

An executive summary drawn from Sarah L. Desmarais and Jay P. Singh, Instruments for Assessing Recidivism Risk: A Review of Validation Studies Conducted in the U.S. (New York: Council of State Governments Justice Center, 2013).

INTRODUCTION

The rates of crime, incarceration, and correctional supervision are disproportionately high in the U.S. and translate into exorbitant costs to individuals, the public, and the state. Within three years of release from jail or prison, two-thirds of offenders are rearrested and half are incarcerated for a new crime or parole violation.¹ Though many offenders recidivate, a considerable proportion do not. Thus, there is a need to identify those offenders at greater risk of recidivism and to allocate resources and target risk management and rehabilitation efforts accordingly. Risk assessment, a crucial component to implementing evidence-based recidivism reduction strategies, is the process of estimating the likelihood an offender will recidivate by identifying those offenders at higher risk and in greater need of interventions. Assessment results, based on ratings of empirically or theoretically based risk and/or protective factors, can be used to determine intervention targets, appropriate programming level and intensity, and supervision level. There is overwhelming evidence to suggest that assessments of risk completed using structured approaches produce estimates that are both more accurate and more consistent across assessors compared to subjective or unstructured approaches.² More and more, structured risk assessment approaches are being used in correctional agencies.³

The nearly 100-year history of risk assessment instrument development has been documented on several occasions.⁴ The focus and structure of risk assessment tools have shifted significantly over time. Studies have identified the important evolution of risk assessment and predictability of criminal behavior,* from Burgess' (1928)⁵ seminal work on the application of a systemized prediction methodology through the progression of the four distinct generations summarized below.

The **first generation** of risk assessment is best described as unstructured professional judgment, in which the assessor relies on his or her professional training and information gathered from the incarcerated individual, official records, or other sources to inform his or her evaluation of risk for recidivism. It is "unstructured" insofar as there is no set checklist or protocol for completing the risk assessment, though assessors may indeed complete structured interviews during the risk assessment process. This method of assessment was widely accepted for decades prior to the development of structured risk assessment tools in the 1970s. Today, it is less frequently used but nonetheless remains a prominent risk assessment strategy despite evidence that unstructured assessments are not particularly accurate.

* Often defined in terms of reoffending.

Following decades of research focused on identifying factors that increase risk of recidivism, **second generation** tools represent a drastic advance in risk assessment technology. Second generation tools are actuarial in nature and evaluate primarily historical and static factors (e.g., sex, age, and criminal history). Rather than subjective judgments of recidivism risk, instruments such as the Salient Factor Score (SFS) and Violent Risk Appraisal Guide (VRAG) instead guide assessors to consider a set list of risk factors to arrive at a numerical risk of recidivism.

The **third generation** of risk assessment is characterized by the development of tools that include dynamic factors and criminogenic needs, and may use an actuarial or structured professional judgment approach. Third generation tools, such as the Level of Service Inventory-Revised (LSI-R), the Self-Appraisal Questionnaire (SAQ), and the Historical-Clinical-Risk Management-20 (HCR-20), still guide assessors to consider static factors; however, by including potentially dynamic items, such as attitude and substance use, they may be sensitive to change in risk levels over time and can assist in identification of treatment targets. These tools are sometimes referred to as “risk-need” instruments and, unlike second-generation assessments, tend to be theoretically and empirically based as opposed to wholly data-driven.

Most recently, **fourth generation** risk assessments explicitly integrate case planning and risk management into the assessment process. As such, the primary goal of the fourth generation extends beyond assessing risk and focuses on enhancing treatment and supervision. Examples of fourth generation tools include the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), Ohio Risk Assessment System (ORAS), and Wisconsin Risk and Needs Tool (WRN). Like the third generation, this generation of risk assessment instruments allows for the role of professional judgment while remaining grounded in research and theory.

This history and scientific research supports a greater use of risk assessment tools within a variety of clinical, law enforcement, judicial, and correctional contexts. Though variability exists in performance across instruments and assessment procedures, the tools perform a critical role in helping to manage public safety. These tools are designed to implicitly ground and offer support to sound and empirically based decision making, service, and supervision through multiple phases in the criminal justice system: intake, pre-release, release/community supervision, and case closure.

The increasing use of risk assessment instruments to guide decision making not only reflects the national, state, and local leadership advocacy on reentry and recidivism reduction but the commitment to legislative action and statutory changes. Growing community support and media attention has similarly demanded the adoption of these tools. For example, the Bureau of Justice Assistance, a component of the Office of Justice Programs, U.S. Department of Justice, which provides leadership and services in grant administration and criminal justice

policy development to support local, state, and tribal justice strategies to achieve safer communities, has strongly encouraged agencies responding to funding solicitations to adhere to the principles of effective correctional intervention, including the use of risk assessment instruments. Through extensive research, the current work identified no fewer than 66 risk assessment tools. While not comprehensive of the field or marketplace, this research has identified 19* standardized criminal justice—specific instruments in broad use to assess the risk of general offending and 47 instruments designed for jurisdiction-specific implementation.

Similar to other disciplines and fields of analysis, the work and practitioners of risk assessment in criminal justice settings are highly dispersed and often highly specialized. The development of distinct risk instruments demonstrates the diversity of users as well as the distinct practices of each state and local government, as they relate to crime-reduction laws, statutes, and strategies. Risk assessment instruments share several key commonalities that represent the foundation of evidence-based practices and principles to improve public safety.

With that in mind, the goal of this guide is to provide a framework to help criminal justice and social service agencies, practitioners, and policymakers evaluate these tools. It outlines the components and parameters of risk assessment tools validated and implemented in correctional settings in the U.S. Additionally, this guide provides a review of and catalogs the available knowledge regarding the accuracy and predictive validity of the risk assessment instruments for adult offenders. Furthermore, the guide presents recommendations of the steps that might be taken to improve public safety outcomes associated with the implementation of criminal justice risk assessment tools.

Throughout the U.S., policymakers, practitioners, researchers, and government officials rely heavily on risk assessment tools to focus criminal justice best practices, direct resources, and support recidivism reduction strategies. It is clear that as “laboratories of innovation,” states and local jurisdictions have helped to spearhead and facilitate risk-instrument development and implementation. Moreover, the federal government—notably, the Department of Justice—has spurred risk-instrument adoption as a best practice and priority for state reentry plans and recidivism-reduction efforts. The synergy of government and correctional leadership and commitment reflects a principle articulated by President Johnson’s Crime Commission: “the greatest need” in criminal justice “is the need to know.”

* Criteria for instruments to be included in the review were: a) designed to assess the likelihood of general recidivism (i.e., new offenses and violation of conditions); b) intended for assessing adult offenders (18 years of age and older); c) used in correctional settings in the United States; and d) validated in the United States. Instruments were excluded from our review if they: a) were designed to assess the likelihood of adverse outcomes for specific offenses (e.g., sexual offenses, violent offenses, spousal assault); b) were intended for assessing juvenile offenders (less than 18 years of age); c) were not used in correctional settings in the United States; d) had not been validated in the United States.; or e) were developed for use in a specific institution or ward.

I. RELIABLE PREDICTIONS: DO THE INSTRUMENTS WORK?

THE IMPACT ON RESOURCES AND RELATION TO OFFENDER RISK, NEED, AND RESPONSIVITY (RNR)

Policymakers' *need to know* and the subsequent strategies for public safety and recidivism reduction might begin with a simple question: Do risk assessment instruments reliably predict recidivism?

The short answer, according to years and volumes of research, is resoundingly: **yes**.⁶ But we must be mindful of what saying *yes* may mean. Adoption of a risk assessment tool goes hand-in-hand with fundamentally altering approaches to reentry and correctional management, supervision, services, and more broadly criminal justice practice. Ultimately, the process of implementing risk assessments within an agency should consist of more than simply adding a tool to the agency portfolio; it should result in a shift of corrections culture, practices, and policies.

The resource-allocation practices of criminal justice agencies have not significantly changed, despite the fact that nearly two-thirds of offenders recidivate following release.⁷ Prison and community programs remain equally distributed across populations in jails, in prisons, or under community supervision, despite risk level. Officials historically prioritize services and treatment for people who demonstrate a willingness and initiative to participate in services. Purposeful or not, correctional personnel and institutions tend to perpetuate an ineffective one-size-fits-all approach to offender management and rehabilitation.

Research across jurisdictions and settings indicates the need for a different model, grounded in the effective use of risk assessment tools. In fact, multiple studies show that prioritizing resources for individuals with a greater likelihood of recidivating is key to improved outcomes. Specifically, focusing resources on individuals with higher risks and needs can lead to a significant reduction in recidivism, while conversely, intensive interventions for low-risk individuals has been shown to be an ineffective use of resources. Furthermore, exposure of programs to low-risk offenders may actually make matters worse and result in harm by increasing association with higher-risk offenders and disrupting relationships with pro-social supports and structures.⁸

With studies⁹ indicating the need to target and provide intensive services to higher-risk offenders as a smart, cost-effective public safety strategy, decision makers face the challenge of redirecting limited public correctional resources to the right offenders. The shift in fiscal reallocation, though not necessarily viewed favorably by individuals/constituents without

theoretical knowledge or perspective, is critical. Reducing the present scale and scope of correctional costs has become an unavoidable necessity at all of levels of government.

Indeed, it is estimated that in 2007, \$74 billion was spent on corrections in the U.S.¹⁰ When both direct and indirect costs are considered, estimates of annual costs have reached as high as \$1.7 trillion.¹¹ According to state and federal data, corrections spending has outpaced budget growth in education, transportation, and public assistance, with Medicaid costs rising more quickly than state corrections spending.¹² To address these increasing costs and associated deficits, criminal justice leaders and agencies have gradually adopted the Risk-Need-Responsivity (RNR) model of offender assessment and rehabilitation.¹³ The RNR model has increased efficacy in reducing recidivism and subsequently reduced taxpayer costs for prisons, jails, and unnecessary programs.¹⁴

The RNR model includes three principles: risk, need, and responsivity.

- *Risk Principle*: Focus supervision and services on the people most likely to reoffend.
- *Need Principle*: Address an individual's greatest criminogenic needs, defined as factors that contribute to risk but can change over time (e.g., social networks, thinking patterns, housing, substance use, finances, etc.).
- *Responsivity Principle*: Adapt interactions and services so that they enhance an individual's ability to learn and acquire new attitudes and skills.

To most effectively implement the RNR model with fidelity, systems/organizations must have access to valid and reliable assessments of recidivism risk. There are several specific issues and critical factors related to risk assessments that will be detailed in Section II: Risk Assessment Instruments.

Chart A: A COLLECTION OF RISK ASSESSMENT INSTRUMENTS IN USE

Community Risk/Needs Management Scale (CRNMS)

Correctional Assessment and Intervention System (CAIS)

Correctional Offender Management Profile for Alternative Sanctions (COMPAS)

Dynamic Factors Intake Assessment (DFIA)

Inventory of Offender Risks, Needs, and Strengths (IORNIS)

Level of Service instruments, including Level of Service/Case Management Inventory (LS/CMI), Level of Service/Risk Need, Responsively (LS/RNR), Level of Service Inventory (LSI), Level of Service Inventory-Revised (LSI-R), and Level of Service Inventory-Revised: Screening Version (LSI-R:SV)

Offender Assessment System (OASys)

Offender Group Reconviction Scale (OGRS)

Ohio Risk Assessment System, including the Ohio Risk Assessment System-Pretrial Assessment Tool (ORAS-PAT), Ohio Risk Assessment System-Community Supervision Tool (ORAS-CST), Ohio Risk Assessment System-Community Supervision Screening Tool (ORAS-CSST), Ohio Risk Assessment System-Prison Intake Tool (ORAS-PIT), and Ohio Risk Assessment System-Reentry Tool (ORAS-RT)

Federal Post Conviction Risk Assessment (PCRA)

Recidivism Risk Assessment Scales (RISc)

Risk Management System (RMS)

Risk of Reconviction (ROC)

Statistical Information of Recidivism Scale (SIR)

Salient Factor Score instruments

Self-Appraisal Questionnaire (SAQ)

Service Planning Instrument (SPIn) & Service Planning Instrument-Women (SPIn-W)

Static Risk and Offender Needs Guide (STRONG)

Wisconsin Risk/Needs (WRN) & Wisconsin Risk and Needs-Revised (WRN-R)

II. RISK ASSESSMENT INSTRUMENTS

As noted, the types and characteristics of assessment instruments in correctional and criminal justice settings are highly dispersed, and often highly specialized and customized, while sharing several commonalities. The 19 identified instruments captured in Chart A measure criminal risk.

