, & Kalb, 2001).

Another pattern in Table 3.9 is defined by the correlations between previous arrests and the scales Social Adjustment ($r = 0.22$), Criminal Personality ($r = 0.15$), Criminal Associates ($r = 0.23$) and Substance Abuse ($r = 0.22$) (Stouthamer-Loeber, Loeber, Wei, Farrington, & Wikstrom, 2002).

There are modest, significant correlations between the assault infractions item from COMPAS Core and the scales Criminal Associates and Peers ($r = .19$), Vocational Educational Problems ($r = 0.15$), Social Environment ($r = 0.13$), Social Adjustment ($r = 0.15$), and Criminal Personality ($r = 0.16$). In their meta-analysis, Gendreau, Goggin, and Law (1997)

Table 3.9: Concurrent correlations between COMPAS Core Scales and criminal history indicators in the Wisconsin Division of Community Corrections sample.

COMPAS Scale	Age-at-First	Prior Arrests	Parole Returns	Prior Prisons	Assault Infractions
Criminal Associates	−0.28	0.23	0.17	0.17	0.19
Substance Abuse	−0.07	0.22	0.14	0.15	0.06
Financial Problems	−0.07	0.08	0.06	0.05	0.04
Voced Problems	−0.24	0.16	0.13	0.11	0.15
Family Crime	−0.22	0.10	0.04	0.02	0.07
Social Environment	−0.16	0.10	0.11	0.13	0.13
Leisure	−0.13	0.04	0.02	0.00	0.05
Residential Instability	−0.09	0.10	0.11	0.09	0.11
Social Adjustment	−0.27	0.22	0.13	0.11	0.15
Social Isolation	−0.02	0.08	0.07	0.06	0.07
Criminal Thinking	−0.13	0.06	0.03	0.02	0.10
Criminal Personality	−0.24	0.15	0.11	0.09	0.16

With n=25,773, a correlation of .013 is significant at $p < .05$ (2-tailed).

found that antisocial attitudes and criminal peers were important individual level predictors of prison misconduct.

There are small, significant correlations between the number of returns to custody for a parole violation and the scales Criminal Associates and Peers ($r = 0.17$), Substance Abuse ($r = 0.14$), Vocational Educational Problems ($r = 0.13$), Residential Instability ($r = 0.11$), and Social Adjustment ($r = 0.13$). Substance abuse, residential stability, and employment and education have been identified in past research as important factors associated with reentry success (Nelson, Deess, & Allen, 1999; Petersilia, 2003; Solomon, Visher, La Vigne, & Osborne, 2006; Travis, 2005). At least one study using self-report and qualitative methods found that housing and employment problems did not distinguish between parole violators and successes (Bucklen & Zajac, 2009).

Overall, the observed relationships between the COMPAS Core scales and criminal history indicators in the Wisconsin DCC sample provide evidence of the convergent validity of the scales. These correlations comport with relationships between risk factors and serious and violent trajectories observed in developmental criminological research (Herrenkohl et al., 2000; Tolan & Gorman-Smith, 1998).The significant correlations we have pointed out are somewhat attenuated by variability in the base rates of the paired variables. These modest associations are typical of correlations between need scales and criminal involvement variables observed in many criminal justice research contexts.

3.2.3 Content Validity

Content validity refers to the coverage of key factors that are relevant in the criminogenic domain. COMPAS Core was designed to have greater coverage of relevant scales than the

LSI-R. Content validity has a major role in any assessment field. It refers to the extent to which an assessment comprehensively includes and assesses the key factors in a domain of interest. The LSI-R includes 10 important criminogenic factors that assess constructs supported in the literature.

A study conducted by Farabee et al. (2010) found that 9 out of these 10 LSI-R scales are clearly matched to a similar scale in COMPAS Core. Thus, in terms criminogenic scale coverage (content validity), COMPAS Core matches virtually all scales contained in the LSI-R. However, the COMPAS Core system additionally includes another 14 scales that can be utilized or turned on/off by an agency depending on its information needs. These additional scales are supported empirically and cover constructs such as anger/hostility, history of non-compliance, low social supports, and socialization failure.

3.3   Internal Consistency Reliability

For a scale to be useful it must be reliable. For example, if one were to carry out repeated testing of a given respondent with different questions or tests, approximately the same scale value should be obtained on each re-test. Generally, if the items entering a scale are highly correlated (internally consistent), then the summated scale will be reliable. Internal consistency reliability - typically assessed by Cronbach's Alpha Coefficient - is a widely used and popular reliability approach. It is often used as a counterpart to test-retest reliability. By convention alphas of 0.70 and above indicate acceptable internal consistency for most applications in the behavioral sciences. Low alphas indicate the scale has too few items or the items don't have much in common and possibly measure more than one construct (Nunnally & Bernstein, 1994).

Table 3.10 shows the summary statistics and alpha coefficients for the COMPAS Core scales in a sample of prison intake assessments in the Michigan Department of Corrections. We have consistently found similar results in prison and probation study samples across numerous jurisdictions.

The low alphas on Violence History (0.53) and Current Violence (0.52) reflect the fact that these are indexes composed of different types of offenses that do not necessarily correlate with each other. A low alpha does not indicate a problem because the items are not expected to be highly correlated as they are in a scale. Family Crime (0.62) is a similar type of index of problems experienced by family members.

Social Adjustment (0.54) and Criminal Opportunity (0.66) are higher order scales. They are not unidimensional. Low internal consistency is less of a concern for these scales. They are composed of two or three underlying constructs each. Cronbach's alpha is less useful for higher order scales, since the multidimensionality of the higher order scales makes it difficult to ascertain what low alpha coefficients indicate. Conversely high alphas do not necessarily indicate unidimensionality.

Table 3.10: Summary statistics and alpha coeffients for the COMPAS Core scales in a prison intake sample from the Michigan Department of Corrections .

	Items	N	Min	Max	Mean	SD	Alpha
Criminal Involvement	4	15,315	0.00	19.00	8.82	4.66	0.75
Noncompliance History	5	15,315	0.00	21.00	4.49	4.23	0.65
Violence History	9	15,315	0.00	16.00	2.13	2.37	0.52
Current Violence	7	15,315	7.00	13.00	8.21	1.28	0.53
Criminal Associates	7	15,315	7.00	22.00	9.75	2.66	0.71
Substance Abuse	10	15,315	10.00	20.00	12.81	2.40	0.76
Financial Problems	5	15,315	5.00	15.00	8.21	2.34	0.70
VocEd Problems	11	15,315	11.00	30.00	19.60	3.89	0.71
Family Crime	6	15,315	6.00	12.00	7.57	1.50	0.62
Social Environment	6	15,315	6.00	12.00	7.54	1.82	0.81
Leisure	5	15,315	5.00	17.00	7.86	3.52	0.86
Residential Instability	10	15,315	9.00	30.00	13.26	3.70	0.71
Social Adjustment	15	15,315	12.00	37.00	20.23	3.44	0.54
Socialization Failure	13	15,315	7.00	32.00	12.10	3.76	0.69
Criminal Opportunity	14	15,315	13.00	39.00	21.23	4.45	0.66
Social Isolation	8	15,315	8.00	40.00	16.90	4.85	0.83
Criminal Thinking	10	15,315	10.00	45.00	20.73	4.91	0.80
Criminal Personality	13	15,315	13.00	58.00	31.84	5.71	0.70

3.4 Test-Retest Reliability

In an independent study by Farabee et al. (2010) the COMPAS Core scales showed very high test-retest reliability, with correlations ranging from 0.70 to 1.00, and with an average correlation above 0.80. Thus, the various COMPAS Core sub-scales demonstrated good to excellent reliability over time. An important aspect of the Farabee study was a comparison against the well-known LSI-R. Overall, the average test-retest correlation coefficient for the COMPAS Core scales was 0.88; for LSI-R, the mean as measured in the same study was 0.64.

Chapter 4

Treatment Implications for Scales

Each COMPAS scale has been constructed based on a variety of behavioral and psychological constructs that are of very high relevance to recidivism and criminal careers. Included in this section is a brief description of the area of research/literature that supports the scale content and context. This material supplements the document "Measurement and Treatment Implications of the COMPAS Core Scales."

Interpretation of the scale scores and how they relate to case planning and intervention is a key concept for COMPAS users. The information contained in this section is intended to assist you in your interpretation of the COMPAS scores as you plan for meaningful interventions and plot the course of behavioral change with the individual. Some brief examples of language for case planning are also offered with each need scale description as a means to generate thoughtful, individualized goals and tasks for a person under supervision. The language (not considered a full treatment plan or goal/task statement) in the case planning examples is action oriented in the goals and tasks. The concept of "how" is defined through behavioral statements. For example, how will the person find emergency housing, or how will the person find new, healthy friends.

4.1   Risk Scales

In this section we describe the Risk Scales in COMPAS. We have developed risks scales for general recidivism, violent recidivism, and pretrial misconduct. There are additional risk scales under development. Northpointe's Research Department also conducts outcomes studies with clients and develops and validates customized risk assessment tools.

4.1.1 Pretrial Release Risk

The Pretrial Release Risk Scale (PRRS) was developed through a pretrial release outcomes study conducted in a large sample of felony defendants assessed with COMPAS in Kent County, Michigan Pretrial Services (Dieterich, 2010). The PRRS was constructed to predict failure to appear (FTA) and new felony arrest among defendants on pretrial release. The development sample included both supervised and unsupervised pretrial releases.

Prior pretrial risk assessment research has consistently identified a set of factors that are predictive of pretrial failure. The most common risk factors include current charges, pending charges, prior arrest history, previous pretrial failure, residential stability, employment status, community ties, and substance abuse (VanNostrand, 2003). We selected items from the COMPAS assessment and included them as candidates for risk model development on the basis of this prior research.

One purpose of pretrial release risk assessment is to sort a pretrial caseload into low-, moderate-, and high-risk groups based on the likelihood of failure to appear in court or commit a new crime pending trial. Use of the risk assessment tool by pretrial services agencies should result in more consistent and equitable decisions regarding release and conditions of release. The use of objective risk assessment tools is recommended by the National Association of Pretrial Services Agencies (2004). The risk assessment tool should be empirically derived and have demonstrated predictive validity in the jurisdiction in which it is deployed. The factors that enter into the risk assessment score should be consistent with applicable state statutes.[1] These and other guiding principles for pretrial risk assessment are outlined in Pretrial Services Legal and Evidence-based Practices (VanNostrand, 2007).

The current PRRS-II is a modified version of the PRRS. The PRRS-II includes eight risk factors (felony top charge, pending case, prior failure to appear, prior arrest on bail, prior jail sentence, drug abuse history, employment status, and length of residence).

4.1.2 General Recidivism

The recidivism risk scale was developed to predict new offenses subsequent to the COMPAS assessment date. The outcome used for the original scale construction was a new misdemeanor or felony offense within two years of the COMPAS administration date.

The scale inputs include criminal involvement (prior arrests and prior sentences to jail, prison, and probation), vocational/educational problems, drug history, age-at-assessment, and age-at-first-arrest. All of these risk factors are well known predictors of recidivism.

Decile scores 1 through 4 (Low Risk Level) may be regarded as low risk since they are clearly lower than "average." Decile Scores 5 through 7 (Medium Risk Level) may be regarded as medium risk since they are in the middle of the distribution and represent cases that are very close to "average" for the total population of the agency. Decile Scores of 8 and above (High Risk Level) may be regarded as high risk since they are in the top third of the distribution.

---

[1]For example in New York a pretrial risk assessment instrument cannot be based on age, gender, or marital status (Division of Probation and Correctional Alternatives, 2007).