How the tools predict risk, with what targets and in what population, varies. Risk assessment instruments include items that represent characteristics of the offender (e.g., physical health, mental health, and attitudes), his or her physical and/or social environment (e.g., neighborhood, family, and peers), or circumstances (e.g., living situation and employment status) that are associated with the likelihood of offending. Most frequently, recidivism risk assessment instruments focus on psychological and social characteristics. Despite the fact that macro-level dynamics, such as service, system, and societal variables, may also affect risk, these variables are rarely included in these risk assessment instruments.

Risk factors are those characteristics that increase risk of offending, whereas *protective factors* are those individual strengths that can serve to reduce risk and provide pathways out of criminal behavior. Inclusion of protective factors in risk-assessment instruments is relatively rare. However, there is mounting evidence that protective factors contribute unique information and improve predictive validity above and beyond consideration of risk factors alone.¹⁵

Risk and protective factors can either be static or dynamic. *Static factors* are historical or otherwise unchangeable characteristics (e.g., history of antisocial behavior) that help establish overall levels of risk and can help identify the level of intervention required. *Dynamic factors* are changeable characteristics (e.g., substance abuse) that establish relative levels of risk and help identify intervention targets; these factors can be either relatively *stable*, changing relatively slowly over time (e.g., antisocial cognition), or *acute*, changing relatively quickly over time (e.g., mood state).¹⁶ Identifying dynamic factors

allows for interventions to be appropriately targeted to the right offenders. Researchers Andrews and Bonta¹⁷ identified eight items as the most “powerful” risk factors for recidivism across offenders and situations. These “Central Eight” (Chart B) have been widely accepted as the most important domains to be assessed and targeted in risk assessment and management efforts.

Research consistently supports that targeting dynamic factors (criminogenic needs) adds incrementally to the predictive validity of static factors and that the former may be more relevant to short-term outcomes and rehabilitation efforts,¹⁸ whereas the latter relate more to longer-term outcomes and risk classification.¹⁹ Thus, there are important benefits to considering both static and dynamic factors in assessing recidivism risk.

Static factors are considered in the 19 instruments reviewed and dynamic factors are considered in 15 of the 19. Protective factors, however, are considered only in two instruments (IORNS and SPIn-W) (see Table 1). Moreover, there are differences in the type of items captured in risk tools and their relation to the Central Eight risk factors. While substance abuse is universally assessed across all 19 instruments reviewed, housing, recreation, and specific information on mental health domains are not usually incorporated into most tools. Table 2 outlines the content domains included in specific risk assessment instruments.

Chart B

The Central Eight—the most powerful risk factors for offenders and situations (Andrews & Bonta, 1998)

- History of antisocial behavior
- Antisocial personality pattern
- Antisocial cognition
- Antisocial associates
- Family and/or marital problems
- School and/or work problems
- Leisure and/or recreation problem
- Substance abuse

Table 1. Type of Factors Included in Risk Assessment Instruments

INSTRUMENTS	TYPES OF ITEMS			
	Risk	Protective	Static	Dynamic
COMPAS	•		•	•
IORNS	•	•	•	•
LSI-R	•		•	•
LSI-R:SV	•		•	•
ORAS-PAT	•		•	•
ORAS-CST	•		•	•
ORAS-CSST	•		•	•
ORAS-PIT	•		•	•
ORAS-RT	•		•	•
PCRA	•		•	•
RMS	•		•	•
SAQ	•		•	•
SFS74	•		•	
SFS76	•		•	
SFS81	•		•	
SPIIn-W	•	•	•	•
STRONG ^a	•		•	
WRN	•		•	•
WRN-R	•		•	•

^a The STRONG includes three parts: Static Risk Assessment, Offender Needs Assessment, and Offender Supervision Plan; values reflect only the first part, which is the component used to assess risk of recidivism.

TABLE 2: Risk Assessment Instrument Content Domains

INSTRUMENTS	Attitudes	Associates/ Peers	History of Antisocial Behaviour	Personality Problems	Relationships	Work/ School	Recreation/ Leisure Activities	Substance Use Problems	Mental Health Problems	Housing Status
COMPAS	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IORNS	✓	✓	✓	✓	✓	✓		✓	✓	
LSI-R	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
LSI-R:SV	✓	✓	✓		✓	✓		✓	✓	
ORAS-PAT			✓			✓		✓		✓
ORAS-CST	✓	✓	✓	✓	✓	✓	✓	✓		✓
ORAS-CSST		✓	✓			✓		✓		
ORAS-PIT		✓	✓	✓	✓	✓	✓	✓	✓	✓
ORAS-RT	✓		✓	✓	✓	✓		✓	✓	
PCRA	✓	✓	✓		✓	✓		✓		
RMS	✓	✓	✓	✓	✓	✓		✓	✓	✓
SAQ	✓	✓	✓	✓				✓		
SFS74			✓			✓		✓		✓
SFS76			✓			✓		✓		
SFS81			✓					✓		
SPIIn-W	✓	✓	✓		✓	✓	✓	✓	✓	✓
STRONG ^a			✓					✓		
WRN	✓	✓	✓		✓	✓		✓	✓	
WRN-R	✓	✓	✓		✓	✓		✓	✓	

^a The STRONG includes three parts, Static Risk Assessment, Offender Needs Assessment, and Offender Supervision Plan; values reflect only the first part, which is the component used to assess risk of recidivism.

In addition to measuring the likelihood of rearrest, reincarceration, or revocation, risk assessments can inform supervision and programming strategies for specific populations of offenders from pre-sentence to community supervision. Risk assessments predict the risk of a certain behavior within an established time period (often one or three years). There may be a need for a jurisdiction to use adjunct assessments, that address specific outcomes (e.g., violence, sexual offending) in conjunction with an assessment of general recidivism. Pretrial risk assessments often measure the risk for reoffending in addition to the risk for an offender to fail to appear in court at a scheduled date. Criminogenic risk assessment is not intended to provide detailed information around topics such as behavioral health or sex offending; instead, it generally serves to alert staff of areas where a more detailed assessment may be needed. Agencies seeking to predict specific behaviors may need to utilize specialized risk assessments (for behavioral health, violence, sex offending, etc.). Sex offenders often have low overall risk, so many agencies opt to conduct additional sex offender assessments (e.g., STATIC-99, ACUTE, STABLE) to gather more information related to the person's sexual offending pattern. It is not always feasible or advisable, however, to conduct issue-specific assessments around every conceivable topic, especially if relevant programming is not available or if the additional information will not change how the case is managed.²⁰ Hence, agencies should carefully consider and define the intended purpose of the selected risk tool. Moreover, it is imperative that agencies consider the evidence supporting the validity of assessments.

III. THE RISK INSTRUMENT AND PREDICTION

Though the types and characteristics of valid risk instruments are fairly well established, deciphering what's clear about risk—both in terms of the predictability of specific tools and more broadly, defining risk scores can be more complex.

Some of this complexity may stem from the varying expectations of users/providers and transferring the model into “real world” settings.²¹ To further explain, users may expect instruments to explicitly and precisely predict, with 100% accuracy, an offender’s participation in future criminal activity; however, we can never with certainty predict the likelihood of a future event. Yet, for those who recently adopted the instrument, there may exist a lack of understanding, training, or education on the theory and principles behind risk assessment tools, especially given the high turnover in corrections, public safety, and the human service industry.

To illustrate, consider the impact of age on risk of recidivism. It is well established by research²² that the propensity to participate in criminal activity significantly declines with increased age regardless of offense type. Yet, research on the culture of correctional institutions, as well as that on stigma, labeling, and operational norms in correctional settings, suggests that officials tend to direct more punishment toward offenders in specific offense categories, despite their limited knowledge of the offenders’ history, risk and/or protective factors.²³ What appears increasingly clear is that decision makers operate in an arena of uncertainty. As a result, decision makers may seek immediate solutions that have substantial public support with the goal of reducing the possibility of negative consequence.*

One attempt to mitigate uncertainty and improve decision-making is the integration of “structured approaches” into the process of conducting risk assessments. There is overwhelming evidence to suggest that assessments of risk completed using structured approaches produce estimates that are both more accurate and consistent across assessors compared to subjective or unstructured approaches.²⁴ The use of structured approaches to

There is overwhelming evidence to suggest that assessments of risk completed using structured approaches produce estimates that are both more accurate and consistent across assessors compared to subjective or unstructured approaches.

* For further explanation of this concept see James G. March and Herbert A. Simon, *Organizations* (Oxford: Wiley, 1958) as well as Jeffrey Lin, Ryken Grattet, and Joan Petersilia, “Back-End Sentencing’ and Reimprisonment: Individual, Organizational, and Community Predictors of Parole Sanctioning Decisions,” *Criminology* 48, no. 3 (2010): 759–795.

classify higher-risk individuals within the general offender population also produces better outcomes compared to unstructured approaches.²⁵ Correctional agencies are increasingly recommending—and many now require—the use of structured risk assessment approaches.²⁶

Approaches to Conducting Structured Risk Assessments

There are two broad approaches to structured risk assessment: actuarial and structured professional judgment. The approaches differ in how each addresses/responds to the following decision points: 1) which risk factors to consider and how to measure them; 2) how to combine risk factors; and 3) how to generate the final risk estimate.²⁷ The strengths and limitations of actuarial and structured professional judgment are reviewed below.

Actuarial Risk Assessment

The actuarial approach represents a mechanical model of risk assessment that is largely focused on historical or static risk factors. When an actuarial instrument is used to assess risk, an offender is scored on a series of items that were most strongly associated with recidivism in the development sample. The offender's total score is cross-referenced with an actuarial table that translates the score into an estimate of risk over a specified timeframe (e.g., 10 years). This estimate represents the percentage of participants in the instrument's development study who received that risk score and recidivated. For example, if an offender receives a score of +5 on an instrument that is translated into a risk estimate of 60 percent over 10 years, this means that 60 percent of those individuals who received a total score of +5 in the instrument's original study went on to recidivate within the specified time period. This does not mean that *this individual offender* has a 60 percent chance of recidivating over a period of 10 years. Understanding these distinctions is important; yet, they are frequently overlooked in practice.

Strengths of the actuarial approach include:

- ***Objectivity.*** No human judgment is involved in estimating risk once items have been rated. Items are typically straightforward and easy to rate (e.g., age, gender, number of prior offenses).
- ***Accuracy.*** Actuarial assessments are more accurate than unstructured assessments.
- ***Transparency.*** Information used to inform risk estimates is explicitly included in the instrument. Items are weighted in a predetermined manner to compute total scores and estimate risk.

- *Speed.* Items included in actuarial instruments can usually be scored using information available in official records.

The chief drawback of the actuarial approach is the application of group-based statistics and norms to individual offenders. Beyond the quandary-related potential statistical issues,^{*} this is a concern because it cannot be determined whether the specific percentage derived by averaging across the group applies to the individual or whether the probability found in the development study applies to this individual. Using the same example provided earlier, if 60 percent of the individuals who received a score of +5 recidivated over a 10-year period, then 40 percent did not. Actuarial assessments cannot help distinguish whether an offender receiving a score of +5 is among the 60 percent or 40 percent. Additionally, these tools may discount specific factors that do not systematically increase (or decrease) recidivism risk across the population but are relevant to a particular offender's level of risk. Lastly, many actuarial assessments have limited utility for individualized risk reduction and rehabilitation efforts due to their focus on historical or static factors (e.g., age, gender, number of prior offenses). With the focus on past behavior and not accounting for offenders changing for the better these tools are not useful for intervention planning or reassessment to measure individual progress.²⁸

Structured Professional Judgment

In contrast to the actuarial approach, the structured professional judgment (SPJ) approach focuses on creating individualized and clear risk formulations. As a result these assessments may support more comprehensive and integrated risk management plans. SPJ instruments guide assessors to estimate risk level (e.g., low, moderate, or high) through consideration of a set number of factors that are empirically and theoretically associated with the outcome of interest. Although offenders are scored on individual items, total scores are not used to make the final judgments of risk. Instead, assessors consider the relevance of each item to the individual offender, as well as whether there are any case-specific factors not explicitly included in the list. Strengths of the structured professional judgment approach include:

- *Professional discretion.* Assessors consider the relevance of factors to the individual offender to inform final estimates of each factor. Case-specific factors can also be taken into consideration.
- *Accuracy.* Structured professional judgment assessments are more accurate than unstructured assessments (and comparable in accuracy to actuarial assessments).

^{*} See Stephen D. Hart, Christine Michie, and David J. Cooke, "Precision of Actuarial Risk Assessment Instruments: Evaluating the 'Margins of Error' of Group v. Individual Predictions of Violence," *British Journal of Psychiatry* 190, no. 49 (2007): s60–s65.

- *Transparency.* Assessors rate a known list of factors according to specific guidelines. Additional items considered can be added to the assessment form.
- *Risk communication and reduction.* Risk formulations provide information regarding the anticipated series of stressors and events that may lead to adverse outcomes and over what period of time, which can inform risk management strategies and identify treatment targets.