It is important to note that the risk scores are generally taken from static information and that current level of needs, for example, substance abuse or other issues, can influence a person's likelihood of acting out or recidivating. In a later discussion, the concept of Low risk/High needs will be covered. General recidivism refers to a broad range of potential acts, therefore, versatility is an element for consideration. The COMPAS Typologies document delineates the typologies that have been discovered through research at Northpointe. One trait that lends itself to recidivism is versatility.

4.1.3 Violent Recidivism

This scale was originally developed in COMPAS Core assessment data on a large sample of probation and presentence investigation (PSI) cases. The scale was subsequently added to COMPAS Reentry. The scale inputs include history of violence, history of non-compliance, vocational/educational problems, the person's age-at-assessment and the person's age-at-first-arrest. The strong association of these factors with future violence has been established in previous research and holds true for people who are considered "non-disordered" (Gendreau et al., 1996). Additionally, meta-analytic results from studies with disordered persons show that a history of violent crime is one of the more potent predictors of violent recidivism (Bonta, Law, & Hanson, 1998).

Decile scores 1 through 4 (Low Risk Level) may be regarded as low risk since they are clearly lower than "average." Decile Scores 5 through 7 (Medium Risk Level) may be regarded as medium risk since they are in the middle of the distribution and represent cases that are very close to "average" for the total population of the agency. Decile Scores of 8 and above (High Risk Level) may be regarded as high risk since they are in the top third of the distribution.

Some offenders, based on their past history of violent acts may score in the high range, yet, show low or medium needs areas. Consideration for the current status of the offender and the support network in place is, as always, recommended, yet in the case of a person who scores high on this scale, special supervision conditions may be deemed necessary.

4.1.4 Recidivism Risk Screen

The Recidivism Risk Screen (RRS) is a brief recidivism risk scale developed to predict a new misdemeanor or felony offense arrest within two years. The RRS consists of five salient risk factors (age, age at first arrest, number of prior arrests, employment status, and prior parole revocations). The RRS is particularly useful to agencies that apply a triage strategy as part of their risk and need assessment protocol to improve efficiency and reduce workload. The RSS is suitable as a prescreen in correctional facilities to select high risk cases for further assessment using a more comprehensive scale set from the Northpointe Suite. The RSS can also be used in community corrections settings to screen candidates for administrative supervision or lower supervision levels. The RSS is not intended as a substitute for the standard risk scales in the Northpointe Suite. The General Recidivism Risk and Violent Recidivism Risk scales measure aspects of risk (both general and violent recidivism) not covered by the

RRS. Used in combination with the Current Violence Scale, the General Recidivism Risk and Violent Recidivism Risk scales provide a complete recidivism risk profile.

4.1.5   On Counter-Intuitive Predictions

Sometimes the COMPAS risk score for a particular person does not match the practitioner's expectations or clinical judgment regarding the level of risk posed by that person. A case in point is when an offender with no prior violence history scores medium or high on the Violent Recidivism Risk Scale. Or, conversely, an offender with some violent history scores low on the Violent Recidivism Risk Scale. This section explains how this occurs and why it is not an indication that the risk scale has failed to work properly.

The COMPAS risk scales are actuarial risk assessment instruments. Actuarial risk assessment is an objective method of estimating the likelihood of reoffending. An individual's level of risk is estimated based on known recidivism rates of offenders with similar characteristics.

The Violent Recidivism Risk Scale is constructed from the following characteristics that we found to be predictive of new person offenses (misdemeanor or felony):

- History of Noncompliance Scale

- Vocational Education Scale

- Current age

- Age-at-first-arrest

- History of Violence Scale

Each item is multiplied by a weight (w). The size of the weight is determined by the strength of the item's relationship to person offense recidivism that we observed in our study data. The weighted items are then added together to calculate the risk score:

Violent Recidivism Risk Score $= (\text{age} * -w) + (\text{age-at-first-arrest} * -w) + (\text{history of violence} * w) + (\text{vocation education} * w) + (\text{history of noncompliance} * w)$

The strong association of each of these inputs with person offense recidivism that we observed in our studies has been established by many other researchers in criminal justice. Meta-analytic results show that violent criminal history, education and vocational problems, current age, and age-at-first-arrest are consistent predictors of violent recidivism. The Violent Recidivism Risk Scale has items in common with many risk assessment instruments in use in corrections, including the Level of Service Inventory-Revised (LSI-R); the General Statistical Information on Recidivism (GSIR); the Violence Risk Appraisal Guide (VRAG)

and the Sex Offender Risk Appraisal Guide (SORAG); and the Self-Appraisal Questionnaire (SAQ).

Your auto insurance company uses a similar risk prediction approach to estimate your risk of having an accident. Besides your age and accident history, the equation includes other characteristics such as credit rating and gender. If you are under 25, male, and have poor credit, you may be classified as high risk even though you have never had an accident.

In the context of Violent Recidivism Risk, if you are young, unemployed and have an early age-at-first-arrest and a history of supervision failure, you could score medium or high on the Violence Risk Scale even though you never had a violent offense arrest.

It is possible for a person's score on the Violent Recidivism Risk Scale to deviate considerably from what one would expect given the person's score on the History of Violence Scale. Consider a hypothetical person who scores high (D10) on History of Violence (2 prior misdemeanor assault offense arrests, 1 prior domestic violence offense arrest, 1 violent property offense arrest, and 1 prior weapons offense arrest); medium (D6) on vocation / education problems, and low on noncompliance history (D1). This person has a late age at onset (age at first arrest = 33 yrs) and he is 51 years old. He has no history of noncompliance (D1) and no vocation or education problems. All of these factors subtract substantially from his Violent Recidivism Risk score, which falls into decile 3 (D3). Note that age is one of the best predictors of violent recidivism, and it carries a lot of weight in the Violent Recidivism Risk Scale calculation. If our hypothetical person were 25 years old and his age at first arrest were 16 years old, his Violent Recidivism Risk score would jump to D8 (High).

Why Is the Current Offense Not Included in the Risk Score?

The Recidivism Risk Scale does not include current violent offense in its calculation. When an offender with a current violent offense obtains a Low Score on the Violent Recidivism Risk Scale, the Low Score may appear counterintuitive. The Violent Recidivism Risk Scale was trained to predict general violent recidivism (misdemeanor or felony person offense). During model development we generally find that violent current offense does not significantly improve the prediction of general violent recidivism. However, an appreciation of the nature and circumstances of the current offense remains essential for effective case management. Current violent offenses are captured by the Current Violence Scale.

What About Domestic Assault or Sex Assault Offenses?

For both domestic assault and sex assault, details about the current offense are important for understanding the risk of recidivism. If the current offense is domestic assault or sexual assault, then it is recommended to use an index-offense-specific risk tool to assess risk of recidivism. COMPAS includes secondary assessments for this purpose, including the Vermont Assessment of Sex Offender Risk-2 (VASOR-2) (McGrath & Hoke, 2001; McGrath, Lasher, Cumming, Langton, & Hoke, 2014) and the STATIC 99 (Hanson, 1997; Hanson & Thornton, 2000) for use with adult male sex assault offenders and the Revised Domestic Violence

Screening Instrument (DVSI-R) for use with adult domestic violence offenders (Williams & Grant, 2006; Stansfield & Williams, 2014).

What Percent of the Assessments will have a Counterintuitive Pattern?

There are two counterintuitive patterns: (1) An offender with no prior violence history scores high on the Violent Recidivism Risk Scale and (2) An offender with high violent history scores low on the Violent Recidivism Risk Scale. The relative frequency of these patterns depends on the relative frequency of violent history in the agency population. If a large percent of the agency population has low violent history then pattern 1 is more likely. If a large percent of the agency population has high violent history then pattern 2 is more likely. The alignment between the agency data and the norm data will affect the proportion classified as high (or low) on the Violent Recidivism Risk Scale, which will also affect the likelihood of counterintuitive scores.

Cases that have a counterintuitive pattern of History of Violence and Violent Recidivism Risk should be examined closely and considered for an override. Persons who exhibit pattern 1 are more likely to have early age at onset and younger age at assessment, and possibly a history of noncompliance and vocational/educational problems. Persons who exhibit pattern 2 are more likely to have late age at onset and older age at assessment, with minimal history of noncompliance and few vocational/educational problems. In all cases a holistic framework to case formulation should be applied that takes into account the varied aspects of the offender as measured by the COMPAS risk and needs scales.

General Comments on Risk Prediction

Risk assessment is about predicting group behavior (identifying groups of higher risk offenders) - it is not about prediction at the individual level. Your risk score is estimated based on known outcomes of groups of offenders who have similar characteristics.

The Violent Recidivism Risk Scale could be constructed in such a way that a high (low) score can only be obtained for someone who has (doesn't have) a history of violent offense arrests. This could be accomplished for example by constructing the Violent Recidivism Risk Scale entirely (or almost entirely) of violent history items. However, based on our own research and that of many other researchers, a scale that depends too heavily on violent history items will not have good predictive power.

Our risk scales are able to identify groups of high-risk offenders - not a particular high-risk individual. We identify groups of offenders who score high, medium or low-risk. We expect that the high-risk group will have higher recidivism rates for violent offenses relative to the low-risk group - this, in fact, has been demonstrated in our outcomes studies.

It is also important to note that we would expect staff to disagree with an actuarial risk assessment (e.g. COMPAS) in about 10% of the cases due to mitigating or aggravating circumstances which the computer is not sensitive to. In those cases staff should be encouraged to use their professional judgment and override the computed risk as appropriate - documenting it in COMPAS with the Override Reason - for monitoring by supervisory staff.

4.2 Criminogenic Need Scales

Need scales measure a criminogenic need and help with case planning. In the following section we briefly describe each COMPAS Core need scale and give examples of the goals and tasks that might be put into a case-plan.

4.2.1 Cognitive Behavioral

This is a higher order scale that incorporates the concepts and items included in the Criminal Associates, Criminal Opportunity, Criminal Thinking, Early Socialization, and Social Adjustment scales.

This scale, as mentioned above, includes grouped items which represent areas of need that can best be addressed in settings that include cognitive restructuring approaches. Concurrent drug/alcohol treatment or other interventions that address immediate needs are recommended, a balanced approach is necessary to avoid overwhelming the person with interventions. For some people, implementing interventions before they are on community supervision is the best approach, as they will have the opportunity to focus on changing their thoughts, feelings and behavior in a controlled setting without the challenges of a community setting. When a person scores in the medium and high ranges of this scale, considerations for their world view must be made, beginning with the question, "does this person see a need for change?".

Table 4.1: Case Planning example for Cognitive Behavior

Goal	Build new and increase healthy coping skills
Task	Immediate Needs: Identify sources/triggers of my anger, frustration, and feelings of being overwhelmed. Make separate lists for each feeling, include what was going on in my immediate surroundings at that moment, who else was there, stressful incidents, and any other information I think is significant.
Task	Ongoing Needs: Use my healthy coping skills (from my skills list/optional actions) to problem-solve in situations where I feel stressed, angry, overwhelmed or when I recognize my triggers to use old behavior to get through a situation.

4.2.2 Criminal Associates/Peers

An involvement with anti-social friends and associates is one of the "big five" risk factors for criminality to emerge in meta-analytic research (Gendreau et al., 1996). Affiliating with aggressive and criminal others is a significant risk factor for further violence and crime. This is consistent with both social learning theory and sub-cultural theories of crime (Andrews, Zinger, et al., 1990; Elliot, Huizinga, & Ageton, 1985).

This scale assesses the degree to which a person associates with other persons who are involved in drugs, criminal offenses or gangs, and determines whether they have a history of arrests and incarceration. A high score would identify persons who are involved in a network of highly delinquent friends and associates.

This domain is considered a strong area of influence for people in the criminal justice system. Interventions in this area can be difficult for the person as their identity with a group as well as a support system, albeit criminally involved, will be altered. Gang influence is particularly difficult as a real level of threat could exist for the person who, by leaving/taking a break from gang life, may be viewed as disrespecting those who have brought him/her to this point in life. Compliance, rather than change is likely for some people, yet, it is a step forward with respect to safety and recidivism.

Table 4.2: Case Planning example for Criminal Associates/Peers

Goal	Increase my association with pro-social, healthy friends
Task	Immediate Needs: Identify traits and behavior of positive, healthy friends and family members
Goal	Reduce interactions with anti-social, potentially harmful friends
Task	Identify friends and family who I tend to get into trouble with, include any co-defendants or criminally involved associates
Task	Create a plan to avoid interaction with criminally oriented friends/family, include statements regarding what my actions will be if I come into contact with the friends/family I have listed as "trouble" for me.

4.2.3   Criminal Involvement

The degree of criminal involvement has consistently emerged as a major risk factor for predicting ongoing criminal behavior. It is the most important of the major risk factors that have emerged in various meta-analysis studies (Gendreau et al., 1996; Andrews & Bonta, 1994). Early juvenile delinquency involvement has also been linked to ongoing criminal behavior (Moffitt, 1993).

This scale is defined by the extent of the person's involvement in the criminal justice system. A high score indicates a person who has had multiple arrests, multiple convictions, and prior incarcerations. The items centrally defining this scale are the number of arrests and number of convictions. A low score identifies the person who is either a first-time arrest or has minimal criminal history. Thus, the central meaning of this scale is the extensiveness of the criminal history.

Arrest history is useful here to see patterns (persons, places, things, time of year) and other related elements that could be antecedents to recidivism and perhaps causal factors (thoughts, feelings, beliefs, attitudes) that can be impacted by intervention. Cognitive behavioral approaches seem to work best in this life area to re-set a person's response to triggers and patterned responses.

Case planning will be similar to criminal associates/peers, criminal personality and criminal opportunity and some cognitive behavioral goals. See the goals listed in Table 4.2, 4.4 and 4.3.

4.2.4   Criminal Opportunity

We have developed a higher order scale to assess the concept of criminal opportunity. This scale emerges from those criminological theories that stress the importance of routine daily activities and the importance of occupying certain social roles (marriage, parenting, being an employee). These roles tend to structure a person's daily activities in a pro-social manner, fostering social bonds and associated local social controls. The theoretical background to this scale includes routine activities theory that emphasizes the importance of immediate local daily activities that place a person in high risk or high opportunity situations (L. E. Cohen & Felson, 1979). The second theoretical theme contributing to this scale is early social control theory (Hirschi, 1969) which emphasizes the importance of social bonds as inhibitors or constraints to crime. The third theoretical strand in which the concept of opportunity is important is the "life cycle" theory of Sampson et.al. (1994). This asserts that age related desistance from crime is linked to life cycle changes that increase both social bonds (wives, children, jobs) and the immediate social controls of associated roles.

This higher order scale assesses criminal opportunity by using items that represent a combination of the following: time in high crime situations, affiliation with high risk persons who often engage in illegal activities, an absence of pro-social or constructive activities (e.g. working, spending time with family, etc.), an absence of social ties, high boredom, high restlessness and being in a high risk age group. The central items include: being unemployed, living in a high crime area, having friends who engage in drug use, and having no constructive activities.

A variety of life areas are represented within this scale. Interventions can be put in place in concurrent waves—for example, seeking out new friends and activities that are pro-social and have positive elements such as learning new skills, helping others, gaining awareness and acting on the awareness at the same time. Structure is a key ingredient in reworking previously idle or non-constructive time. Performance measures as a means of accountability and tracking behavior are also useful tools in this area.

Table 4.3: Case Planning example for Criminal Opportunity

Goal	Increase positive activities
Task	Immediate Needs: Set a date and time for any new activities to help me follow through with the plans I make for new, positive activities.
Task	Ongoing Needs: Develop career aspirations, goals, and identify potential role models as a way to connect with others outside of my family as a means to move forward. Create a plan with each item listed, including dates, for behavioral actions on my part.

4.2.5 Criminal Personality

Several personality dimensions have emerged from recent research as significantly related to persistent criminality. These dimensions involve impulsivity, risk-taking, restlessness and boredom, absence of guilt (callousness), selfishness and narcissism, interpersonal dominance, anger and hostility, and a tendency to exploit others (Hare, 1991; Cooke, Forth, & Hare, 1998). Bonta (1996) reports that criminal personality was the second most important dynamic factor in predicting recidivism. Bandura (1996) also reports validating similar personality dimensions. Criminal personality was one of the "big five" risk factors for criminality in the meta-analysis of Gendreau (1996). The well known General Theory of Crime proposed by Gottfredson and Hirschi (1990) similarly invokes the personality concept of "low self-control," which includes similar dimensions of criminal personality. Prior research has demonstrated a modest but significant relationship between psychopathy, low self-control (variously defined) and both violence and general criminal behavior (Quinsey et al., 1998). Quinsey el al. (1998) include the PCL (Hare, 1991) within their violence risk predictive system – the VRAG.

The items in this scale cover the main dimensions identified as components of the criminal personality (e.g. impulsivity, no guilt, selfishness/narcissism, a tendency to dominate others, risk-taking, and a violent temper or aggression).

Personality is a complex concept and many social scientists believe personality is "set" in childhood/adolescence. Given that many factors come together to create personality, the idea of criminal personality is no less complicated. There are patterns seen in persons who exhibit criminal personality traits. Intervention then, is based on cognitive behavioral approaches that examine and offer alternatives to thoughts, feelings, beliefs and resultant criminal behavior. A specific diagnosis of anti-social personality disorder is not necessary when considering intervention, the area of focus is as listed above – what is the process the person undergoes while deciding to engage in criminal behavior, what is his/her rationale, and what is he/she willing to do about making changes?

Table 4.4: Case Planning example for Criminal Personality

Goal	Build new and increase positive coping and communication skills.
Task	Immediate Needs: Journal my behavior in the areas of thoughts, feelings, attitudes and resultant behavior when I feel stressed, angry, or that something unfair has happened to me. Do my journal entries daily for 5 days and bring to my next probation appointment.

4.2.6 Criminal Thinking Self-Report

Antisocial attitudes and beliefs are identified among the "big five" risk factors in meta-analysis studies of factors that predict crime (Gendreau et al., 1996). However, there is no agreement on the particular attitudinal dimensions or cognitions that are the most useful for predictive purposes. Various studies focus on aspects of thinking style, attitudes toward criminal

justice, neutralization and excuses, tolerance for law violation, cognitive justifications, etc. Clearly, this area could require a highly extensive inventory to map the full range of cognitive dimensions relative to crime. In the absence of such consensus, we adapted the approach of Bandura (1996). Bandura's approach assesses several key cognitive dimensions that justify, excuse, and minimize any damage caused by the person's behavior/crime.

This scale brings together several cognitions that serve to justify, support, or provide rationalizations for the person's criminal behavior. These dimensions include moral justification, refusal to accept responsibility, blaming the victim, and rationalizations (excuses) that minimize the seriousness and consequences of their criminal activity. These include rationalizations such as: drug use is harmless because it doesn't hurt anybody else, criminal behavior can be justified by social pressures, theft is harmless if those stolen from don't notice or don't need what was taken, etc.

The concepts discussed above as they relate to the Criminal Personality scale are also present in this scale, and have been identified in further detail through the person's own self-report. A distinct pattern of rationalizations for criminal and/or harmful behavior is present for those who score in the probable and highly probable categories. Interventions that focus primarily on cognitive behavioral approaches tend work best with those who evidence significant criminal thinking.

Table 4.5: Case Planning example for Criminal Thinking Self-Report

Goal	Modify criminal thinking, develop a positive attitude toward various life areas (see specific goals).
Task	Immediate Needs: Create a list of what works for me (positive thoughts and activities) and what doesn't (negative thoughts and activities) that keep me in the same cycle of getting into trouble.

4.2.7 Current Violence

This scale forms part of the general criminal history and measures the degree of violence in the present offense. The central item that defines the scale is the presence of an assaultive felony. Other key items involve whether or not a weapon was used, if there was injury to a person, etc.

Research has shown that the level of violence in the instant offense is not a good predictor of future crime. Keeping in mind the degree and type of violence in the instant offense as compared to the person's history of violence and current level of functioning/needs scores is good practice. One area for clear consideration is that of family violence and how this will affect any kind of living arrangement for community-based supervision.

Table 4.6: Case Planning example for Current Violence

Goal	Build new and increase healthy coping skills.
Task	Enroll in and successfully complete an Anger Management Program. Bring my homework and updates from the class to each probation appointment.
Task	Create a list of my healthy and unhealthy coping skills, then list next to each one the usual outcome when I use that reaction or response.

4.2.8 Family Criminality

From a social learning theory perspective, participation in criminal behavior may be facilitated by significant others who model such behavior. Research has consistently demonstrated that delinquency and adult crime are both associated with parent criminality (West, 1973; Lykken, 1995). Children may learn that violent and deviant behavior "work" in the context of their family. Aside from the social learning and role modeling perspective, other intergenerational mechanisms may operate to transmit values and behaviors from parent to child. Genetic influences, for example, may operate to transmit anti-social personality disorder and criminality (Lykken, 1995). COMPAS therefore includes a measure of family criminality focusing on the criminality and drug use history of the mother, father, and siblings.

This scale assesses the degree to which the person's family members (mother, father, and siblings) have been involved in criminal activity, drugs, or alcohol abuse. The items cover: arrests of each family member, whether they have been in jail or prison, and whether the parent or parental figure has a history of alcohol or drug problems.

Families can be significant positive resources for any person in the criminal justice system. The presence of family criminality, however, can create a dichotomous situation in that, on the one hand the family is a source of support, comfort, and hope, and on the other hand, they may also be criminally involved and their support revolves around their criminal activity and belief systems.

Table 4.7: Case Planning example for Family Criminality

Goal	Eliminate criminal involvement with family members.
Task	Immediate Needs: List/identify family members (those who I have a relationship with and spend time with) who are criminally involved.
Task	Ongoing Needs: Create a time line of my involvement with these family members and the consequences/benefits of spending time with them, e.g. when did it happen and what happened while we were together.

4.2.9 Financial Problems

This concept appears as one of the more modest risk factors in the Gendreau et al. (1996) meta-analysis. It is linked to lower social class, poor housing, community disorganization, and other factors. Homicides, for example, are disproportionately found in high poverty areas. Numerous social dimensions related to poverty are linked to high crime, including residential mobility, family disruption, single parent families, crowded housing conditions, and higher opportunity for violence (Sampson & Lauritsen, 1994). The measure of poverty and financial problems in COMPAS focuses on the struggle to survive financially, problems paying bills and other issues related to a shortage of money.

This scale assesses the degree to which a person experiences poverty and financial problems. It assesses whether the person worries about financial survival, has trouble paying bills, and has conflicts with friends or family over money.

Unpredictable economic times may play a role in this area, however, a person's pattern of earning (or not) and spending money is an important element. Education on money management and fulfilling court ordered financial commitments is part of the necessary approach when considering interventions. Assuming someone knows how to manage their finances is an erroneous starting place, vocational training may also play a role in creating a successful change plan.

Table 4.8: Case Planning example for Financial Problem Scale

Goal	Gain financial stability/independence.
Task	Immediate Needs: Apply for financial assistance/emergency shelter and/or food stamps (use other resources as referred by PO).
Task	Immediate Needs: Inform my supervisor at work about my probation appointments and any terms and conditions that might impact my ability to do my job.