Criticisms of SPJ include the potential reintroduction of decision-making biases in the final risk judgments and diminished predictive accuracy with the inclusion of dynamic risk factors. Reassessments should be conducted periodically to account for the change in dynamic risk factors over time. Additionally, these instruments are time-consuming and take comparatively longer to administer than actuarial assessments, item ratings are often more nuanced, and necessary information might not be readily available on file to code all items.

Despite the criticisms, recent reviews show that actuarial and structured professional judgment instruments produce assessments with commensurate rates of validity in predicting recidivism.²⁹

Table 3: Validity of Total Scores in Predicting Different Forms of Recidivism

INSTRUMENTS	OUTCOMES					
	<i>k</i>	General Offending (including violations)	<i>k</i>	General Offending (excluding violations)	<i>k</i>	Violations Only
COMPAS	–	–	5	Good	1	Fair
LSI-R	3	Good	26	Fair-Good	7	Good
LSI-R:SV	–	–	2	Fair-Good	–	–
ORAS-PAT	1	Fair	2	Fair	2	Good
ORAS-CST	–	–	1	Excellent	–	–
ORAS-CSST	–	–	1	Excellent	–	–
ORAS-PIT	–	–	1	Good	–	–
ORAS-RT	–	–	1	Good	–	–
PCRA	–	–	2	Excellent	–	–
RMS	–	–	1	Good	1	Good
SFS74	–	–	–	–	–	–
SFS76	1	Excellent	–	–	–	–
SFS81	6	Excellent	–	–	–	–
SPIn-W	1	Poor	–	–	–	–
STRONG	–	–	1	Excellent	–	–
WRN	–	–	8	Fair-Good	1	Excellent
WRN-R	–	–	1	Good	–	–

Notes. *k* = number of samples. General Offending = new arrest, charge, conviction, or incarceration; Violations = technical violation, probation revocation, or breach of conditions.

Predictive Validity

Of the 19 instruments reviewed in this guide, no instrument produced more accurate assessments/outcomes in comparison to the others, with predictive validity varying based on the reported performance indicator. Predictive validity is defined as the extent to which predictions derived from the instrument are accurate and/or consistent with other independently validated measures.³⁰ In other words, how well does the selected instrument predict reoffending for specific groups? Predictive validity is categorized as poor, fair, good, and excellent. Poor predictive validity suggested poor performance in accurate risk predictions. Tools coded as excellent demonstrated excellent predictive validity. The following indicators were included in the review: risk classifications, offender subgroups (sex, race/ethnicity, and diagnostic categories), and context (research versus “real world” practice).

Specifically, for general offending *including* violations, predictive validity ranged from poor for SPIn-W assessments to excellent for SFS76 and SFS81 assessments. For general offending *excluding* violations, total scores for over two-thirds of instruments had either good or excellent predictive validity. Moreover, predictive validity ranged from fair for ORAS-PAT assessments to excellent for the ORAS-CST, ORAS-CSST, PCRA, and STRONG assessments (see Table 3).

When taking into account the differences in offenders (gender, race/ethnicity, and mental health), other patterns of predictive validity emerged.

Predictive Validity across Offender Subgroups

Gender. Overall, predictive validity ranged from fair to excellent across gender. Some instruments performed equally well for male and female offenders; for instance, COMPAS assessments demonstrated good predictive validity for both genders; STRONG assessments demonstrated excellent validity for both male and female offenders; and predictive validity for the ORAS instrument for which comparisons were possible—namely, the ORAS-CST, ORAS-CSST, ORAS-PIT, and ORAS-RT—ranged from good to excellent for both groups. Table 4 presents the validity of total scores in predicting recidivism by the offender’s gender.

Other instruments showed differential performance by offender gender. In particular, LSI-R assessments showed good predictive validity for male offenders, but only fair predictive validity for female offenders. Similarly, LSI-R:SV assessments presented only fair predictive validity for female offenders but ranged from fair to good in their predictions for male counterparts.

Few risk tools were evaluated exclusively for male or female offenders (Table 4). Predictive validity of SFS76 and SFS81 assessments were evaluated for male offenders, only; the SFS76 demonstrated excellent validity, while validity of SFS81 assessments ranged from good to excellent. WRN total scores were also evaluated for male offenders and showed fair validity. Designed for women, the SPIn-W showed good validity.

No studies reported predictive validity of assessments by offender gender for the IORNS, ORAS-PAT, PCRA, RMS, SAQ, SFS74, or WRN-R.

Race/ethnicity. Comparisons by

offender race/ethnicity were only possible for assessments completed using the COMPAS and LSI-R. For COMPAS assessments, predictive validity was good for white and black offenders. For LSI-R assessments, predictive validity ranged from poor to good across white, black, Hispanic, and non-white offenders, with performance varying largely depending on sample size and performance indicator rather than race/ethnicity. Together, these findings fail to provide evidence of differential performance of COMPAS and LSI-R assessments as a function of offender race/ethnicity.

Diagnostic categories. No comparisons of predictive validity within or across instruments as a function of mental state, substance use, or personality disorders were possible. Even when these sample characteristics were reported, predictive validity was not provided by subgroup. As for race/ethnicity, there is a critical need for research examining risk assessment accuracy between mentally disordered and non-disordered offenders as well as across diagnostic subgroups. That said, prior meta-analytic work has found the predictors of recidivism to be comparable for mentally disordered offenders,³¹ suggesting that assessments also may perform comparably.

Table 4. Validity of Total Scores in Predicting Recidivism by Offender Sex

INSTRUMENTS	OFFENDER SEX			
	k	Male	k	Female
COMPAS	2	Good	2	Good
LSI-R ^a	9	Good	8	Fair
LSI-R:SV	2	Fair-Good	1	Fair
ORAS-CST	1	Excellent	1	Good
ORAS-CSST	1	Good	1	Excellent
ORAS-PIT	1	Good	1	Good
ORAS-RT	1	Good	1	Excellent
SFS76 ^b	1	Excellent	–	–
SFS81 ^c	–	Good-Excellent	–	–
SPIn-W ^{d,e}	–	–	2	Good
STRONG	1	Excellent	1	Excellent
WRN	1	Fair	–	–

IV. PRAGMATIC CONSIDERATIONS: IMPLEMENTATION AND PRACTICE

A myriad of pragmatic considerations should be taken into account when determining the appropriate risk instrument for a correctional or criminal justice setting. Whether the use of the tool is in a prison or jail, at admission or release, by court or hearing officers or the parole board, or in the field by community probation or parole officers, it is necessary to examine factors such as target populations for assessment, timetables for re/assessment, how scores guide supervision levels, facility resources, personnel, and quality assurance. Furthermore, agencies must query the availability and accessibility of official records and collateral information (e.g., arrest records, institutional misconduct records, etc.), the amount of time required to conduct the assessment, and the capacity and capability of staff to perform the assessment process with fidelity.

Some agencies and institutions have made existing investments to the use of instruments, and so the integration of existing data, systems, and practices becomes another important factor. As a result, organizations may have to allow increased conversion and/or integration time prior to fully implementing the risk assessment tools. In regard to data and systems integration, some risk instruments offer “suites” that facilitate the integration of supplemental products, data, and services. Other tools are stand-alone assessments and do not have the design specifications that allow for integration with existing structures. In addition, if using a proprietary tool, agencies should frequently consult and communicate with the owner/developer to ensure the capability and approval to implement changes.

Moreover, depending on where the tool is being administered, the length of and time required to complete the assessment can be a critical factor, especially when considering use at short-term facilities such as reception locations. Most risk assessments involve a semi-structured interview with an offender, which requires that the offender is physically present (or available via videoconference), coherent, cooperative, and able to understand the nature of the questions being asked. Some tools also use self-assessment, which usually requires literacy equivalent to a sixth-grade reading level.³² Other tools rely solely on information available in the offender’ file.

Accessibility: What information is available?

Length: How much time is required to complete the assessment?

Developers have provided time estimates for the assessment process, yet, these may vary based on setting, conditions, training/familiarity with the tool, and resources (e.g., data). Table 5 shows the characteristics and duration of the assessment process used in the nine

research studies that analyzed administration time for the instruments listed above. But one cautionary note: risk assessments were completed by professionals in forensic services for approximately three-quarters of the studies (82 percent); the remaining assessments were conducted by researchers (15 percent), and two studies were self-administered. The assessments were most often completed in a prison setting (28 percent) or in the community (38 percent), but at times they were conducted in jail (10 percent), a clinic or hospital (4 percent), or at another facility (6 percent). In terms of supervision level, roughly one-third of assessments (36 percent) were performed during community supervision, one-quarter were completed pre-release (26 percent), and the remainder were conducted either prior to incarceration (11 percent) or at admission (10 percent). Data to complete the instruments were captured from file reviews in 24 samples (33 percent), interviews in 12 samples (17 percent), and offender self-report in two samples (3 percent).

Of the tools reviewed in this guide, administration time was reported for only the following five instruments: LSI-R, LSI-R:SV, IORNS, COMPAS, and SAQ. For the LSI-R, administration time ranged from 30 to 60 minutes for assessments conducted in the context of “real world” practice³³ and 45 to 90 minutes in research studies.³⁴ The LSI-R:SV was reported to have a mean administration time of 10 minutes when completed in practice.³⁵ In the same study, the IORNS required 15 minutes to complete; however, this estimate included only the interview portion of the assessment. Across three studies, administration time for the COMPAS varied from 43 to 165 minutes.³⁶ SAQ assessments were reported to take approximately 20 minutes.³⁷ To reiterate, the availability, accessibility, and quality of data and professional capacity and aptitude of the assessor may affect the duration of time to administer the assessment and level of human error.

Table 5. Characteristics of Risk Assessment Instruments

INSTRUMENTS	CHARACTERISTICS				
	<i>k</i>	Items	Intended Population(s)	Intended Outcome(s)	Time (minutes)
COMPAS	3	70	Any Offender	General Offending & Violations	10—60
IORNS	1	130	Any Offender	General Offending & Violations	15—20
LSI-R	25	54	Any Offender	General Offending & Violations	30—40
LSI-R:SV	2	8	Any Offender	General Offending & Violations	10—15
ORAS-PAT	3	7	Any Offender	General Offending	10—15
ORAS-CST	1	35	Any Offender	General Offending	30—45
ORAS-CSST	1	4	Any Offender	General Offending	5—10
ORAS-PIT	1	31	Any Offender	General Offending	Unknown
ORAS-RT	1	20	Any Offender	General Offending	Unknown
PCRA	2	56	Any Offender	General Offending & Violations	15—30
RMS	2	65	Any Offender	General Offending	Unknown
SAQ	2	72	Any Offender	General Offending	15
SFS74	3	9	Parolees	General Offending	Unknown
SFS76	4	7	Parolees	General Offending	Unknown
SFS81	8	6	Parolees	General Offending	Unknown
SPIn-W	2	100	Any Offender	General Offending	Unknown
STRONG ^a	1	26	Any Offender	General Offending	Unknown
WRN	9	53	Any Offender	General Offending	Unknown
WRN-R	1	52	Any Offender	General Offending	Unknown

Notes. k = number of samples; General Offending = new arrest, charge, conviction, or incarceration; Violations = technical violation, probation revocation, or breach of conditions.

^a The STRONG includes three parts: Static Risk Assessment, Offender Needs Assessment, and Offender Supervision Plan; values reflect only the first part, which is the component used to assess risk of recidivism.

V. WHAT ISN'T KNOWN: LIMITATIONS AND FURTHER RESEARCH

This empirical review identified no fewer than 66 risk assessment tools: 19 instruments in broad use to assess the risk of recidivism and at least 47 instruments designed for use in specific jurisdictions. It was supported by an initial and aggressive analysis of studies published in peer-reviewed journals, as well as government reports, doctoral dissertations, and Master's theses. In total, this effort resulted in 173 records that were filtered to a final count of 53 studies (72 samples).

Despite the breadth and depth of this review, there are still important limits to what has been established in research about risk assessment tools. Some of those limitations are provided below along with areas of interest as next steps for research partners and the broader criminal justice and correctional communities, and as suggestions for future research. There is a significant need for independent research and the development of a refined understanding of the importance of risk assessment tools' impact on predicting recidivism for target populations.

The identified areas of concern are:

- **Guiding risk management:** Most studies included in this review reported findings on whether recidivists generally received higher risk estimates than did non-recidivists (known as *discrimination*). Very few studies reported definitive conclusions on whether those offenders who were identified as being at high risk for recidivism went on to recidivate during the specified follow-up period or whether those offenders who were identified as low risk did not recidivate (known as *calibration*). However, the absence of these conclusions is not unique to the studies included in the current review; a recent review found that calibration estimates were reported in less than a fourth of violence risk assessment studies.³⁸ Discrimination and calibration are two sides of the same coin—both representing important qualities of an instrument's predictive validity—but address different issues.³⁹

Though many of the instruments included in the present review were shown to have acceptable levels of predictive validity, the goal of risk assessment is not simply to predict the likelihood of recidivism, but, ultimately, to *reduce* the risk of recidivism by determining intervention targets, appropriate programming level and intensity, and supervision level. To do so, the information derived during the risk assessment process must be used to guide risk management and rehabilitation efforts, with adherence to the RNR model. Specifically, it should assess offenders' risk of recidivism, with more restrictive and intensive efforts focused on high-risk offenders; match treatment and

rehabilitation efforts to offenders' individual criminogenic needs (as identified in the risk assessment process); and deliver services in a manner that is responsive to individual learning styles, motivations, personalities, and strengths.

- **Consistency across assessors:** The research rarely examined inter-rater reliability, defined as the degree to which a rater scores an assessment consistently and the degree of scoring consistency between different raters when assessing the same individual, of available recidivism risk assessment instruments. With the exception of LSI-R and LSI-R:SV, there was no available information regarding whether assessments completed using the reviewed instruments were consistent across assessors/raters. This is not trivial; inter-rater reliability has been referred to as “the most basic requirement for a risk assessment instrument.”⁴⁰ Inter-rater reliability is especially critical on factors that require a judgment call on the part of the assessor. To check inter-rater reliability of assessments, agencies can present the same case to multiple staff members and have those staff score assessments, compare scores, and resolve discrepancies. Inter-rater reliability can be conducted online as well as in person.
- **Further research on race, gender, and mental health:** Few studies examine predictive validity within specific offender subgroups. Only a handful of studies included in this review compared validity by offender gender or race/ethnicity. No study examined predictive validity across psychiatric diagnostic categories. Due to limited empirical support, there is insufficient evidence to conclude that risk instruments perform comparably or are equally applicable to specific offender subgroups. As described earlier, actuarial instruments estimate risk of recidivism through comparison of a given offender's total score against the recidivism rates of offenders with the same (or a similar) score in the *construction sample* (the individuals studied to determine the relationship between the risk variables and likelihood of recidivism). Race/ethnicity and gender are important factors associated with recidivism that may not be accounted for in these actuarial models. There is considerable evidence to suggest that race/ethnicity and gender are potentially important sources of assessment bias.⁴¹
- **Independent research:** Nearly a third of the studies included in this review were conducted by the developer/owner of the instrument under investigation. To note, performance, or accuracy in predicting recidivism, of the instruments was significantly better in studies conducted by the tool developer than in studies conducted by independent researchers. *Allegiance effects* (the likelihood that the developer of a tool or treatment will find more positive results than independent researchers) were unable to be tested due to the relatively small number of studies per instrument. Though the

reasons for allegiance effects are unclear,^{*} there is a gap as far as independent evaluation of the predictive validity of risk assessments.

- **Increased understanding of violations:** More than two-thirds of studies used a prospective study design, which is an optimal approach for examining predictive validity, and the average length of follow-up was approximately two years (23.5 months). Studies were most frequently conducted in midwestern states followed by the southwestern and northeastern regions of the U.S. Nearly 70 percent of the studies examined general recidivism as the outcome; roughly a quarter considered a variety of outcomes, with few focusing specifically on violations of conditions. As a result, the knowledge of the validity of recidivism risk assessment instruments in predicting violations as opposed to other forms of recidivism is limited.

^{*} E.g., bias, fidelity. See Grant T. Harris, Marnie E. Rice, and Vernon L. Quinsey, "Allegiance or Fidelity? A Clarifying Reply," *Clinical Psychology: Science and Practice* 17, no. 1 (2010): 82–89.

CONCLUSION

In working to identify the most appropriate risk assessment instrument in practice, consider these central/primary questions:

What is the outcome of interest?

This review revealed that some instruments performed better than others in predicting particular recidivism outcomes. Specifically, the SFS instruments performed particularly well in predicting general offending *including* violations, whereas the ORAS-CST, ORAS-CSST, PCRA, and STRONG were excellent predictors of general offending *excluding* violations. WRN assessments fared the best as predictors of exclusively violations.

Who is the target population for assessment?

Certain instruments were developed to assess specific populations; for example, the SFS instruments are designed for use with parolees. Also, several instruments appear to perform better for particular subgroups of offenders than for other subgroups of offenders. The LSI instruments, for instance, produced assessments with only fair predictive validity for female offenders, though predictive validity was generally good for male offenders. Other instruments, such as the COMPAS, ORAS, and STRONG, produced assessments with good validity for both male and female offender populations.

What information is available, how should the assessment be conducted (semi-structured interview, record reviews), and what is the estimated length of time required to complete the assessment process?

Instruments such as the IORNS are completed based solely on offender self-report; other instruments, such as the PCRA and COMPAS, combine information derived from a number of sources, including self-report, interview, and review of official records and collateral information. And some instruments, such as the ORAS, can be completed exclusively based on file review. Similarly, the time required to complete a risk assessment will depend not only on the nature and amount of information required but also on the number of items included. The number of items varied broadly across instruments from 4 items (ORAS-CSST) to 130 items (IORNS).

Decision makers should consider whether the amount of required time and information necessary to complete the assessment will be available on a consistent basis for those who will be conducting and managing the intake and/or assessment process. Staff should engage in a

continual review of the pertinent literature in the field and receive refresher training courses on assessment administration, especially those works on risk assessment. Organizations should consider the methodology of the study, sample size, strength of the empirical support for inter-rater reliability and predictive validity, generalizability of findings, and possible sources of bias. It is recommended that decision makers work closely with researchers to ensure that the measures and scoring threshold for the tool is appropriate for the intended population.

Finally, it is important to remember that the goal of risk assessment is not simply to predict the likelihood of recidivism, but, ultimately, to reduce the risk of recidivism. To do so, the data derived during the risk assessment process must be used to inform risk management, treatment planning, and rehabilitation efforts. Risk assessment results should guide offender classification and staff decisions for client treatment by identifying key responsibility factors and offender characteristics.

Lastly, with increased empirical evidence on the proper use and limitations of risk assessments and the appeal for implementation of evidence-based practices that respond to the needs of offenders, agencies will have to consider organizational readiness to change and engage and educate stakeholders on RNR principles. Adopting a risk assessment instrument should contribute to a shift of culture and alter the approach to classification, treatment, reentry, and supervision. To maximize the adoption process, increase departmental effectiveness, produce fiscal efficiencies, and improve offender outcomes and produce safer communities, agencies should develop strategic messaging and carefully deploy the adoption and use of the risk assessment tool. The planning process will require systems to assess areas of strength, identify existing barriers/gaps, and finally commit leadership and stakeholders to build support for evidenced-based practices to improve public safety.

APPENDIX: SUMMARY OF FINDINGS BY INSTRUMENT

In this section, each risk assessment instrument is described and findings of U.S. studies examining predictive validity are summarized. Instruments are presented in alphabetical order.

Correctional Offender Management Profiling for Alternative Sanctions

Description

The Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is an actuarial risk assessment instrument intended to assess risk for general offending and technical violations across offender populations (Brennan, Dieterich, and Ehret 2009).

The COMPAS contains static and dynamic risk factors. Content areas assessed include attitudes, associates or peers, history of antisocial behavior, personality problems, circumstances at school or work, leisure or recreational activities, substance use problems, mental health problems, and housing, divided across 22 scales (Blomberg et al. 2010). Scores on the self-report assessment, data from official records, and information from interview are used to arrive at an overall risk score for each offender.

COMPAS assessments are completed through a combination of a computer-assisted self-report questionnaire, an interview conducted by a trained assessor, and data collected from the offender's records. The instrument can be purchased from Northpointe at northpointeinc.com. Assessors must undergo a two-day training session that covers practical use, interpretation of results, and case-planning strategies in order to administer the COMPAS. Advanced training options that focus on the theoretical underpinnings of offender assessments, gender responsiveness, motivational interviewing, and other topics are available.

U.S. Research Evidence

In total, four studies have evaluated predictive validity of COMPAS assessments in U.S. samples. Blomberg and colleagues (2010) found that individuals identified as higher risk were indeed more likely to recidivate; specifically, 7 percent of those identified to be low risk recidivated, 16 percent of those identified as medium risk recidivated, and 27 percent of those identified as high risk recidivated. In other samples, predictive validity was determined to be good for general offending (Brennan, Dieterich, and Ehret 2009) and fair for violations (Farabee and Zhang 2007). Predictive validity for male and female offenders has ranged from good to excellent (Brennan, Dieterich, and Ehret 2009).

There have been no comparisons of predictive validity in U.S. samples between total scores and risk classifications, assessments completed in research and practice contexts, or by offender race/ethnicity. There also have not been any U.S. evaluations of inter-rater reliability.

Practical Issues and Considerations

For the self-report portion of the assessment, the computer upon which the offender completes the questionnaire must have Internet access and run on Windows. The assessor must receive training to be qualified to administer the structured interview.

Selected References and Suggested Readings

Thomas Blomberg et al., *Validation of the COMPAS Risk Assessment Classification Instrument* (Tallahassee, FL: Florida State University Center for Criminology and Public Policy Research, 2010), available at criminologycenter.fsu.edu/p/pdf/pretrial/Broward%20Co.%20COMPAS%20Validation%202010.pdf.

Tim Brennan, William Dieterich, and Beate Ehret, "Evaluating the Predictive Validity of the COMPAS Risk and Needs Assessment System," *Criminal Justice and Behavior: An International Journal* 36, no. 1 (2009): 21–40.

David Farabee and Sheldon Zhang, *COMPAS Validation Study: First Annual Report* (Los Angeles: California Department of Corrections and Rehabilitation, 2007), available at cdcr.ca.gov/Adult_Research_Branch/Research_Documents/COMPAS_Validation_Dec_2007.pdf.

David Farabee et al., *COMPAS Validation Study: Final Report* (Los Angeles: California Department of Corrections and Rehabilitation, 2010), available at cdcr.ca.gov/Adult_Research_Branch/Research_Documents/COMPAS_Final_Report_08-11-10.pdf.

Federal Post Conviction Risk Assessment

Description

The Federal Post Conviction Risk Assessment (PCRA) is an actuarial risk assessment instrument intended to assess risk for general offending and technical violations across offender populations (Johnson et al. 2011).

The PCRA contains 56 static and dynamic risk factors. Content areas assessed include attitudes, associates or peers, history of antisocial behavior, relationships, circumstances at work or school, and substance use problems. Self-report assessment scores are combined with probation officer assessment scores to arrive at an overall risk score.

PCRA assessments are composed of two components: 1) the Officer Assessment, and 2) Offender Self-Assessment. The self-report questionnaire consists of items that are “scored” and “unscored.” The 15 scored items are those that have been shown in studies conducted by the Administrative Office of U.S. Courts (Administrative Office) to predict recidivism and contribute to the overall risk score. The 41 unscored items have been shown in other research to predict recidivism, but have not been evaluated by the Administrative Office. They are included to inform intervention strategies, but do not contribute to the risk scores. Assessments must be administered by probation officers who have passed the online certification test created and offered by the Administrative Office; the Administrative Office prohibits uncertified assessors from accessing the PCRA. Prior to the online certification, probation officers must complete 16 hours of training. They also must renew their certification every year. The PCRA is available through the Administrative Office at uscourts.gov.

U.S. Research Evidence

One study has assessed the predictive validity of PCRA assessments in two large U.S. samples. Johnson et al. (2011) found excellent predictive validity in both. There have been no comparisons of predictive validity between assessments completed in research and practice contexts, by offender gender, or by offender race/ethnicity. There also have not been any U.S. evaluations of inter-rater reliability.

Practical Issues and Considerations

Overall, research evidence is limited. There have been no evaluations of the reliability and predictive validity of PCRA assessments beyond the initial construction and validation study. Whether findings generalize to other samples is unknown. Independent replication is needed.

Selected References and Suggested Readings

Administrative Office of the United States Courts, Office of Probation and Pretrial Services, *An Overview of the Federal Post Conviction Risk Assessment* (Washington, DC: Administrative Office of the United States Courts, 2011), [available at uscourts.gov/uscourts/FederalCourts/PPS/PCRA_Sep_2011.pdf](https://uscourts.gov/uscourts/FederalCourts/PPS/PCRA_Sep_2011.pdf).

James L. Johnson et al., "The Construction and Validation of the Federal Post Conviction Risk Assessment (PCRA)," *Federal Probation* 75, no. 2 (2011): 16–29.

Inventory of Offender Risk, Needs, and Strengths

Description

The Inventory of Offender Risk, Needs, and Strengths (IORNS) is an actuarial risk assessment instrument intended to assess risk for general offending and technical violations across offender populations (Miller 2006a).

The IORNS contains 130 static, dynamic, risk, and protective factors. Content areas assessed include attitudes, associates or peers, history of antisocial behavior, personality problems, relationships, circumstances at school or work, substance use problems, mental health problems, and housing. Individual item responses are summed to create Static, Dynamic and Protective indexes as well as an Overall Risk index. There also are two validity scales.