4.2.10 History of Non-Compliance

This scale focuses on the number of times a person has failed when he or she has been supervised in the community (probation or parole). The central defining items are the number of times that probation or parole has been violated or revoked. Related items include the number of times a new charge or arrest has occurred while the person was on probation and the number of returns to custody for parole violations.

This scale focuses on the number of times the person has failed when he or she has been placed on a community-based status. The central defining item is the number of times probation or parole has been suspended or revoked. Related items include the number of times the person has failed to appear for a court hearing, the number of times a new charge/arrest or technical rules violation has occurred while on probation, parole and prior community corrections program placement failures (i.e. electronic monitoring, community service work, day reporting, etc.) Thus, the scale involves the risk of technical rules violation failure leading to revocation of probation, pretrial release, or community corrections placement status.

Different states/agencies have different thresholds for supervision violation and suspension/revocation. While policy decisions do effect the person's history "on paper" it is also important to understand the person's willingness and ability to successfully complete community-based supervision. Clearly articulated expectations with terms and conditions of supervision and case planning are key factors in laying the groundwork for success. Behaviorally stated goals and a high degree of structure with room for individual differences and learning curves could enhance a person's success rate.

Table 4.9: Case Planning example for History of Non-Compliance

Goal	Attend all probation meetings as scheduled.
Task	Immediate Needs: Client and PO agree upon appointments for two week intervals including attendance at Cog/Behavioral group 1x week. Client to use pocket calendar for personal reminder of all appointments, during this two week period (March 10-24, 2010) client is to attend 2 scheduled appointments at this office (2/12 and 2/19 at 3pm) and the cog group on 2/15 and 2/22 at 6pm.
Note	All case planning activities should include tangible sanctions should the person fail to comply or engage in change behavior, and in the cases when a very high degree of structure is put in place, those sanctions may be stated on the case plan.

4.2.11 History of Violence

A history of violent behavior has been demonstrated to be one of the most powerful predictors of future violence (Farrington, 1991; Parker & Asher, 1987). The likelihood of future violence appears to steadily increase with each instance of a prior violent incident. Each prior arrest for violent behavior increases the likelihood of further violence. Similarly, a history of juvenile violence has been found to be a predictor of adult violence (Farrington, 1991).

The aim of this scale is to reflect the seriousness and extent of violence in an individual's criminal history. It focuses on the frequency with which violent felony offenses have occurred, the use of weapons, and the frequency of injuries to victims. The frequency of several specific violent offenses are also included in the scale (e.g., robbery, homicide, and assaultive offenses).

Multiple episodes of violence may suggest the need for further psychological evaluation. The accumulation of multiples (events, victims, types of crimes against persons/animals) creates a pattern of serious concern. Interventions may be targeted at cognitive behavioral constructs to manage behavior, and highly structured supervision may be preferred by the supervising agency.

While we are not going to change the past, we can teach people to intervene in old thought processes and put in place, new, healthier thoughts that lead to pro-social responses rather than reactions that always follow the same patterns.

Table 4.10: Case Planning example for History of Violence

Goal	Increase my healthy responses to events that trigger an angry reaction for me.
Task	Immediate Needs: List the way I have shown my thoughts and feelings in the past.
Task	Immediate Needs: Describe what happens when I lose self-control.
Task	Immediate Needs: Describe what happens when I use positive, self-control responses.

4.2.12 Leisure/Boredom

Aimlessness in the use of leisure time is linked to several theories of crime. For example, it is a component of Hirschi's early Social Control theory representing an aspect of weak external social bonding (Hirschi, 1969). Aimless use of leisure time is also included as a risk factor in the LSI (Andrews & Bonta, 1994). The General Theory of Crime (M. R. Gottfredson & Hirschi, 1990) includes aimlessness and the related concept of proneness to boredom within the dimension of low self-control or criminal personality. It is also linked to routine activities theory by the maxim of "Idle hands are the devil's workshop" (Osgood, Wilson, O'Malley, Bachman, & Johnston, 1996).

This scale assesses the degree to which the person experiences feelings of boredom, restlessness, or an inability to maintain interest in a single activity for any length of time. Thus, this scale may be regarded as reflecting a psychological dimension rather than representing the amount of constructive opportunities in the person's community environment.

As noted above, the issue is not necessarily time management, but the person's value of experiences and relationships. Creating an understanding of these elements may be a first step toward making changes for the individual. Some social or information processing issues may be identified through further assessment, and these issues can then be addressed accordingly.

Table 4.11: Case Planning example for Leisure/Boredom

Goal	Learn about the relationship between my level of participation with other people/events/interests and my ability to be involved in things outside of work or other required activities.
Task	Immediate Needs: Create a plan for getting involved with my friends who participate in the basketball league at the rec center. List the night and time of the league and the person who I can talk to get on a team. Ask my friend to go with me if I feel like I need support in joining the team.

4.2.13 Residential Instability

An unstable lifestyle is one aspect of the second factor of Hare's Psychopathy Checklist and this is an obvious risk factor for crime and violence (Hare, 1991). Additionally, low social ties and an unstable residential address are often used in pre-trial risk assessment instruments to predict risk of flight. The absence of social ties, and the presence of social isolation are also seen in Social Control theory as the absence of restraints on deviant behavior that result from weak social bonding. In addition, since change and stress are correlated, an unstable lifestyle may be stressful. Finally, personal stress/distress appears as a risk factor with modest predictive validity in meta-analysis studies (Gendreau et al., 1996).

The items in this scale measure the degree to which the individual has long term ties to the community. A low score on this scale indicates a person who has a stable and verifiable address, local telephone and long term local ties. A high score would indicate a person who has no regular living situation, has lived at the present address for a short time, is isolated from family, has no telephone, and frequently changes residences.

Community-based supervision requires a verifiable address. The reality is that some individuals end up in shelters right after release, or, they don't have the financial means to secure acceptable living quarters for months after sentencing/release. The historical nature of the person's residential stability is good information while the person is incarcerated in that planning can be put into place to avoid the pitfalls aforementioned. Renewing and/or creating family contacts and other potential support resources can be used as realistic goals in establishing residential stability.

Table 4.12: Case Planning example for Residential Instability

Goal	Seek and obtain sustainable living situation.
Task	Ongoing Needs: Develop a workable budget that includes housing costs that I did not list under my immediate needs such as pets, additional furnishings, any agreements that I can lawfully enter into to help reduce the cost of my rent.

4.2.14 Social Adjustment

Interpersonal problems may exist in each main social institution (family, school, work, etc.) A pattern of interpersonal problems may indicate poor social skills. The present higher order scale was constructed to assess the recurrence of interpersonal problems across various social contexts. Social skills training is often advocated as a treatment approach in preventing further violence and crime. Social adjustment problems are also implicated in several theoretical perspectives of criminal behavior (e.g., weak social bonding in social control theory (Hirschi, 1969), stress (Gendreau et al., 1996) social cognitive models of crime (Dodge, Pettit, McClaskey, & Brown, 1986; Dodge, 1998) and the erosion of social capital (Hagan, 1998)).

This scale is higher order in the sense that it uses items from other scales that crosscut several domains. It aims to capture the degree to which a person is unsuccessful and conflicted in his/her social adjustment in several of the main social institutions (school, work, family, marriage, relationships, financial.) A high score indicates a person who has been fired from jobs, had conflict at school, failed at school or work, has conflict with family, exhibits family violence, cannot pay bills, has conflicts over money, etc. Thus, the common theme is problematic social relationships across several key social institutions.

Areas for intervention will depend on the most pressing issue and need for support in that area. Creating a sense of connectedness and responsibility for self and to others is a foundational element of many cognitive behavioral approaches. Structuring communication expectations and methodologies for the individual may be a starting place, many programs provide sequenced awareness and practice options. The supervision professional may work with the individual in identifying other community-based pro-social activities, as well.

Table 4.13: Case Planning example for Social Adjustment

Goal	Increase positive social supports with family, friends, and community.
Task	Immediate Needs: Create a plan for increasing my time spent with positive, pro-social friends and family members.

4.2.15   Social Environment

Living in a high crime neighborhood is a well-established correlate of both delinquency and adult crime (Thornberry, Huizinga, & Loeber, 1995; Sampson & Lauritsen, 1994). This risk factor fits into several theoretical models of crime and delinquency (e.g., social disorganization, social learning, and sub-cultural theories). Disorganized and high crime communities are characterized by perceived high crime rates, gangs, easy access to drugs, and inadequate housing.

This scale focuses on the amount of crime, disorder, and victimization potential in the neighborhood in which a person lives. High crime is indicated by the presence of gangs, ease of obtaining drugs, the likelihood of being victimized, a belief that a weapon is needed for protection, and so on.

Few scales reflect areas where the person has no direct control over the identified issues, however, this scale is based on environmental factors that the individual has to cope with on a daily basis. Problem-solving around the possibility of relocating or finding a safer living arrangement may be paramount. Other risk factors come into play when considering the person's social environment (criminal opportunity, criminal peers, family criminality, residential instability, etc.) and these factors may become more of a primary focus should they be identified as active in the person's life.

The Social Environment and Social Isolation scales will typically use case planning language similarly. Increasing positive family and peer relationships, as we have seen in other scales is a primary focus, as well as involvement in specific activities.

4.2.16 Social Isolation

Positive social supports appear to serve several functions that may reduce crime and violence. Social support may act as a protective factor or mediator of stress, since stress and anxiety may predispose a person towards anger and violence. Positive social support has been shown in research to act as a protective factor against risk of violence even in high risk environments (Estroff & Zimmer, 1994). As described below, the COMPAS social isolation scale is bipolar in that it serves to identify social isolation/loner behavior on one pole and strong social supports at the other pole.

This scale assesses the degree to which the person has a supportive social network and is both accepted and well integrated into this network. The scale is scored such that a high score represents an absence of support, and the presence of feelings of social isolation and loneliness. The defining items include: feeling close to friends, feeling left out of things, the presence of companionship, having a close best friend, feeling lonely, etc.

As mentioned in other social support areas, intervention can be across many dimensions and impact the person on both the awareness and practice levels. Strategies might include finding a mentor, joining known pro-social or support groups, learning new skills/hobbies, and creating new social connections where the person's new, healthy behavior will be expected by those involved in the activities.

4.2.17 Socialization Failure

Socialization failure during childhood and adolescence has been consistently linked to crime and delinquency. Problems in the family and inadequate parenting are the critical background issues (Lykken, 1995). We have constructed a higher order factor in COMPAS that builds on the early onset of delinquency, problem behavior in school (dropout, suspensions, fighting, etc.), inadequate parental socialization, and early drug use. These are all well known risk factors for later criminality (Chaiken, Chaiken, & Rhodes, 1994; Lykken, 1995) and all represent early socialization problems. Lykken (1995) in particular, explores the link between socialization failure and criminal behavior in his concept of the sociopath.

This scale combines items reflecting family problems, early school problems, and early delinquency, all of which suggest socialization failure (how the person was socialized growing up). The intent is to examine socialization breakdown through its early indicators in school, delinquency, and family problems. A high score would represent a person whose parents were jailed or convicted or had alcohol or drug problems. In addition, a high score is associated with early behavior problems in school (being expelled, failing grades, skipping classes, fighting) and would also manifest serious delinquency problems.

This scale looks at the history or pathway that was involved in the person's upbringing that may have significantly affected his/her view of the world in terms of trust, respect for reasonable authority, value of others, and the development of beliefs and attitudes that are active and present today. High scoring individuals may need cognitive restructuring programs to assist in an awareness of, and change plan for, some of the beliefs and attitudes that lead to troublesome behavior for the person.