The IORNS is a true/false self-report questionnaire completed by the offender and requires a third-grade reading level. The IORNS manual indicates that assessments take 15 to 20 minutes to administer and 20 to 25 minutes to score. There are no training requirements for assessors, provided the purchaser of the exam has a degree in forensic or clinical psychology or psychiatry as well as certification in psychological testing. The purchaser is also responsible for overseeing the scoring of the assessment. IORNS assessments are available through Psychological Assessment Resources (parinc.com). Costs include those associated with the manual, interview guides, and assessment forms. For further information on pricing, see parinc.com.

U.S. Research Evidence

Predictive validity of IORNS assessments have been evaluated in only one U.S. sample conducted by the author of the instrument. Miller (2006b) found that offenders with higher

Overall Risk Indices were in jail more frequently and had more non-violent arrests than those with lower scores. Similarly, those offenders who had more halfway house rule violations have significantly lower Overall Risk and Dynamic Needs Indices.

There have been no comparisons of predictive validity in U.S. samples between assessments completed in research and practice contexts, by recidivism outcome, offender gender, or offender race/ethnicity. There also have not been any U.S. evaluations of inter-rater reliability.

Practical Issues and Considerations

Though findings are promising, predictive validity of IORNS assessments has only been evaluated in one study conducted by the instrument developer; independent replication is needed.

Selected References and Suggested Readings

Holly A. Miller, *Inventory of Offender Risk, Needs, and Strengths (IORNS)* (Odessa, FL: Psychological Assessment Resources, 2006a).

Holly A. Miller, "A Dynamic Assessment of Offender Risk, Needs, and Strengths in a Sample of Pre-Release General Offenders," *Behavioral Sciences & the Law* 24, no. 6 (2006b): 767–782.

Level of Service Instruments

Description

The Level of Service family of instruments includes the Level of Service Inventory-Revised (LSI-R) and the Level of Service Inventory-Revised: Screening Version (LSI-R:SV), which are actuarial risk assessment instruments intended to assess risk for general offending and technical violations across offender populations (Andrews and Bonta 1995, 1999).

The LSI-R contains 54 static and dynamic risk factors. Content areas include attitudes, associates or peers, history of antisocial behavior, personality problems, relationships, circumstances at school or work, leisure or recreational activities, substance use problems, mental health problems, and housing. Item responses are scored and summed for a total score from 0 to 54.

The LSI-R:SV contains eight static and dynamic items selected from the LSI-R. Content areas assessed include attitudes, associates or peers, history of antisocial behavior, personality

problems, relationships, circumstances at school or work, and substance abuse problems. Individual item responses are scored and summed for a total score ranging from 0 to 9. This score is used to determine if the offender requires a full LSI-R assessment.

LSI-R and LSI-R:SV assessments are completed through interview and file review, a process estimated to require approximately 30 to 40 minutes for the LSI-R and 10 to 15 minutes for the LSI-R:SV (though studies reported longer completion times—see below). The assessor does not need formal training, but scoring must be overseen by someone who has at least B-level qualifications (i.e., post-secondary training in psychological assessment). The LSI-R and LSI-R:SV materials are available through Multi-Health Systems (mhs.com). Costs include those associated with the manual, interview guides, and assessment forms. For further information on pricing, see mhs.com.

U.S. Research Evidence

Predictive validity of LSI-R total scores has been evaluated in 25 U.S. samples. Performance has ranged from poor to good, with the median on the cusp of fair and good. No studies have examined the predictive validity of the risk classifications (as opposed to total scores). LSI-R total scores seem to perform slightly better for men than for women, though performance is in the fair-good range for both. U.S. studies have not shown differences in validity as a function of race/ethnicity. Predictive validity for total scores completed in the context of research and practice is also comparable. Validity in predicting general offending is slightly better than for technical violations. In the one U.S. study reporting inter-rater reliability data, agreement ranged from poor to excellent across content domains, but was excellent overall (Simourd 2006).

Predictive validity of the LSI-R:SV has only been examined in two U.S. samples with mixed results: one study showed fair performance (Walters 2011) and the other showed good performance (Lowenkamp, Lovins, and Latessa 2009). The LSI-R:SV seems to perform better for men (good predictive validity) than for women (fair predictive validity). There have been no comparisons of predictive validity between total scores and risk classifications, assessments completed in research and practice, by offender race/ethnicity, or by recidivism outcome. Because the LSI-R:SV is a self-report instrument, inter-reliability is not relevant.

Practical Issues and Considerations

Researchers and professionals have reported administration times that deviate from the LSI-R manual's estimate of 30 to 40 minutes, including an average completion time of 60 minutes in

one sample (Holsinger et al. 2004) and ranges from 45 to 90 minutes in two others (Evans 2009; Lowenkamp, Lovins, and Latessa 2009).

There is considerable variation in the cut-off scores used for the risk categories. The manual encourages altering cut-off scores based on the characteristics of offense groups, but research should be conducted prior to implementation to examine the predictive validity of revised cut-off scores (Kim 2010).

A recent addition to the Level of Service family of instruments is the Level of Service/Case Management Inventory (LS/CMI), an actuarial risk assessment with 43 items intended to aid professionals in late adolescent and adult offender management. No studies examining the LS/CMI met inclusion criteria for this review.

Selected References and Suggested Readings

Donald A. Andrews and James L. Bonta, *LSI-R: The Level of Service Inventory-Revised* (Toronto: Multi-Health Systems, 1995).

Donald A. Andrews and James L. Bonta, *LSI-R: SV Level of Service Inventory Revised: Screening Version User's Manual* (Toronto: Multi-Health Systems, 1999).

Stephanie A. Evans, "Gender Disparity in the Prediction of Recidivism: The Accuracy of the LSI-R Modified," (master's thesis, University of Alabama, 2009).

Alexander M. Holsinger, Christopher T. Lowenkamp, and Edward J. Latessa, "Validating the LSI-R on a Sample of Jail Inmates," *Journal of Offender Monitoring* 17, no. 1 (2004): 8–9.

Hye-Sun Kim, "Prisoner Classification Re-visited: A Further Test of the Level of Service Inventory-Revised (LSI-R) Intake Assessment." (PhD diss., Indiana University of Pennsylvania, 2010).

Christopher T. Lowenkamp, Brian Lovins, and Edward J. Latessa, "Validating the Level of Service Inventory—Revised and the Level of Service Inventory: Screening Version With a Sample of Probationers," *Prison Journal* 89, no. 2 (2009): 192–204.

David Simourd, *Validation of Risk/Needs Assessments in the Pennsylvania Department of Corrections* (Mechanicsburg, PA: Pennsylvania Department of Corrections, 2006).

Glenn D. Walters, "Predicting Recidivism with the Psychological Inventory of Criminal Thinking Styles and Level of Service Inventory–Revised: Screening Version," *Law and Human Behavior* 35, no. 3 (2011): 211–220.

Ohio Risk Assessment System

Description

The Ohio Risk Assessment System (ORAS) is composed of five actuarial risk assessment instruments intended to assess risk for recidivism across offender populations (Latessa et al. 2009): the 7-item Pretrial Assessment Tool (ORAS-PAT), the 4-item Community Supervision Screening Tool (ORAS-CSST), the 35-item Community Supervision Tool (ORAS-CST), the 31-item Prison Intake Tool (ORAS-PIT), and the 20-item Prison Reentry Tool (ORAS-RT). Each includes static and dynamic risk factors and is designed for use at a specific stage in the criminal justice system. Assessments identify criminogenic needs and place offenders into risk categories. An additional sixth instrument, the Prison Screening Tool (ORAS-PST), is designed to identify low-risk inmates who do not need the full ORAS-PIT assessment.

Item responses are scored and summed to create total scores that are compared against risk classification cut-off values. The ORAS-PAT has a range from 0 to 9, the ORAS-CSST from 0 to 7, the ORAS-CST from 0 to 49, the ORAS-PIT from 3 to 29, and the ORAS-RT from 0 to 28. Each tool considers the offender's history of antisocial behavior, circumstances at school or work, and substance abuse problems; some also evaluate additional domains, such as attitudes (e.g., ORAS-CST, ORAS-RT), and mental health problems (e.g., ORAS-PIT, ORAS-RT).

The ORAS tools are completed through a structured interview and analysis of official records; the ORAS-CSST, ORAS-PIT, and ORAS-RT additionally use self-report questionnaires. Assessors must complete a two-day training package that accompanies the tool prior to administering any assessments. The ORAS is published by the Ohio Department of Rehabilitation and Correction (drc.ohio.gov). The system is non-proprietary and can be obtained from the Center of Criminal Justice Research at the University of Cincinnati (uc.edu/corrections/services/risk-assessment.html).

U.S. Research Evidence

ORAS-PAT total scores demonstrated fair validity in predicting arrest in the construction sample and good validity in the validation sample (Latessa et al. 2009). A second evaluation found fair predictive validity for ORAS-PAT assessments, good validity for ORAS-PIT and ORAS-RT

assessments, and excellent validity for ORAS-CCST and ORAS-CST assessments (Lowenkamp, Lemke, and Latessa 2008). ORAS-PST assessments have not been included in these evaluations.

Predictive validity of ORAS assessments differs somewhat in relation to offender gender. Specifically, ORAS-CST assessments performed slightly better for male than female offenders, though predictive validity was excellent in both cases. Conversely, ORAS-PIT and ORAS-RT assessments performed better for female (excellent predictive validity) than male offenders (good). ORAS-CSST assessments, in contrast, have shown comparable predictive validity for both male and female offenders. The ORAS-PAT total scores have demonstrated better validity in predicting technical violations (good) than general offending (fair).

There have been no comparisons of predictive validity between total scores and risk classifications, between assessments completed in research and practice contexts, or by offender race/ethnicity. There also have been no evaluations of inter-rater reliability.

Practical Issues and Considerations

Though findings are promising, there has been relatively little research on the predictive validity of the ORAS, with only one evaluation of four of the tools in the suite and two of the fifth tool. What's more, there have been no reports on the inter-rater reliability of the assessments. Finally, all extant research has been completed by the study developers; independent replication is needed.

Selected References and Suggested Readings

Edward Latessa et al., *Creation and Validation of the Ohio Risk Assessment System: Final Report* (Cincinnati, OH: University of Cincinnati Center for Criminal Justice Research, 2009), available at ocjs.ohio.gov/ORAS_FinalReport.pdf.

Christopher T. Lowenkamp, Richard Lemke, and Edward Latessa, "The Development and Validation of a Pretrial Screening Tool," *Federal Probation* 72, no. 3 (2008): 2–9.

Risk Management Systems

Description

The Risk Management Systems (RMS) is an actuarial risk assessment instrument intended to be used to assess risk for general offending across offender populations (Dow, Jones, and Mott

2005). The RMS currently contains 67 static and dynamic risk factors; however, when it was validated, the instrument included only 65 items. The assessment is split into four parts: 1) Needs (24 items), 2) Risk (9 items), 3) Mental Health (10 items), and 4) Other-External (24 items). Content areas assessed include attitudes, associates or peers, history of antisocial behavior, personality problems, relationships, circumstances at school or work, substance abuse problems, mental health problems, and housing.

The RMS is administered using a computer-based questionnaire. As such, the assessor is removed from the initial assessment process; individual item responses are statistically analyzed to calculate risk of recidivism. Risk scores for violence and recidivism range from 1.00 (Low) to 2.00 (High), at 0.01 intervals. However, there are no established cut-off scores for risk categories, so the assessor must interpret the subsequent level of risk/supervision required. RMS assessment materials can be purchased through Syscon Justice Systems (syscon.net). For information on pricing see syscon.net.

U.S. Research Evidence

Predictive validity of RMS assessment has been reported in two U.S. studies; performance has ranged from good (Kelly 2009) to excellent (Dow, Jones, and Mott 2005). The risk classifications have notably better predictive validity (excellent) compared to total scores (good). Validity is comparable for predicting general offending and technical violations. RMS assessments appear to have better predictive validity when completed in research studies (excellent) than in the context of “real world” practice (good); however, this comparison is confounded by use of risk classifications in one study and total scores in the other.

There have been no U.S. comparisons of predictive validity by offender gender or race/ethnicity. There also have not been any U.S. evaluations of inter-rater reliability.

Practical Issues and Considerations

The initial development and validation study intended the tool to be used to assess risk for general offending (Dow, Jones, and Mott 2005); a later study established the validity of RMS assessments in predicting technical violations (Kelly 2009). Overall, further independent research is needed to replicate and establish the generalizability of findings, as well as to determine the validity of different cut-off scores.

Selected References and Suggested Readings

Edward Dow, Charles Jones, and Jack Mott, "An Empirical Modeling Approach to Recidivism Classification," *Criminal Justice and Behavior* 32, no. 2 (2005): 223–247.

Bridget Kelly, "A Validation Study of Risk Management Systems," (master's thesis, University of Nevada, Las Vegas, 2009, available at digitalscholarship.unlv.edu/thesesdissertations/128).

Deborah Koetzle Shaffer, Bridget Kelly, and Joel D. Lieberman, "An Exemplar-Based Approach to Risk Assessment: Validating the Risk Management Systems Instrument," *Criminal Justice Policy Review* 22, no. 2 (2011): 167–186.

Salient Factor Score

Description

The Salient Factor Score (SFS) is an actuarial risk assessment tool intended to assess risk for general offending across offender populations. This tool is specifically designed to determine whether an offender should be granted parole or not.

There are at least four versions of the SFS. Items have been adapted over the years to be consistent with research findings; however, each version of the instrument only measures static risk factors. The SFS74 contains nine items, and content areas include history of antisocial behavior, circumstances at work or school, substance use problems, and housing. The SFS76 contains seven items and content areas include history of antisocial behavior, circumstances at work or school, and substance use problems. The SFS81 contains six items, and content areas include history of antisocial behavior and substance use problems. The SFS98 includes six items, and the only content area included is history of antisocial behavior. Unlike the prior versions, the SFS98 also considers whether the offender was older than 41 at the time of the current offense.

The SFS instruments are completed through review of official records. Item ratings are summed to arrive at an overall risk score, with a *higher* score indicating *lower* risk. These total scores are then used to place offenders within one of four risk categories: very good risk, good risk, fair risk, and poor risk. For further information contact the United States Parole Commission (justice.gov/uspc).

U.S. Research Evidence

Predictive validity SFS74, SFS76, and the SFS81 assessments have been examined in 15 U.S. samples. Validity of SFS74 and SFS76 assessments in predicting general offending has ranged from good to excellent. SFS81 assessments generally have also shown excellent predictive validity across studies, though the odds ratio was notably low in one evaluation (Howard 2007). No U.S. evaluations of the predictive validity of SFS98 assessments were discovered.

There have been no U.S. comparisons of the predictive validity of the SFS instruments between total scores and risk classifications, between assessments completed in research and practice contexts, or by offender race/ethnicity. There also have been no evaluations of inter-rater reliability.

Practical Issues and Considerations

Though items are relatively straightforward to code, investigations of inter-rater reliability are needed to establish the consistency of assessments across assessors.

Jurisdiction-specific adaptations include the Connecticut Salient Factor Score.

Selected References and Suggested Readings

Peter B. Hoffman, "Twenty Years of Operational Use of a Risk Prediction Instrument: The United States Parole Commission's Salient Factor Score," *Journal of Criminal Justice* 22, no. 6, (1994): 477–494.

Peter B. Hoffman and Sheldon Adelberg, "The Salient Factor Score: A Nontechnical Overview," *Federal Probation* 44, no. 1 (1980): 44–52.

Brooke K. Howard, "Examining Predictive Validity of the Salient Factor Score and HCR-20 Among Behavior Health Court Clientele: Comparing Static and Dynamic Variables," (PhD diss., Pacific Graduate School, 2007).

Self-Appraisal Questionnaire

The Self-Appraisal Questionnaire (SAQ) is an actuarial risk assessment instrument to assess risk for general offending among male offenders (Kingston, 2005).

The SAQ contains 72 dynamic and static risk factors. Content areas include attitudes, associates or peers, history of antisocial behavior, personality problems, and substance abuse problems.

Items are divided across seven subscales. Scores on six subscales are calculated to provide an overall risk score. A seventh anger subscale is not used to assess risk for recidivism. Therefore, of the 72 total items, 67 items are used to predict recidivism. Total scores are used to place offenders in one of four risk categories: low, low-moderate, high-moderate, and high.

The SAQ is a true/false self-report questionnaire. There are five items that can be used to assess the validity of an offender's answers by comparing them against official records. The SAQ takes approximately 15 minutes to administer and 5 minutes to hand score. The assessor does not need formal training, but scoring must be overseen by someone who has at least B-level qualifications (i.e., post-secondary training in psychological assessment). The SAQ can be purchased from Multi-Health Systems Inc. at mhs.com. Costs include those associated with the manual and assessment forms. For further information on pricing, see mhs.com.

U.S. Research Evidence

Two studies have evaluated the predictive validity of the SAQ in U.S. samples. These studies used low-, moderate-, and high-risk categories rather than the four categories suggested by the assessment developer. Mitchell and Mackenzie (2006) found poor validity of the SAQ assessments in predicting rearrest and failed to find differences in total scores between recidivists and non-recidivists. In contrast, using a longer follow-up period and a larger sample, Mitchell, Caudy, and Mackenzie (2013) found that SAQ assessments predicted time to first reconviction, though the effect size was small.

There have been no comparisons of predictive validity in U.S. samples between total scores and risk classifications, between assessments completed in research and practice, by offender gender, or by race/ethnicity. Because the SAQ is a self-report instrument, inter-reliability is not relevant.

Practical Issues and Considerations

The SAQ requires a fifth-grade reading level. Prior studies of the validity of SAQ assessments in predicting violent outcomes, including institutional violence and violent recidivism (e.g., Campbell, French, and Gendreau 2009), as well as violent and non-violent recidivism in Canadian samples (e.g., Kingston, MacTavish, and Loza-Fanous 2007) have shown more promising results than those reported herein vis-à-vis validity in predicting non-violent offending in U.S. samples.

Selected References and Suggested Readings

Mary A. Campbell, Sheila French, and Paul Gendreau, "The Prediction of Violence in Adult Offenders: A Meta-Analytic Comparison of Instruments and Methods of Assessment," *Criminal Justice and Behavior* 36, no. 6 (2009): 567–590.

Wagdy Loza Kingston, *The Self-Appraisal Questionnaire (SAQ): A Tool for Assessing Violent and Non-Violent Recidivism* (Toronto: Mental Health Systems, 2005).

Wagdy Loza Kingston, Angèle MacTavish, and Amel Loza-Fanous, "A Nine-Year Follow-Up Study on the Predictive Validity of the Self-Appraisal Questionnaire for Predicting Violent and Nonviolent Recidivism," *J Interpers Violence* 2007 22, no. 9 (2007): 1144–55.

Ojmarrh Mitchell, Michael S. Caudy, and Doris Layton Mackenzie. "A Reanalysis of the Self-Appraisal Questionnaire: Psychometric Properties and Predictive Validity," *International Journal of Offender Therapy and Comparative Criminology* 57, no. 4 (2013): 445–459.

Ojmarrh Mitchell and Doris Layton Mackenzie, "Disconfirmation of the Predictive Validity of the Self-Appraisal Questionnaire in a Sample of High-Risk Drug Offenders," *Criminal Justice and Behavior* 33, no. 4 (2006): 449–466.

Service Planning Instruments

Description

The Service Planning Instrument (SPIn) is an actuarial risk assessment tool intended to assess risk for offending and to identify the service needs of male offenders. The SPIn-W was developed for use with female offenders.

Both the SPIn and SPIn-W are self-report, computer-based instruments. The SPIn includes 90 static, dynamic, risk, and protective factors. Content areas assessed include attitudes, associates or peers, history of antisocial behavior, relationships, circumstances at school or work, substance use problems, mental health problems, and housing. The SPIn-W includes 100 static, dynamic, risk, and protective factors. Content areas include attitudes, associates or peers, history of antisocial behavior, relationships, circumstances at school or work, leisure or recreational activities, substance use problems, mental health problems, and housing.

For both instruments, software is used to calculate an offender's risk score, which is presented graphically and narratively. The assessor must compare responses on static items to the offender's official records. Assessors are required to attend a two-day training session.

Additional two-day training programs to help administrators better prepare for the case-planning process, as well as data workshops, refresher courses, technical support, and quality assurance are also available. The SPIn and SPIn-W can be purchased from Orbis Partners Inc. (orbispartners.com). For information on pricing, see orbispartners.com.

U.S. Research Evidence

Predictive validity of SPIn assessments has not been assessed in U.S. samples. Two studies have evaluated predictive validity of the SPIn-W assessments; performance ranged from poor to excellent.

There have been no comparisons of predictive validity in U.S. samples between total scores and risk classifications, between assessments completed in research and practice contexts, by outcome, or by offender race/ethnicity. There also have been no U.S. evaluations of inter-rater reliability.

Practical Issues and Considerations

Overall, evidence regarding the predictive validity of SPIn-W assessments is both limited and mixed. There is no evidence to support the predictive validity of SPIn assessments, nor inter-rater reliability for either the SPIn or SPIn-W. More research is needed.

Selected References and Suggested Readings

Cathleen Meaden, "The Utility of the Level of Service Inventory-Revised Versus the Service Planning Instrument for Women in Predicting Program Completion in Female Offenders," (master's thesis, Central Connecticut State University, 2012), available at content.library.ccsu.edu/cdm/ref/collection/ccsutheses/id/1771.

Bart Millson, David Robinson, and Marilyn Van Dieten, *Women Offender Case Management Model* (Washington, DC: U.S. Department of Justice, National Institute of Corrections, 2010), available at cjinvolvementwomen.org/sites/all/documents/Women%20Offender%20Case%20Management%20Model.pdf.

Static Risk and Offender Needs Guide

The Static Risk and Offender Needs Guide (STRONG) is an actuarial risk assessment instrument intended to assess risk for general offending across offender populations (Barnoski and Drake 2007).

The STRONG consists of three parts: the Static Risk Assessment, which contains 26 static risk factors; the Offender Needs Assessment, which contains 70 dynamic risk and protective factors; and the Offender Supervision Plan, which is auto-populated based on the results of the Offender Needs Assessment. Content areas assessed in the Static Risk Assessment include history of antisocial behavior and substance use problems. Items scores are used to create three separate scores: Felony Risk Score, Non-Violent Felony Risk Score (high property risk/high drug risk), and Violent Felony Risk Score. These three scores are then used to classify offenders in one of five categories: high-risk violent; high-risk property; high-risk drug; moderate risk; and low risk. Content areas assessed in the Offender Needs Assessment include attitudes, associates or peers, personality problems, relationships, circumstances at work or school, substance use problems, mental health problems, and housing. Ratings on items included in the Offender Needs Assessment are not used to inform risk assessments but instead guide the development of interventions designed to reduce risk of future criminal justice involvement.

STRONG assessments are completed by assessors using a web-based interface. Assessors must complete an initial training program as well as routine booster training sessions. The STRONG was developed by Assessments.com in collaboration with the Washington Department of Corrections. A very similar version can be purchased for use in other jurisdictions through assessments.com.

U.S. Research Evidence

Only one U.S. study has evaluated the predictive validity of STRONG assessments; assessments demonstrated excellent predictive validity overall, as well as for male and female offenders separately (Barnoski and Drake 2007). Predictive validity has not been examined as a function of offender race/ethnicity, type of recidivism outcome, or between assessments completed in the context of research versus practice. There have been no evaluations of inter-rater reliability.

Practical Issues and Considerations

Though findings are promising, predictive validity of STRONG assessments has only been evaluated in one study conducted by the instrument developer; independent replication is needed.

Selected References and Suggested Readings

Robert Barnoski and Elizabeth K. Drake, *Washington's Offender Accountability Act: Department of Corrections' Static Risk Instrument* (Olympia, WA: Washington State Institute for Public Policy, 2007), available at wsipp.wa.gov/rptfiles/07-03-1201R.pdf.

Wisconsin Risk and Needs Scales

Description

The Wisconsin Risk and Needs Scales (WRN) is an actuarial risk assessment instrument intended to assess risk for general offending and technical violations across offender populations. A revised version (WRN-R) was designed specifically for use with probationers and parolees (Eisenberg, Bryl, and Fabelo 2009).

The WRN contains 53 static and dynamic risk factors. Content areas assessed include attitudes, associates or peers, history of antisocial behavior, relationships, circumstances at work or school, substance use problems, and mental health problems. Individual item scores are scored and summed for a total risk score ranging from 0 to 52. The total score is used to place the offender in a risk category based on predetermined cut-offs: Low = 0 to 7; Medium = 8 to 14; and High = 15+.

The WRN-R retained 52 of the WRN's items and covers the same content areas. The weights of the different factors have been revised from the original WRN based on the results of a validation study, and the revised total risk score has a range of 0 to 25. The total score is used to estimate risk level based on new cut-offs: Low = 0 to 8; Medium = 9 to 14; and High = 15+.

WRN assessments are completed using information obtained through interview. The WRN is non-proprietary and available at no cost through Justice Systems Assessment & Training (j-satresources.com/Toolkit/Adult/adf6e846-f4dc-4b1e-b7b1-2ff28551ce85).