Table 4.14: Case Planning example for Early Socialization Failure

Goal	Build new and increase my positive coping skills and responses.
Task Activity	Attend and successfully complete cognitive behavioral program. Complete first exercise in workbook by 3-20-10 and bring the completed exercise to the next probation appointment. Participate in the cog group by engaging the exercise on My Thoughts, and participating in the role play discussion.
Note	In the case of a structured, sequenced program, case planning will often be stated as in the example above.

4.2.18   Substance Abuse

Numerous published research studies have established that substance abuse is a significant risk factor for both general criminal behavior and violent behavior. Substance abuse emerged as one of the major risk factors in the meta-analysis studies of Gendreau et al. (1996).

The present scale is a general indicator of substance abuse problems. A high score suggests a person has drug or alcohol problems and may need substance abuse intervention. The items in this scale cover prior treatment for alcohol or drug problems, drunk driving arrests, blaming drugs or alcohol for present problems, drug use as a juvenile, and so on.

The cut points on this scale are lower than the other needs scales due to the design of the scale. A person who scores in the Probable range (3-4) is considered a person who is in need of further evaluation (i.e. ASI, SASSI, etc.) and a person who scores in the Highly Probable range (5-10) may have a serious alcohol or drug problem requiring a structured treatment approach. Because of the high incidence of drug/alcohol abuse within the criminal justice population, a primary intervention for many individuals to impact recidivism is assisting the person to attain and maintain sobriety.

Substance abuse typically intersects every life area for a person. Therefore, cognitive behavioral restructuring and life skills planning may be needed following, or, in some cases during, treatment. Case planning language varies in this area between the example shown under the Socialization Failure (Table 4.14) scale regarding structured, sequenced steps, and, the use of supervision focused goals and tasks as listed in Table 4.15.

Table 4.15: Case Planning example for Substance Abuse

Goal	Maintain Sobriety
Task	Attend AA meetings 3 times per week and show my attendance card to my PO at each meeting.
Task	Call in for UA/BAC testing daily and report by 5pm on the day I am to do my testing.

4.2.19   Vocation/Education

Another of the "big five" risk factors for crime and recidivism prediction in the Gendreau et al. (1996) meta-analysis is labeled "social achievement." This concept is an amalgam of educational attainment, vocational skills, job opportunities, a record of stable employment, good income, and, more generally, the level of legitimate economic opportunity. Basically, persons with more social capital have higher "life chances" than other persons who may have very restricted opportunities for success (Hagan, 1998; Coleman, 1990). The family is of critical importance in building social capital. Parents either transmit positive and substantial social capital to their child or fail in the socialization process. This scale is a higher order factor in COMPAS, using items from both educational and vocational domains. Individuals differ greatly in access to social capital or other resources. Social capital is somewhat dynamic. It can be built or destroyed. For example, a record of serious criminal behavior or high school dropout will clearly diminish life chances and social resources, whereas completing a job skills training course or obtaining a GED may increase these chances.

This higher order scale assesses the degree of success or failure in the areas of work and education. A high score represents a lack of resources. Those who score high will present a combination of failure to complete high school, suspension or expulsion from school, poor grades, no job skills, no current job, poor employment history, access only to minimum wage jobs, etc. Thus, the scale represents a lack of educational and/or vocational resources.

A score in the Probable range is significant in that a person may be struggling to seek and maintain employment that meets his/her skill set, ability, and interests. Vocational stability plays a significant role in success on community supervision. Intervention can therefore be initiated during incarceration or upon release. Education, or additional training may be the reasonable answer to assisting the person to maintain employment. Therefore, it is important to look at the whole picture in this domain when assessing paths and barriers to success.

Table 4.16: Case Planning example for Vocation/Eduation

Goal	Develop vocational skills
Task	Immediate Needs: Ask myself what it will take to meet the goals I am setting, identify barriers that come from others/situations, and those that I have put in place.
Task	Immediate Needs: Identify methods to break down the barriers that I have put in place, use my resources (supervisor, PO, instructor) to move forward with my plan.
Task	Enroll in vocational training program using the funding source I found when I contacted the instructor at the school.

4.2.20   The Lie Scale and Random Responding Test

These validity tests provide alerts that the person being assessed by COMPAS is possibly "faking good" or is responding randomly.

Items in the Lie Scale include questions about feeling unhappy or angry with the options across a Likert scale that include "never." Since it is highly unlikely that a person has never felt unhappy or angry, the selection of "never" would suggest they are not telling the truth, or perhaps they are being careless with their responses. If several of the items on the Lie scale are given extreme answers, the criminal justice professional is then alerted to the possibility that the person is not responding truthfully.

The Random Responding scale is based on 37 highly correlated pairs of COMPAS scale items. Some items appear more than once in the pairs as they relate to more than one construct. Random responding has the effect of breaking these correlations. The cutting score was internally set up to detect the 5% of the respondents at the extreme end of the distribution who might be answering the questionnaire in a random fashion.

Chapter 5

Typology

The fact that people respond differently to different treatments has been labeled as responsivity and repeats the conventional wisdom that "one man's meat is another man's poison." It indicates that the wrong treatment may make things worse and creates a need for careful matching of people to specific treatments. This is central to both "What Works" and to the Risk-Need-Responsivity (RNR) model. It also underlies Evidence-Based Practice (EBP), since incorrect matching of a person to treatment may sabotage the effectiveness of virtually any intervention. Thus, a challenge for treatment providers is to match intake assessments to service plans in order to achieve good outcomes. Andrews et al. (2006) recently acknowledge that specific responsivity or differential matching is the least explored of all the RNR principles. The traditional strategy for "'matching" has been to develop treatment-relevant classifications to guide differential matching (Warren, 1971; Megargee & Bohn, 1979; S. Baird, Heinz, & Bemus, 1979). Most of these classification efforts failed because of a variety of technical problems (Harris & Jones, 1999).

However, we have developed risk and need typologies to facilitate the goals of specific responsivity and to guide the matching of interventions to client needs in the context of the COMPAS system. We have developed treatment-relevant typologies for both males and females. These are now included as a standard component of the COMPAS software. These typologies use advanced pattern recognition, cross-validation procedures and multiple methods to verify the stability of the typologies. Each person is now automatically classified on the basis of "best fit" to one of several standard and replicated needs profiles. The class profile of each person is automatically produced as part of the standard report to help treatment staff conceptualize the "kind" of client they are dealing with, and to develop a service plan to meet the specific responsivity needs of that unique individual. It is important to realize that no person is a perfect match to his/her class and will be unique in his/her overall pattern of risks and needs. However, his/her assigned prototype membership will suggest a beginning "framework" for a case plan that may then be customized according to the unique risk and need patterns of each person. Thus, the default treatment plans for each prototype will provide treatment staff a useful initial guide to the most likely kind of service plan for each individual.

The scales required to determine a type in the COMPAS Core typology are: Criminal As-

sociates, Substance Abuse, Financial Problems, Vocation/Education, Family Criminality, Social Environment, Leisure/Boredom, Residential Instability, Social Isolation, Criminal Attitudes, Criminal Personality, and Age at Assessment.

5.1   Interpretation

Questions may arise as to how to interpret the COMPAS typology assignment and how to integrate it into the case plan. Overall, we suggest that the typology results should be interpreted in the context of the other three key classification elements that are provided in the overall COMPAS Risk Assessment. These are as follows:

1. Risk Potential Scales (Predictive levels): These two (red) scales represent overall risk potential scales. They include separate risk scales for Violent Recidivism and General Recidivism.

2. Risk and Need Profiles (Prior history): Next, the profile chart provides the person's decile scores on all background scales (e.g., criminal history, drugs, peers, family, work/education, etc.). These provide the basic data elements that drive risk predictions, needs assessment and treatment plans.

3. Explanatory Typology: This provides the closest fit of each person to one of eight prototypical categories. The eight types represent different kinds of people. It is important to remember that the profile chart of any individual person will never be an exact match to his closest prototype. Many people are hybrids that may not fit well into any typology.

These three elements may be used collectively to guide case formulation and to understand what is "going on" with a case, and to select supervision levels and treatment interventions.

Other important elements that may influence case formulation are as follows:

Recommended Level of Supervision: The recommended level of supervision is found in the Assessment Summary section. The Violence and Recidivism risk potential factors are the main drivers of this recommendation.

Overrides of the supervision level: Overrides of the calculated recommended supervision level are clearly appropriate when it is felt that the automated procedure is either over- or under-estimating the risk level. This is especially true when the screener can identify the presence of mitigating or aggravating factors. Examples of mitigating factors are such things as your own street knowledge of the person, age and any extended periods of crime free behavior, etc. Aggravating factors are such things as severity of offense, gang membership, your knowledge of their street behavior, of non-apprehended crimes, or concerns on the Lie Scale or Random Response Score (as applicable).

Common Prototypes versus Anomalous cases: There are several things to understand about the typology label:

1. The typologies represent "Common Types" of people: We have found that there are eight common categories or prototypical offending and behavior patterns that often re-appear in criminal justice populations. These eight prototypes are described in the software and the software assigns each client to their nearest prototype. However, please remember that no individual is ever an exact match to his/her typology. In most cases there will be a good match to the closest fitting category, but will always have some differences to the ideal prototype. However, some cases will NOT be a good match to any prototype, or may straddle the boundaries between two prototypes. These boundary or hybrid cases are not given a prototype assignment and must be interpreted as unique cases.

2. What to do with the poor fitting/boundary cases: With boundary or hybrid cases, the typology should be ignored, or used as a starting point for a more individualized interpretation. Such boundary types are often harder to interpret and are more complex. If the screener's judgment clearly disagrees with the computer-assigned prototype then an override is appropriate. The anomaly should be reported and the counselor will interpret the case using the individual's case chart and other relevant information to determine processing and treatment plans.

3. Typology Purposes are explanatory and for treatment planning: A main purpose of the typology is to give an alert if a case belongs to one of the major case types (e.g., a young streetwise gang member; an older repeat drinking driver, etc). If a case is a good fit this may help in understanding the case and it's treatment needs since such kinds of cases will have been seen before.

4. The Typology is not a risk classification! The typology emphasizes explanatory and need profiles and treatment: The typology prototypes represent diverse profiles of need factors, and are not designed as a predictive risk classification. Thus, the typology alone should not be used to determine risk levels but it may often help in risk and placement decisions if used in conjunction with the risk scales.

5.2 Male Typology

5.2.1 Type Descriptions

Category 1 - Chronic drug abusers – most non-violent

The central theme of this prototype is long-term substance abuse and non-violent offences. For example, serious substance abuse and use of alcohol/drugs at the current arrest. Problems often begin in adolescence, for example, with first arrests around age 16 or 17. Factors underlying this type may include mixtures of family criminality, family disorganization, out-of-home placements and some juvenile socialization problems. The profile appears in all ethnic groups, but especially young Anglos. The social context does not suggest total social exclusion. For example, some members have relatively few social risk factors and some strengths such as low poverty, educational-vocational resources, stable residence in good neighborhoods and are not isolated, bored or socially rejected. Anti-social personality and extreme criminal attitudes are mostly absent.

Official criminal histories support this profile. This type averages of 3 to 4 prior arrests mostly for drug use or trafficking. This category is mostly non-violent with relatively low current violence, low weapon offences and low victim injuries – although in some cases the current charge includes assault. There is little evidence of domestic violence and sex offences.

Category 2 - Low risk situational – fighting/domestic violence caution

This type has several economic and educational "strengths" suggesting an apparently normal citizen. They mostly avoid criminal associates and follow a low risk lifestyle. However, some members of this group are involved in serious violence, thus caution is warranted. These persons generally are not raised in high crime families, avoid drugs, have stable addresses in safe areas and few financial problems. Personality and criminal attitudes appear average. The profile offers no clear social or criminogenic explanation for offending or for violence. This pattern may reflect the well known accidental or situational event that unexpectedly occurs to create serious violence and an arrest situation.