U.S. Research Evidence

Predictive validity of the WRN assessments have ranged from fair (Eisenberg, Bryl, and Fabelo 2009) to excellent (Connolly 2003). WRN assessments appear to perform better for predictive technical violations (excellent) than general offending (good). Comparisons between predictive validity of assessments completed in research versus practice failed to identify any differences. There have been no comparisons of predictive validity in U.S. samples between total scores and

risk classifications, or by offender gender or race/ethnicity. There have also been no U.S. evaluations of inter-rater reliability.

To date, predictive validity of the WRN-R has been evaluated in one U.S. study; assessments demonstrated good predictive validity.

Practical Issues and Considerations

A high percentage of offenders are classified as high risk using the WRN due to the heavy weight given to convictions for an assaultive offense in the past five years. There is concern that such overclassification is “counter to the goal of risk classification: to differentiate the population by risk and allocate resources accordingly” (Eisenberg, Bryl, and Fabelo 2009, iv).

In 2004, a new, automated assessment and case management system called the Correctional Assessment and Intervention System (CAIS) was developed based upon the WRN and the Client Management Classification tools (Baird, Heinz, and Bemus 1979). This CAIS is an actuarial risk assessment instrument intended to assess risk for general offending and technical violations across offender populations, as well as to be used in the development of case management plans. Its predictive validity has not yet been evaluated.

Selected References and Suggested Readings

Christopher Baird, Richard C. Heinz, and Brian J. Bemus, *The Wisconsin Case Classification/Staff Deployment Project: A Two-Year Follow-up Report* (Madison, WI: Wisconsin Department of Corrections, 1979).

Michele M. Connolly, “A Critical Examination of Actuarial Offender-Based Prediction Assessments: Guidance for the Next Generation of Assessments.” (PhD diss., University of Texas at Austin, 2003), available at repositories.lib.utexas.edu/bitstream/handle/2152/522/connollymm036.pdf?sequence=2.

Mike Eisenberg, Jason Bryl, and Tony Fabel, *Validation of the Wisconsin Department of Corrections Risk Assessment Instrument* (New York: Council of State Governments Justice Center, 2009), available at csgjusticecenter.org/wp-content/uploads/2012/12/WIRiskValidationFinalJuly2009.pdf.

OTHER TYPES OF INSTRUMENTS USED TO ASSESS RECIDIVISM RISK

Violence Risk Assessment Instruments

Violence risk assessment instruments, such as the Historical-Clinical-Risk Management-20 (HCR-20) (Webster et al. 1997) and Violence Risk Appraisal Guide (VRAG) (Quinsey et al. 2006), are intended to assess risk of future violence specifically, but also are frequently used to assess risk of non-violent recidivism.

HCR-20

The HCR-20 is a structured professional judgment scheme composed of 20 static and dynamic items that assess historical risk factors, clinical risk factors, and risk management factors. The individual item ratings are used to inform a final professional judgment of low, moderate, or high risk. Only one study has evaluated the validity of HCR-20 assessments in predicting recidivism in a U.S. sample (Barber-Rioja et al. 2012). Overall, the assessment total score was found to have excellent validity in predicting both general offending and technical violations. The HCR-20 has been widely validated outside the U.S. (see kdouglas.files.wordpress.com/2007/10/hcr-20-annotated-biblio-sept-2010.pdf). Materials for the HCR-20 are proprietary and must be purchased.

VRAG

The VRAG is an actuarial instrument designed for use with previously violent, mentally disordered offenders. It consists of 12 items that gather information on static and dynamic risk factors. Individual item responses are weighted and summed for a total score, which is then used to estimate level of risk based on an actuarial table. The predictive validity of VRAG assessments for both general offending and violations also has been evaluated in only one U.S. sample (Hastings et al. 2011). Validity in predicting general offending ranged from good to excellent for male offenders and fair to good for female offenders. Validity in predicting technical violations ranged from fair to good for male offenders and poor to fair for female offenders. Like the HCR-20, much research completed outside the U.S. has examined the validity of VRAG assessments. For more information, visit mhcp.on.ca. The VRAG is available at no cost.

References and Suggested Readings

Barber-Rioja et al., "The Utility of the HCR-20 and PCL:SV in the Prediction of Diversion Noncompliance and Reincarceration in Diversion Programs," *Criminal Justice and Behavior* 39, no. 4 (2012): 475–492.

Seena Fazel et al., "Use of Risk Assessment Instruments to Predict Violence and Antisocial Behaviour in 73 Samples Involving 24,827 People: Systematic Review and Meta-Analysis," *BMJ: British Medical Journal* 345, no. 1 (2012): 19.

Mark E. Hastings et al., "Predictive and Incremental Validity of the Violence Risk Appraisal Guide Scores with Male and Female Jail Inmates," *Psychological Assessment* 23, no. 1 (2011): 174–183.

Vernon L. Quinsey et al., *Violent Offenders: Appraising and Managing Risk* (2nd ed.) (Washington, DC: American Psychological Association, 2006).

Christopher D. Webster et al., *HCR-20: Assessing Risk for Violence Version 2* (Burnaby, BC: Simon Fraser University, Mental Health, Law, and Policy Institute, 1997).

Personality Assessment Instruments

Personality assessment instruments, such as the Psychopathy Checklist-Revised (PCL-R) (Hare 2003), the Psychopathy Checklist: Screening Version (PCL:SV) (Hart, Cox, and Hare 1995), and the Personality Assessment Instrument (PAI) (Morey 1991) evaluate personality constructs that correlate with criminal offending.*

PCL Instruments

The PCL-R is a 20-item actuarial assessment that can be used to diagnosis psychopathy, a form of antisocial personality disorder characterized by a persistent pattern of severe and refractory callous-unemotionality. Individual items are scored through file review and semi-structured interview, then summed for a total score ranging from 0 to 40 (where 30+ indicates the presence of psychopathy). The PCL:SV is a shorter, 12-item version. Again, individual item ratings are scored and summed, with a cutoff score of 18 typically used for classification of

* For a meta-analytic review see Jay P. Singh and Seena Fazel, "Forensic Risk Assessment: A Metareview," *Criminal Justice and Behavior* 37, no. 9 (2010): 965–988.

psychopathy. Research demonstrates excellent correspondence between the two measures in correctional samples (Guy and Douglas 2006). Validity of PCL-R and PCL:SV assessments in predicting recidivism has been evaluated extensively in the U.S., with performance ranging from poor to good (e.g., Gonsalves, Scalora, and Huss 2009; Salekin et al. 1998; Walters and Duncan 2005). For more information on the PLC-R and PCL:SV, see hare.org/scales/, where materials can also be purchased.

PAI

The PAI contains 344 self-report items that are divided into 22 validity, clinical, treatment consideration, and interpersonal scales. Individual item responses within the scales are hand scored and assessed in conjunction with interpretive guidelines included in the professional manual (Morey 2007). In U.S. studies assessing the predictive validity of the PAI, the assessment scale scores had fair to good validity in predicting general offending (e.g., Barber-Rioja et al. 2012; Walters 2009; Walters and Duncan 2005). For an overview and bibliography, see www4.parinc.com/Products/Product.aspx?ProductID=PAI, where materials can also be purchased.

Other Personality Assessment Instruments

Other instruments including the California Psychological Inventory: Socialization Scale (CPI:SO), Lifestyle Criminality Screening Form (LCSF), Minnesota Multiphasic Personality Inventory (MMPI), Neuroticism, Openness to Exposure Personality Inventory-Revised (NEO-PI-R), and the Peterson, Quay, and Cameron Psychopathy Scale (PQC) can produce valid assessments of recidivism risk, though performance varies widely.*

References and Suggested Readings

Virginia Barber-Rioja et al., "The Utility of the HCR-20 and PCL:SV in the Prediction of Diversion Noncompliance and Reincarceration in Diversion Programs," *Criminal Justice and Behavior* 39, no. 4 (2012): 475–492.

Robert D. Hare, *Hare Psychopathy Checklist-Revised (PCL-R): Second Edition, Technical Manual* (Toronto: Multi-Health Systems, 2003).

* See Glenn D. Walters, "Predicting Criminal Justice Outcomes with the Psychopathy Checklist and Lifestyle Criminality Screening Form: a Meta-Analytic Comparison," *Behavioral Sciences & the Law* 21, no. 1 (2003): 89–102; Glenn D. Walters, "Risk-Appraisal Versus Self-Report in the Prediction of Criminal Justice Outcomes: A Meta-Analysis," *Criminal Justice and Behavior* 33, no. 3 (2006): 279–304.

Stephen D. Hart, David N. Cox, and Robert D. Hare, *The Hare Psychopathy Checklist: Screening Version*, 1st ed. (Toronto: Multi-Health Systems, 1995).

Valerie M. Gonsalves, Mario J. Scalora, and Matthew T. Huss, "Prediction of Recidivism Using the Psychopathy Checklist—Revised and the Psychological Inventory of Criminal Thinking Styles within a Forensic Sample," *Criminal Justice and Behavior* 36, no. 7 (2009): 741–756.

Leslie C. Morey, *Personality Assessment Inventory Professional Manual*, 2nd ed (Lutz, FL: Psychological Assessment Resources, 2007).

Randall T. Salekin et al., "Psychopathy and Recidivism Among Female Inmates," *Law and Human Behavior* 22, no. 1 (1998): 109–128.

Glenn D. Walters, "The Psychological Inventory of Criminal Thinking Styles and Psychopathy Checklist: Screening Version as Incrementally Valid Predictors of Recidivism," *Law and Human Behavior* 33, no. 6 (2009): 497–505.

Glenn D. Walters and Scott A. Duncan, "Use of the PCL-R and PAI to Predict Release Outcome in Inmates Undergoing Forensic Evaluation," *Journal of Forensic Psychiatry & Psychology* 16, no. 3 (2005): 459–476.

Criminal Thinking Questionnaires

Criminal thinking questionnaires, such as the Psychological Inventory of Criminal Thinking Styles (PICTS) (Walters 1995) and the Texas Christian University Criminal Thinking Scales (TCU CTS) (Knight, Simpson, and Morey 2002), are designed to identify attitudes and thought patterns associated with criminal behavior.

PICTS

The PICTS is an 80-item, self-report measure composed of eight thinking pattern scales, two validity scales, four factor scales, two composite scales, and a General Criminal Thinking (GCT) scale. The validity of PICTS scores in predicting general offending has been evaluated in a number of U.S. studies with mixed findings. Performance of the GCT scale scores ranges from poor to good (e.g., Walters 2009a, 2009b, 2011); however, other research suggests the eight thinking pattern scales have poor validity (Gonsalves, Scalora, and Huss 2009).

TCU CTS

The TCU CTS is an actuarial, self-report instrument designed to measure criminal thinking. The instrument contains 37 items distributed across six thinking pattern scales: Entitlement, Justification, Power Orientation, Cold Heartedness, Criminal Rationalization, and Personal Irresponsibility. In one U.S. study, the six thinking pattern scale scores had poor validity in predicting both general offending and technical violations (Taxman, Rhodes, and Dumenci 2011). More information and a copy of the TCU CTS assessment materials are available at no cost from ibr.tcu.edu/pubs/datacoll/cjtrt.html.

References and Suggested Readings

Valerie M. Gonsalves et al., "Prediction of Recidivism Using the Psychopathy Checklist—Revised and the Psychological Inventory of Criminal Thinking Styles within a Forensic Sample," *Criminal Justice and Behavior* 36, no. 7 (2009): 741–756.

Kevin D. Knight, Dwayne Simpson, and Janis T. Morey, *TCU-NIC Cooperative Agreement: Final Report* (Fort Worth, TX: Texas Christian University, Institute of Behavioral Research, 2002).

Faye S. Taxman, Anne Giuranna Rhodes, and Levent Dumenci, "Construct and Predictive Validity of Criminal Thinking Scales," *Criminal Justice and Behavior* 38, no. 2 (2011): 174–187.

Glenn D. Walters, "The Psychological Inventory of Criminal Thinking Styles Part I: Reliability and Preliminary Validity," *Criminal Justice and Behavior* 22, no. 3 (1995): 307–325.

Glenn D. Walters, "Effect of a Longer Versus Shorter Test-Release Interval on Recidivism Prediction with the Psychological Inventory of Criminal Thinking Styles (PICTS)," *International Journal of Offender Therapy and Comparative Criminology* 53, no. 6 (2009a): 665–678.

Glenn D. Walters, "The Psychological Inventory of Criminal Thinking Styles and Psychopathy Checklist: Screening Version as Incrementally Valid Predictors of Recidivism," *Law and Human Behavior* 33, no. 6 (2009b): 497–505.

Glenn D. Walters, "Predicting Recidivism with the Psychological Inventory of Criminal Thinking Styles and Level of Service Inventory-Revised: Screening Version," *Law and Human Behavior* 35, no. 3 (2011): 211–220.