The official criminal history reflects a low risk profile. The group, as a whole, has fewer official arrests, convictions or prior violence than other types. The official data shows lower violence history, lower weapons use, lower non-compliance, fewer probation episodes and almost no burglaries, robberies, The current offense often is for DUI, substance abuse or an assault (fight/no weapons). Many are incarcerated for the first time. However, as noted above, some members of this group have been charged with a serious assault and/or domestic violence. This category occurs in all ethnic or racial groups – a variant is found in Category 8.

Category 3 - Chronic alcohol problems – DUI, domestic violence

The dominant pattern of this category consists of older (40+), mostly relatively well-educated men who function fairly well with stable jobs, finances and residences, but with recurrent

alcohol problems and a history of DUI and/or domestic violence. They show the oldest age at first arrest (27) and are thus late starters. A generally low risk lifestyle is reflected by few criminal peers, educational-vocational and financial success, low crime families, stable and safe addresses and pro-social structured leisure. They mostly avoid high-risk situations and do not appear to hold anti-social attitudes or personalities. Thus, the explanation for their offending would appear to relate to alcohol proneness perhaps in a context of family stress, rather than social exclusion or environmental explanations.

The official data corroborates this pattern showing that this group has the highest score for current DUI arrest and using alcohol (but not drugs) at the current arrest. Overall, they have average criminal involvement and few violent offenses. However, domestic violence also occurs for some of these people. DUI and alcohol abuse are the major problems since the category has lower clusters arrest rates than other clusters for current violence, weapon arrests, assaults, juvenile felony arrests, fraud, property, burglary and robbery offenses. COMPAS risk scales assign this category to low risk, although this is influenced by their older age (since age lowers risk scores in the risk equations). Thus, they may be expected to have a moderate recidivism risk mainly for drug/alcohol related offenses or domestic violence.

Category 4 - Socially marginalized – poor, uneducated, stressed, habitual offenders

The central problem in this type is socio-economic marginalization (e.g., educational-vocational failure, poor job skills, poverty, unstable residence, poor social supports and social isolation). This category is older (average age 37) and occurs in all ethnic groups. The social resources for these men appear reasonable since they mostly do not have high crime families or antisocial peers, do not reside in high crime areas and do not hold extreme criminal attitudes – all of which argue against a social learning explanation and do not suggest a high-risk lifestyle. There is also little evidence of criminal personality.

Many of these cases are chronic repeaters with multiple arrests, probation terms and convictions. Their official criminal history coheres with the above profile in two main ways. First, they are mostly late starters with a late age at first arrest (21), few juvenile felonies and a relative absence of juvenile socialization problems. Second, their offense pattern of fraud larceny (and some drug trafficking) and low robbery, suggests instrumental crime for financial gain, or perhaps coping with poverty and unemployment. Finally, some of these men exhibit prior domestic violence that coheres with prior weapons use and victim injury. Substance abuse and criminal opportunity scores are about average.

Note: Mental health (MH) problems are often linked to social isolation and social adjustment problems. Thus, cases with MH and social withdrawal problems may enter this lonely marginalized category. A mental health assessment is recommended to clarify MH issues.

Category 5 - Criminally versatile – young marginalized persons often gang affiliated

This pattern exhibits multiple risk factors and several co-occurring causal processes linked to criminality. First, is extreme social exclusion/marginalization (e.g., educational-vocational failure, joblessness and poverty). Second is a lack of social control bonds, withdrawal from education and work, boredom and little constructive use of leisure. Third, their high-risk criminal opportunity lifestyle is reflected in weak pro-social bonds, boredom and higher than average gang affiliation. Fourth, social learning is suggested by a pattern of anti-social attitudes, gang membership (for some), early school failure and out-of-home placements, all implying affiliation with other rejected and weakly socialized peers. Finally, many of these cases reflect an anti-social personality that has been empirically linked to family disintegration, family crime, juvenile felonies and early onset shown by many of these cases. These themes reflect the sociopathic type of described by Lykken (1995) and Mealey (1995), and others.

The criminal history of this category coheres with the above high risk profile. This young group (22-23 average age) generally has an early age at first arrest (around 16), higher scores than other types for juvenile felonies, weapons arrests, current violence, current property and sex offense charges. However, there are two anomalies. First, they show relatively low substance abuse. Second they score only average for prior arrests and convictions, perhaps resulting from their youth (i.e., their early stage of a criminal career).

Category 6 - Socially isolated long term substance abuse – multiple minor and mostly non-violent offenses

This group reflects four major criminogenic problems. First, many members exhibit serious long-term substance abuse, suggesting addiction. Second, their extreme marginalization is shown by social isolation, poverty, unstable residence, poor social adjustment, boredom and a lack of pro-social leisure activities. Third, they appear embedded in a criminal drug culture and exhibit high criminal opportunity. Finally, a disposition for criminality is shown by high crime personality and antisocial attitudes. This type occurs in all ethnic groups.

The official criminal history matches this profile in several ways. Chronic criminality is shown by multiple arrests, convictions and probations. Chronic substance abuse is confirmed by alcohol and drug offenses, using hard drugs (heroin, cocaine) as juveniles, being high/intoxicated at current arrest and (in some cases) current drunk driving and/or drug possession charges (but rarely trafficking). This category is difficult to treat as shown by non-compliance, probation/parole revocations and FTA's. They also exhibit above average scores for current fraud, prior domestic violence and burglary/larceny (but, rarely robbery). Criminal violence (except for domestic violence) is rare as shown by relatively low arrests/convictions for weapons offenses and lower scores for assaultive felonies.

Category 7 - Serious versatile high risk individuals

This type has the most serious and violent profile. It may warrant referral for a test such as the Psychopathy Check List (PCL). This profile reflects a chronic, violent and versatile criminal career as well as multiple criminogenic risk factors.

This profile reflects four major causal processes linked to high criminality. 1) A strong personal disposition to crime is shown by anti-social personality, antisocial attitudes/thinking, early onset of crime, parental criminality and versatile criminal offences. 2) Social marginalization is shown by educational/vocational failure, unstable residence, poverty, boredom and weak pro-social ties. 3) Social learning as reflected by anti-social peers, anti-social neighborhood, parent criminal behavior and anti-social thinking. 4) Poor socialization is suggested by parental crime and family disorganization, early juvenile onset, early failure in school, criminal attitudes.

The official criminal history matches this extreme criminogenic profile. It has the most chronic and dangerous criminal career with the highest scores for criminal involvement, juvenile onset, non-compliance and violent and versatile offending. These people have the highest scores for arrests and convictions for robbery, burglary, weapon offenses, assaults, injury to victims, violent felonies, fraud, drug possession and domestic violence arrests.

Category 8 - Low risk situational accidental category

Like Category 2, this category reflects lower criminogenic risks and more pro-social strengths than most other categories. Thus, this profile offers no clear explanation for their engagement in the criminal justice system. Like Category 2, these persons reflect perhaps "normal" folks who became embroiled in a situational-accidental event that led to entry into the criminal justice system. Many members of this category will have less poverty, more adequate jobs and education, more stable residence in safer areas than most persons in this population. They appear mostly to avoid anti-social peers and criminal opportunities and may have pro-social ties. Their attitudes and personalities are not clearly anti-social. They report low drug use (compared to other groups), fewer criminal peers, lower family crime and positive use of leisure.

The criminal history of this category confirms its low risk, non-violent status. Most have few prior arrests and for many this may be their first incarceration. They generally have fewer felonies or weapons offenses, and less history of probation or probation failure. Most are assigned to the lowest risk category by the COMPAS risk models.

The current arrest pattern perhaps explains the situational nature of this category. Specifically, they have the lowest (mostly zero) scores for felony charges, assaultive felonies, weapons offenses, victim injury, family violence, burglary/larceny, robbery and drug offenses. In many cases their arrests are alcohol related, simple assault, drunk driving, non-felony fraud or minor property offense, or a sex offense. Thus, it is prudent to check the details (if available) of the current offense of persons in this category.

An important caution is that a small percentage of this type may be "faking good" as indicated by the Lie Test score. Thus, while many are truly low risk (as confirmed by official

history) a small percentage may be lying. Thus, it is still prudent to show caution with these persons.

5.3 Female Typology

5.3.1 Type Descriptions

Category 1 - Drug problems and anti-social sub-cultural influences – some with relationship conflicts

This group (average age 35) appears locked into a high-risk sub-culture e.g. anti-social peers, anti-social family and residence in a high-risk crime environment. Some reflect early onset of teenage delinquency and cocaine use as a juvenile. Chronic drug problems are suggested by above average scores for previous drug treatments and drug possession charges. Many of these women hold anti-social attitudes. This profile suggests a social learning process where these women are socialized within an anti-social drug sub-culture. However, some strengths are still present for some of these women, e.g., stable housing, adequate use of leisure time and apparently good social support. The group criminal history is about average and not noticeably violent – although the group is above average for jail and probation terms, prior convictions and non-compliance history. For some of these women their current domestic violence charges suggest relationship conflict.

Category 2 - Family disorganization and inadequate parenting – residential instability and minor non-violent offences

This younger group has an average age of 25 years. Early family disorganization, abuse and inadequate parents appear central. Their high scores for family criminality and juvenile out-of-home placements suggests inadequate parenting. Their high juvenile socialization score also suggests early onset of problems. Their adult life challenges include residential instability and social adjustment problems. However, several positive features emerge for some of these women, i.e., lower than average scores for criminal peers, below average scores for criminal attitudes and criminal personality. Many of these women appear to avoid drugs, with relatively few reporting drug treatment or use of drugs as juveniles. The profile suggests some positive social supports and fairly constructive use of leisure time. The criminal history is consistent with the above profile and is mostly non-violent and fairly low for non-compliance. The most common current charge is minor fraud. Mental health issues may be explored given the possibility of early family abuse and/or neglect.

Category 3 - Chronic substance abusers – women with higher social resources than other groups

This older (average age 38) category shows less poverty, more positive education and vocational skills and residence in an apparently safe low crime areas, than other categories. These positive features are consistent with lower than average scores for criminal associates, lower anti-social attitudes and a fairly positive use of leisure time. The group appears to have relatively fewer social adjustment problems, better social supports and a lifestyle that avoids high risks and criminal opportunity. They do not have high scores for criminal personality.

The official data matches this profile with a relatively late onset, mostly minor offenses and few juvenile problems.

DUI is the most frequent current offence among these mostly non-violent women – although some also have domestic a violence record. However, the presence of prior convictions, prior drug offences and frequency of prior treatments for drugs and/or alcohol underscores a clearly chronic substance abuse problem.

Category 4 - Marginalized poor and isolated older women – economic survival crimes

The average age of this group is 40 years. This group is characterized by poverty, social isolation and a lower than average constructive leisure activities. This group has a late onset with an average age at first arrest of 27 years. Their criminal history mainly involves minor fraud. Aside from poverty they show few other criminogenic factors. For example, they fall below average for criminal peers, antisocial attitudes, living in high crime areas or following high opportunity lifestyles. Their family of origin appears relatively law abiding. Their history exhibits few juvenile problems. It appears that their problems mostly emerge in adulthood from poverty and poor social support. Their instrumental crimes such as minor fraud and sex offences may be for economic survival. Their poor social adjustment and social isolation suggest screening for mental health problems. The risk assessment assigns most of these women to a low risk non-violent category.

Category 5 - Young antisocial poorly educated women with some violent offences and early delinquency onset

This younger category (average age is 25) has a limited adult criminal history - with relatively few adult arrests or convictions - but the highest score for a current violent offence, some involving felony and weapons charges. Their criminogenic factors include: early onset of delinquency, above average antisocial personality, antisocial attitudes, poor education/vocational resources, bored/unproductive use of leisure hours and pessimism about finding a good job. Early delinquency is reflected in higher than average juvenile marijuana and alcohol use (but fewer hard drugs), high school dropout and the earliest first arrest. Surprisingly, the group has relatively low affiliation with antisocial peers or gangs; no clear tendency to live in high crime areas, abuse drugs, or to have extreme poverty or a high crime family background. Their relatively low formal adult criminal histories, appear consistent with their average scores on COMPAS risk assessment scales. However, the presence of early onset delinquency and, in some cases, serious current violence suggests caution with this group.

Category 6 - Chronic long term criminal history A – multiple co-occurring social and psychological risk factors

Drugs, extreme socio-economic marginalization, teen onset of problems and extreme problems in social relations characterize this high risk category. The recidivism risk computation

identifies this group as high risk. Multiple criminogenic factors co-occur, including: antisocial peers, antisocial attitudes, antisocial personality, extreme substance abuse, high crime family, poverty, extreme vocational and educational deficits, inability to use leisure time constructively and a tendency to live a high risk life style. Problems started early and these women report the highest levels for out-of-home placements as juveniles, the worst school grades, the highest use of cocaine as a juvenile, the earliest first arrest and the highest number of juvenile felony arrests. This is a non-compliant group with multiple failures and extreme drug problems. Social isolation and social adjustment problems are high. This group commits a variety of offences, including: domestic violence, drug possession, and other assault.

Category 7 - Chronic long term criminal history B – multiple co-occurring problems and high risk

This rare and infrequent group is a more serious version of type 6. While both categories have multiple co-occurring risk and need factors group 7 is systematically higher than group 6. This category has the highest scores for: violence risk, recidivism risk, FTA risk and technical violation risk. They are highest for: overall criminal history, history of non-compliance, current violence and juvenile delinquency indicators. The multiple criminogenic factors include: residential instability, family crime, vocational-educational failures, antisocial attitudes, antisocial personality, social adjustment problems, social isolation/withdrawal, extreme drug use and so on. Compared to Category 6, this group has the highest scores for current violence, injuries to victims, current felony arrests and current robberies. They exhibit extreme poverty, live in higher risk areas and report more gang affiliations.

Category 8 - Late starters with multiple strengths and fewer risk factors – minor non-violent offence history

These women (like pattern 3) reflect higher resources than other groups for educational and vocational scores, jobs, completing high school, living in safer areas, stable housing, better social supports and fewer leisure problems. Their family background appears more pro-social and they report less poverty, antisocial attitudes or personality issues. This group appears to adopt safer lifestyles by avoiding anti-social persons, fewer drug problems and more pro-social leisure activities. While, we may be suspicious of this positive profile, their official criminal history is consistent with this low risk profile showing the lowest criminal involvement and incarcerations, the fewest arrests and convictions, the lowest arrest rate, the lowest felony charges, the lowest pending charges, less non-compliance and the oldest age at first arrest (average age is 27). Current charges reflect minor fraud and DUI. This official data therefore coheres with this low need/risk profile. However, some women in this category may be "faking good." This was detected using the built-in COMPAS validity test for defensive faking-good responses and notice should be taken of this warning.

Bibliography

Andrews, D. A., & Bonta, J. (1994). The psychology of criminal conduct. Cincinnati: Anderson Press.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. Criminal Justice and Behavior, 17(1), 19-52.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or needs assessment. Crime and Delinquency, 52(1), 7-27.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. Criminology, 28(3), 369-403.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May). Machine bias. ProPublica.

Babchishin, K. M., & Helmus, L. M. (2016). The influence of base rates on correlations: An evaluation of proposed alternative effect sizes with real-world data. Behavior Research Methods, 48(3), 1021-1031.

Baird, C. (2009). A question of evidence: A critique of risk assessment models used in the justice system. (Tech. Rep.). Madison, WI: National Council on Crime and Delinquency.

Baird, S., Heinz, R., & Bemus, B. (1979). Project report 14: Two-year follow-up.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. Journal of Personality and Social Psychology, 71(2), 364-374.

Barbaree, H. E., Seto, M., Langton, C. M., & Peacock, E. J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. Criminal Justice and Behavior, 28(4), 490-521.

Barnoski, R., & Aos, S. (2003). Washington's Offender Accountability Act: An analysis of the Department of Corrections' risk assessment (document no. 03-12-1202) (Tech. Rep.). Olympia, WA: Washington State Institute for Public Policy.

Barnoski, R., & Drake, E. K. (2007). Washington's Offender Accountability Act: Department of Corrections' static risk assessment (Tech. Rep.). Olympia, WA: Washington State Institute for Public Policy.

Berk, R. A. (2012). Criminal justice forecasts of risk: A machine learning approach. New York, NY: Springer.

Berk, R. A. (2016). A primer on fairness in criminal justice risk assessments. The Criminologist, 41(6), 6 - 9.

Berk, R. A., & Bleich, J. (2013). Statistical procedures for forecasting criminal behavior: A comparative assessment. Criminology & Public Policy, 12, 513-544.

Berk, R. A., Heidari, H., Jabbari, S., Kearns, M., & Roth, A. (2018). Fairness in criminal justice risk assessment: The state of the art. Sociological Methods & Research. (Advance online publication. doi.org/10.1177/0049124118782533)

Blomberg, T., Bates, W., Mann, K., Meldrum, R., & Nedelec, J. (2010). Validation of the COMPAS risk assessment classification instrument (Tech. Rep.). Tallahassee, Florida: Center for Criminology and Public Policy Research, College of Criminology and Criminal Justice, Florida State University.

Bonta, J. (1996). Risk-needs assessment and treatment. Thousand Oaks, CA: Sage.

Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. Psychological Bulletin, 123(2), 123–142.

Brennan, T., & Dieterich, W. (2008). Michigan Department of Corrections Core COMPAS pilot study: One-year follow-up (Tech. Rep.). Traverse City, MI: Northpointe Institute for Public Management.

Brennan, T., & Dieterich, W. (2009). Testing the predictive validity of the DPCA COMPAS risk scales: Phase I (Tech. Rep.). Traverse City, MI: Northpointe Institute for Public Management.

Brennan, T., Dieterich, W., & Ehret, B. (2009). Evaluating the predictive validity of the COMPAS risk and needs assessment system. Criminal Justice and Behavior, 36, 21-40.

Brennan, T., & Oliver, W. L. (2013). The emergence of machine learning techniques in criminology. Criminology and Public Policy, 12, 551-562.

Bucklen, K. B., & Zajac, G. (2009). But some of them don't come back (to prison!): Resource deprivation and thinking errors as determinants of parole success and failure. The Prison Journal, 89, 239-264.

Bureau of Justice Statistics. (2012). Correctional populations in the United States, 2011. Washington, D.C.: U.S. Department of Justice, Office of Justice Programs.

Chaiken, J., Chaiken, M., & Rhodes, W. (1994). Predicting violent behavior and classifying violent offenders. In A. Reiss & J. Roth (Eds.), Understanding and preventing violence, vol. 4: Consequences and control (pp. 217–295). Washington, D.C: National Academy Press.

Cohen, J., & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. American Sociological Review, 44, 588-608.

Coleman, J. S. (1990). Foundations of social theory. Cambridge,Mass: Harvard University Press.

Committee on Assessing Juvenile Justice Reform. (2013). Reforming juvenile justice: A developmental approach (R. J. Bonnie, R. L. Johnson, B. M. Chemers, & J. A. Schuck, Eds.). Washington, D.C.: The National Academies Press.

Cooke, D. J., Forth, A. E., & Hare, R. D. (Eds.). (1998). Psychopathy: Theory, research and implications for society. Dordecht, Netherlands: Kluwer Academic Publishers.

Copas, J., & Marshall, P. (1998). The offender group reconviction scale: A statistical reconviction score for use by probation officers. Applied Statistics, 47, 159 - 171.

Dahle, K. P. (2006). Strengths and limitations of actuarial prediction of criminal re-offence in a German prison sample: A comparative study of LSI-R, HCR-20 and PCL-R.

International Journal of Law and Psychiatry, 29(5), 431-442.

Desmarais, S. L., & Singh, J. P. (2013). Risk assessment instruments validated and implemented in correctional settings in the United States. (Tech. Rep.). Lexington, KY: Council of State Governments Justice Center.

Dieterich, W. (2010). Kent County Pretrial Services outcomes study: Developing and testing the COMPAS pretrial release risk scale. (Tech. Rep.). Traverse City, MI: Northpointe.

Dieterich, W., Brennan, T., & Oliver, W. L. (2011). Predictive validity of the COMPAS Core risk scales: A probation outcomes study conducted for the Michigan Department of Corrections (Tech. Rep.). Traverse City, MI: Northpointe Inc.

Dieterich, W., Mendoza, C., & Brennan, T. (2016). Compas risk scales: Demonstrating accuracy equity and predictive parity: Performance of the COMPAS risk scales in Broward County (Tech. Rep.). Traverse City, MI: Northpointe Inc.

Dieterich, W., Mendoza, C., Hubbard, D., Ferro, J., & Brennan, T. (2017). COMPAS risk scales validation study: An outcomes study conducted for the Santa Barbara County Probation Department (Tech. Rep.). Traverse City, MI: Northpointe.

Dieterich, W., Mendoza, C., Hubbard, D., Ferro, J., & Brennan, T. (2018). COMPAS risk scales validation study: An outcomes study conducted for the Riverside County Probation Department (Tech. Rep.). Traverse City, MI: Northpointe.

Dieterich, W., Oliver, W., & Brennan, T. (2011). Predictive validity of the Reentry COMPAS Risk scales: An outcomes study with extended follow-up conducted for the Michigan Department of Corrections. (Tech. Rep.). Traverse City, MI: Northpointe Inc.

Division of Probation and Correctional Alternatives. (2007). New York State pretrial release services standards.

Dodge, K. A. (1998). The structure and function of reactive and proactive aggression. In D. J. Pepler & K. H. Rubin (Eds.), The development and treatment of childhood aggression (pp. 201–218). Hillsdale, N.J.: Lawrence Erlbaum.

Dodge, K. A., Pettit, G. S., McClaskey, C. L., & Brown, M. M. (1986). Social competence in children. Monograph of the Society for Research in Child Development, 51(2), 1–85.

Elliot, D. S., Huizinga, D., & Ageton, S. (1985). Explaining delinquency and drug use. Beverly Hills, CA: Sage Publications.

Estroff, S. E., & Zimmer, C. (1994). Social networks, social support, and violence among persons with severe, persistent mental illness. In J. Monahan & H. J. Steadman (Eds.), Violence and mental disorder: Developments in risk assessment (pp. 259–295). Chicago: University of Chicago Press.

Farabee, D., Zhang, S., Roberts, R. E., & Yang, J. (2010). COMPAS validation study: Final report (Tech. Rep.). UCLA Integrated Substance Abuse Programs.

Farrington, D. P. (1991). Childhood aggression and adult violence: Early precursors and later life outcomes. In D. J. Pepler & K. H. Rubin (Eds.), The development and treatment of childhood aggression (pp. 5–29). Hillsdale, NJ: Lawrence Erlbaum.

Farrington, D. P., Jolliffe, D., Loeber, R., Stouthamer-Loeber, M., & Kalb, L. M. (2001). The concentration of offenders in families, and family criminality in the prediction of boys' delinquency. Journal of Adolescence, 24, 579–596.

Fass, T., Heilbrun, K., DeMatteo, D., & Fretz, R. (2008). The LSI-R and the COMPAS. Criminal Justice and Behavior, 35, 1095–1108.

Fischer, C. (1995). The subcultural theory of urbanism: A twentieth year assessment. American Journal of Sociology, 101, 543 - 577.

Flores, A. W., Bechtel, K., & Lowenkamp, C. (2016). False positives, false negatives, and false analyses: A rejoinder to Machine bias: There's software used across the country to predict future criminals. And it's biased against Blacks. Federal Probation, 80(2), 38-46.

Flores, A. W., Lowenkamp, C. T., Smith, P., & Latessa, E. J. (2006). Validating the Level of Service Inventory-Revised on a sample of federal probationers. Federal Probation, 70(2), 44-78.

Gendreau, P., Goggin, C., & Little, T. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! Criminology, 34, 575-607.

Gendreau, P., Goggin, C. E., & Law, M. A. (1997). Predicting prison misconducts. Criminal Justice and Behavior, 24(4), p414 - 431.

Gottfredson, M. R., & Hirschi, T. (1990). A general theory of crime. Stanford: CA: Stanford University Press.

Gottfredson, S. D., & Jarjoura, G. R. (1996). Race, gender, and guidelines-based decision making. Journal of Research in Crime and Delinquency, 33, 49-69.

Gottfredson, S. D., & Moriarty, L. J. (2006). Statistical risk assessment: Old problems and new applications. Crime & Delinquency, 52, 178-200.

Grann, M., Belfrage, H., & Tengstrom, A. (2000). Actuarial assessment of risk for violence: Predictive validity of the VRAG and the historical part of the HCR-20. Criminal Justice and Behavior, 27, 97-114.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. Psychological Assessment, 12(1), 19–30.

Hagan, J. (1998). Life course capitalization and adolescent development. In R. Jessor (Ed.), New perspectives on adolescent risk behavior (pp. 499–517). New York: Cambridge University Press.

Hamilton, Z., Kigerl, A., Campagna, M., Barnoski, R., Lee, S., van Wormer, J., & Block, L. (2017). Designed to fit: The development and validation of the STRONG-R recidivism risk assessment. Criminal Justice and Behavior, 43, 230-263.

Hanson, R. K. (1997). The development of a brief actuarial risk scale for sexual offense recidivism. (User Report 97-04). Ottawa, Ontario: Department of the Solicitor General of Canada.

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. Law and Human Behavior, 24, 119-136.

Hare, R. D. (1991). Manual for the Hare Psychopathy Checklist-Revised. Toronto, Canada: Multi-Health Systems.

Harrell, F. E., Jr. (2001). Regression modeling strategies. New York: Springer.

Harris, P., & Jones, P. R. (1999). Differentiating delinquent youths for program planning and evaluation. Criminal Justice & Behavior, 26(4), 403-434.

Hastie, T., Tibshirani, R., & Friedman, J. (2008). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). New York: Springer.

Herrenkohl, T. I., Maguin, E., Hill, K. G., Hawkins, J. D., Abbott, R. D., & Catalano, R. F. (2000). Developmental risk factors for youth violence. Journal of Adolescent Health, 26, 176–186.

Hirschi, T. (1969). Causes of delinquency. Berkeley, CA: University of California Press.

James, G., Witten, D., Hastie, T., & Tishirani, R. (2013). An introduction to statistical learning with applications in R. New York: Springer.

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2016, November). Inherent trade-offs in the fair determination of risk scores. StatMl Papers. Retrieved from https://arxiv.org/abs/1609.05807v2

Knight, K., Simpson, D. D., & Morey, J. T. (2002). Evaluation of the TCU Drug Screen (Tech. Rep.). Fort Worth, TX: Institute of Behavioral Research Texas Christian University.

Kroner, C., Stadtland, C., Eidt, M., & Nedopil, N. (2007). The validity of the violence risk appraisal guide (VRAG) in predicting criminal recidivism. Criminal Behaviour and Mental Health, 17(2), 89 - 100.

Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. New York: Springer.

Lansing, S. (2012). New York State COMPAS-Probation Risk and Needs Assessment Study: Evaluating predictive accuracy (Tech. Rep.). Albany, New York: New York State Division of Criminal Justice Services, Office of Justice Research and Performance.

Latessa, E. J., Lemke, R., Makarios, M., Smith, P., & Lowenkamp, C. T. (2010). The creation and validation of the Ohio Risk Assessment System (ORAS). Federal Probation, 74(1), 16-22.

Latessa, E. J., Lovins, B., & Makarios, M. (2013). Validation of the indiana risk assessment system: Final Report (Tech. Rep.). University of Cincinnati School of Criminal Justice.

Latessa, E. J., Lux, J. L., Lugo, M., & Long, J. (2016). Examining the validity and reliability of the Ohio Risk Assessment System Community Supervision Tool and Community Supervision Screening Tool (Tech. Rep.). University of Cincinnati.

Latessa, E. J., Machak, S., Lux, J., Newsome, J., Lugo, M., & Papp, J. (2017). The Ohio Risk Assessment System (ORAS): A re-validation & inter-rater reliability study (Tech. Rep.). Cincinnati, OH: The University of Cincinnati Corrections Institute. (Prepared for the Ohio Department of Rehabilitation and Correction)

Linn, S. (2004). A new conceptual approach to teaching the interpretation of clinical tests. Journal of Statistics Education, 12, 1 - 9.

Lovins, B. K., Latessa, E. J., May, T., & Lux, J. (2017). Validating the ohio risk assessment system community supervision tool with a diverse sample from texas. Corrections, 0(0), 1-17. Retrieved from https://doi.org/10.1080/23774657.2017.1361798 doi: 10.1080/23774657.2017.1361798

Lykken, D. (1995). The antisocial personalities. Hillsdale, N.J.: Lawrence Erlbaum.

McGrath, R. J., & Hoke, S. E. (2001). Vermont Assessment of Sex Offender Risk. Waterbury, VT: Author

McGrath, R. J., Lasher, M. P., Cumming, G. F., Langton, C. M., & Hoke, S. E. (2014). Development of the Vermont Asessment of Sex Offender Risk-2 (VASOR-2) Reoffense Risk Scale. Sexual Abuse: A Journal of Research and Treatment, 26, 271-290.

Mealey, L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. Behavioral and Brain Sciences, 18(3), 523-599.

Megargee, E., & Bohn, M. (1979). Classifying criminal offenders: A new system based on the MMPI. Beverly Hills, CA: Sage.

Mendoza, C., Dieterich, W., Oliver, W. L., & Brennan, T. (2016, November). Decision analysis and equity. Paper presented at the Annual Meeting of the American Society of Criminology. Philadelphia, PA.

Moffitt, T. E. (1993). Adolescence-limited and life-course persistent antisocial behavior: A developmental taxonomy. Psychological Review, 100, 674–701.

Monahan, J., & Skeem, J. (2014). Risk redux: The resurgence of risk assessment in criminal sanctioning. Federal Sentencing Reporter, 26, 158-166.

National Association of Pretrial Services Agencies. (2004). National Association of Pretrial Services Agencies standards on pretrial release (3rd edition) (Tech. Rep.). National Association of Pretrial Services Agencies.

Nelson, M., Deess, P., & Allen, C. (1999). The first month out: Post-incarceration experiences in New York City. New York: Vera Institute.

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). New York: McGraw-Hill.

Osgood, D. W., Wilson, J. K., O'Malley, P. M., Bachman, J. G., & Johnston, L. D. (1996). Routine activities and individual deviant behavior. American Sociological Review, 61(4), 635-655.

Palmer, T. (1994). A profile of correctional effectiveness and new directions for research. Albany, NY: SUNY Press.

Parker, J. G., & Asher, S. R. (1987). Peer relations and later personal adjustment: Are low-accepted children at risk. Psychological Bulletin, 102(3), 357–389.

Pepe, M. S. (2003). The statistical evaluation of medical tests for classification and prediction. New York: Oxford University Press.

Petersilia, J. (2003). When prisoners come home: Parole and prisoner reentry. New York: Oxford University Press.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). Violent offenders: Appraising and managing risk. Washington, DC: American Psychological Association.

Reich, W. A., Picard-Fritsche, S., Barber Rioja, V., & Rotter, M. (2016). Evidence-based risk assessment in a mental health court: A validation study of the COMPAS risk assessment (Tech. Rep.). New York: Center for Court Innovation.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's d, and r. Law and Human Behavior, 29(615-620).

Sampson, R., & Lauritsen, J. (1994). Deviant lifestyles proximity to crime and the offender-victim link in personal violence. Journal of Research in Crime and Delinquency, 27, 7–40.

Serin, R. C., & Lowenkamp, C. T. (2015). Selecting and using risk and need assessments (Vol. X) (No. I). Alexandria, VA: National Drug Court Institute.

Singh, J. P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. Behavioral Sciences and the Law, 31, 8-22.

Smith, P., Cullen, F. T., & Latessa, E. J. (2009). Can 14,737 women be wrong? A meta-analysis of the LSI-R and recidivism for female offenders. Criminology and Public Policy, 8, 183-208.

Solomon, A. L., Visher, C., La Vigne, N. G., & Osborne, J. (2006). Understanding the challenges of prisoner reentry: Research findings from the Urban Institute's prisoner reentry portfolio. Washington, DC: Urban Institute Press.

Stansfield, R., & Williams, K. R. (2014). Predicting family violence recidivism using the DVSI-R. Criminal Justice and Behavior, 41, 163-180.

Stouthamer-Loeber, M., Loeber, R., Wei, E., Farrington, D. P., & Wikstrom, P.-O. H. (2002). Risk and promotive effects in the explanation of persistent serious delinquency in boys. Journal of Consulting and Clinical Psychology, 70, 111–123.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. Science, 240, 1285-1293.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. Psychological Science in the Public Interest, 1, 1–26.

Thornberry, T. P., Huizinga, D., & Loeber, R. (1995). The prevention of serious delinquency and violence: Implications from the program of research on the causes and correlates of delinquency. In J. C. Howell, B. Krisberg, J. D. Hawkins, & J. J. Wilson (Eds.), Sourcebook on serious, violent and chronic juvenile offenders (pp. 213–237). Thousand Oaks, CA: Sage.

Tolan, P. H., & Gorman-Smith, D. (1998). Development of serious and violent offending careers. In R. Loeber & D. P. Farrington (Eds.), Serious and violent juvenile offenders: Risk factors and successful interventions (pp. 68–85). Thousand Oaks, CA: Sage.

Travis, J. (2005). But they all come back: Facing the challenges of prisoner reentry. Washington, DC: Urban Institute Press.

VanNostrand, M. (2003). Assessing risk among pretrial defendants in Virginia: The Virginia pretrial risk assessment instrument (Tech. Rep.). Virginia Department of Criminal Justice Services.

VanNostrand, M. (2007). Legal and evidence-based practices: Application of legal principles, laws, and research to the field of pretrial services (Tech. Rep.). National Institute of Corrections and Crime and Justice Institute.

Van Voorhis, P., Wright, E., Salisbury, E., & Bauman, A. (2010). Women's risk factors and their contributions to existing risk/needs assessment: The current status of a gender-responsive supplement. Criminal Justice And Behavior, 37, 261-288.

Vose, B., Cullen, F. T., & Smith, P. (2008). Empirical status of the Level of Service Inventory. Federal Probation, 72, 22-29.

Warren, M. Q. (1971). Classification of offenders as an aid to efficient management and effective treatment. Journal of Criminal Law, Criminology, and Police Science, 62, 239–258.

West, D. J. (1973). Who becomes delinquent? London: Heinemann.

Williams, K. R., & Grant, S. R. (2006). Empirically examining the risk of intimate partner violence: The Revised Domestic Violence Screening Instrument (DVSI-R). Public Health Reports, 121, 400-408.