APPENDIX A

List of Jurisdiction-Specific Risk Assessment Instruments

1. Alabama Risk and Needs Assessment
2. Allegheny County Risk Assessment
3. Arizona Risk Assessment Suite
4. Arkansas Post-Prison Board Transfer Risk Assessment
5. California Parole Violation Decision Making Instrument
6. California Static Risk Assessment
7. Colorado Actuarial Risk Assessment Scale
8. Connecticut Salient Factor Score
9. Delaware Parole Board Risk Assessment
10. Georgia Board of Pardons and Parole's Field Log of Interaction Data
11. Georgia Parole Behavior Response and Adjustment Guide
12. Georgia Parole Decisions Guidelines Grid System
13. Georgia Department of Corrections Offender Tracking Information System
14. Hawaii Risk and Needs Assessment
15. Illinois Risk Assessment Instrument
16. Illinois Risks, Assets and Needs Assessment Tool
17. Indiana Risk Assessment System
18. Kentucky Pretrial Risk Assessment Instrument
19. Kentucky Parole Guidelines Risk Assessment Instrument
20. Iowa Board of Parole Risk Assessment
21. Louisiana Risk Needs Assessment
22. Maryland Public Safety Risk Assessment
23. Michigan Parole Guidelines Score Sheet
24. Mississippi Parole Risk Instrument
25. Missouri Sentencing Assessment Risk Instrument

26. Missouri Parole Board Salient Factor Guidelines
27. Montana Risk Assessment Instrument
28. Nebraska Criminal History Assessment instrument
29. Nevada Parole Risk Assessment
30. New Mexico Risk and Needs Assessment
31. North Carolina Risk Needs Assessment
32. Oregon Criminal History/Risk Assessment
33. Public Safety Checklist for Oregon
34. Orange County Pretrial Risk Assessment
35. Rhode Island Parole Risk Assessment
36. South Carolina Parole Risk Assessment Instrument
37. South Dakota Initial Community Risk/Needs Assessment
38. State of Hawaii LSI-R Proxy
39. Tennessee Offender Risk Assessment/Needs Assessment
40. Tennessee Parole Grant Prediction Scale and Guidelines
41. Texas Parole Risk Assessment Instrument
42. Utah Criminal History Assessment
43. Vermont Parole Board Risk Assessment
44. Virginia Pretrial Risk Assessment Instrument
45. Virginia Risk Assessment Tool
46. Washington Risk Level Classification
47. West Virginia Parole Board Assessment

Glossary of Terms

Actuarial Risk Assessment

Mechanical approach to risk assessment in which offenders are scored on a series of items statistically associated with recidivism risk in the sample of offenders upon whom the instrument was developed. The total score is cross-referenced with a statistical table that translates the score into an estimate of recidivism risk during a specified timeframe.

Area Under the Curve (AUC)

Performance indicator measuring the probability that a randomly selected offender who recidivated during follow-up would have received a higher risk classification using a given risk assessment approach than a randomly selected offender who did not recidivate during follow-up.

Cohen's d

Performance indicator measuring the standardized mean difference between the estimated level of risk or total score of offenders who did and did not recidivate during follow-up.

Dynamic Factor

Changeable characteristics (e.g., substance abuse) that establish a relative level of risk and help inform intervention; they can be either relatively *stable*, changing relatively slowly over time (e.g., antisocial cognition), or *acute*, changing more quickly over time (e.g., mood state).

Kappa (k)

Measure of inter-rater reliability representing the percentage of categorizations (e.g., low, moderate, or high risk) upon which multiple assessors agreed, statistically corrected for chance.

Intra-Class Correlation Coefficient (ICC)

Measure of inter-rater reliability representing the strength of agreement between multiple assessors on *continuous* variables (e.g., total scores), statistically corrected for chance.

Meta-analysis

Systematic review that includes a quantitative synthesis of the findings of *primary research*.

Observed Agreement

Measure of inter-rater reliability representing the percentage of categorizations (e.g., low, moderate or high risk) upon which multiple assessors agreed.

Odds ratio (OR)

An odds ratio (OR) is a measure of association between an exposure and an outcome. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure. Odds ratios are most commonly used in case-control studies.

Parole

Conditional release of a prisoner before the expiration of his or her sentence subject to conditions supervised by a designated parole officer.

Performance Indicator

Statistical measure of predictive validity.

Point-Biserial Correlation Coefficient (r_{pb})

Performance indicator measuring the direction and strength of the association between a *continuous predictor* (e.g., total score) and a *dichotomous outcome* (e.g., recidivating versus not).

Primary Research

Collection of new data that does not already exist.

Probation

Release of an offender from detention or sentence served in the community in lieu of detention, subject to conditions supervised by a probation officer.

Protective Factors

Characteristics of the offender (e.g., physical health, mental health, attitudes), his or her physical and/or social environment (e.g., neighborhood, family, peers), or situation (e.g., living situation) that are associated with a decrease in the likelihood of offending.

Recidivism

The repetition of criminal or delinquent behavior, most often measured as a new arrest, conviction, or return to prison and/or jail for the commission of a crime or for the violation of conditions of supervision.

Risk Assessment

Process of estimating the likelihood an offender will recidivate to identify those at higher risk and in greater need of intervention. Also may assist in the identification of treatment targets and the development of risk management and treatment plans.

Risk Assessment Instrument

Instrument composed of empirically or theoretically based risk and/or protective factors used to aid in the assessment of recidivism risk.

Risk Factors

Characteristics of the offender (e.g., physical health, mental health, attitudes), his or her physical and/or social environment (e.g., neighborhood, family, peers), or situation (e.g., living situation) that are associated with an increase in the likelihood of offending.

Somer's d

Performance indicator measuring the direction and strength of the association between an *ordinal predictor* (e.g., estimate of risk as low, moderate, or high) and a *dichotomous outcome* (e.g., recidivating versus not).

Structured Professional Judgment

Structured approach to risk assessment focused on creating individualized and coherent risk formulations and comprehensive risk management plans. Assessors estimate risk through consideration of a set number of factors that are empirically and theoretically associated with the outcome of interest. Total scores are not used to make the final judgments of risk. Instead, assessors consider the relevance of each item to the individual offender, as well as whether there are any case specific factors not explicitly included in the list.

Static Factors

Historical or otherwise unchangeable characteristics (e.g., history of antisocial behavior) that help establish absolute level of risk.

Systematic Review

A process by which the empirical literature from multiple primary studies on a particular topic meeting pre-determined inclusion and exclusion criteria is descriptively analyzed.

Technical Violation

A breach of the conditions of parole or probation.

Unstructured Risk Assessment

A subjective assessment of recidivism risk based on the assessor's intuition, knowledge of theory, and professional experience.

NOTES

¹ Patrick A. Langan and David J. Levin, *Recidivism of Prisoners Released in 1994* (NCJ 193427) (Washington, DC: Bureau of Justice Statistics, 2002).

² Stefania Ægisdóttir et al., “The Meta-Analysis Of Clinical Judgment Project: Fifty-Six Years Of Accumulated Research on Clinical Versus Statistical Prediction,” *Counseling Psychologist*, 34, no. 3 (2006): 341–382.

³ Jennifer L. Skeem and John Monahan, “Current Directions in Violence Risk Assessment,” *Current Directions in Psychological Science* 20, no. 1 (2011): 38–42.

⁴ Donald A. Andrews and James Bonta, *The Psychology of Criminal Conduct* (3rd ed.) (Cincinnati, OH: Anderson Publishing Company, 2003); Clive R. Hollin, “Risk-Needs Assessment and Allocation to Offender Programmes,” in *Offender Rehabilitation and Treatment: Effective Programmes and Policies to Reduce Reoffending*, ed. James McGuire (Chichester, UK: Wiley, 2002), 309–22.

⁵ Ernest Watson Burgess, “Factors Determining Success or Failure on Parole,” in *The workings of the Indeterminate-Sentence Law and the Parole System in Illinois*, eds. Andrew Alexander Bruce, Albert James Harno, John Landesco, and Ernest Watson Burgess (Springfield, IL: Illinois State Board of Parole, 1928), 221–234.

⁶ Donald A. Andrews, Craig Dowden, and Paul Gendreau, *Clinically Relevant and Psychologically Informed Approaches to Reduced Reoffending: A Meta-Analytic Study of Human Service, Risk, Need, Responsivity, and Other Concerns in Justice Contexts* (unpublished manuscript, 1999); Edward J. Latessa and Christopher T. Lowenkamp, “What are Criminogenic Needs and Why are they Important?” *For the Record* 4 (2005): 15–16; James Bonta and Donald A. Andrews, *Risk-Need-Responsivity Model for Offender Assessment and Rehabilitation* (Ottawa, ON: Public Safety Canada, 2007); Donald A. Andrews and James L. Bonta, *Level of Service Inventory-Revised: Screening Version (LSI-R:SV): User’s Manual* (Toronto: Multi-Health Systems, 1998).

⁷ Langan and Levin, *Recidivism of Prisoners Released in 1994*.

⁸ Christopher T. Lowenkamp and Edward J. Latessa, “Understanding the Risk Principle: How and Why Correctional Interventions Can Harm Low-Risk Offenders,” *Topics in Community Corrections* (2004): 3–8; Latessa and Lowenkamp, “What are Criminogenic Needs and Why are they Important?” 15–16; James Bonta and Donald A. Andrews, *Risk-Need-Responsivity Model for Offender Assessment and Rehabilitation* (Ottawa, ON: Public Safety Canada, 2007).

⁹ Lowenkamp and Latessa, “Understanding the Risk Principle,” 3-8; Latessa and Lowenkamp, “What are Criminogenic Needs and Why are they Important?” 15–16; Bonta and Andrews, *Risk-Need-Responsivity Model for Offender Assessment and Rehabilitation*.

¹⁰ Tracey Kyckelhahn, *Justice Expenditure and Employment Extracts, 2007–Revised* (Washington, DC: Bureau of Justice Statistics, 2012).

¹¹ David A. Anderson, “The Aggregate Burden of Crime,” *Journal of Law and Economics* 42, no. 2 (1999): 611–642.

¹² Pew Center on the States, *One in 31: The Long Reach of American Corrections* (Washington, DC: Pew Charitable Trusts, 2009).

¹³ Bonta and Andrews, *Risk-Need-Responsivity Model*.

-
- ¹⁴ Christopher T. Lowenkamp et al., "Adhering to the Risk and Need Principles: Does It Matter for Supervision-Based Programs?," *Federal Probation* 70, no. 3 (2006): 3–8.
- ¹⁵ Sarah L. Desmarais et al., "Using Dynamic Risk and Protective Factors to Predict Inpatient Aggression: Reliability and Validity of START Assessments," *Psychological Assessment* 24, no. 3 (2012): 685–700.
- ¹⁶ R. Karl Hanson and Andrew J. R. Harris, "A Structured Approach to Evaluating Change Among Sexual Offenders," *Sexual Abuse: A Journal of Research and Treatment* 13, no. 2 (2001): 105–122.
- ¹⁷ Don A. Andrews and James Bonta, *The Psychology of Criminal Conduct*, 2nd ed. (Cincinnati, OH: Anderson, 1998).
- ¹⁸ Catherine M. Wilson et al., "Predictive Validity of Dynamic Factors: Assessing Violence Risk in Forensic Psychiatric Inpatients," *Law and Human Behavior* 37, no. 6 (2013): 377–88.
- ¹⁹ Stephen D. Hart, Christopher D. Webster, and Kevin S. Douglas, "Risk Management using the HCR-20: A General Overview Focusing on Historical Factors," in *HCR-20 Violence Risk Management Companion Guide*, eds. Kevin S. Douglas, Christopher D. Webster, Stephen D. Hart, Derek Eaves, and James R. P. Ogloff (Burnaby, BC: Simon Fraser University, Mental Health, Law & Policy Institute, 2001), 27–40.
- ²⁰ Bree Derrick and Beth Skinner, *Risk Assessment: Moving from the Basics to the Tough Questions* (New York: Council of State Governments Justice Center, 2014).
- ²¹ Don A. Andrews and James Bonta, *The Psychology of Criminal Conduct*, 4th ed. (Newark, NJ: LexisNexis, 2006).
- ²² Robert L. Sampson and John H. Laub, "A Life-Course View of the Development of Crime," *Annals of the American Academy of Political and Social Science* 602, no. 1 (2005): 12–45.
- ²³ John Irwin, *The Warehouse Prison: Disposal of the New Dangerous Class* (Los Angeles: Roxbury Publishing, 2005); Erving Goffman, *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates* (Garden City, NY: Anchor Books, Doubleday and Company, 1961).
- ²⁴ Ægisdóttir et al., "The Meta-Analysis of Clinical Judgment Project," 341–382.
- ²⁵ Cynthia A. Mamalian, *State of the Science of Pretrial Risk Assessment* (Washington, DC: Pretrial Justice Institute, 2011).
- ²⁶ Skeem and Monahan, "Current Directions in Violence Risk Assessment," 38–42.
- ²⁷ John Monahan, "Mandated Community Treatment: Applying Leverage to Achieve Adherence," *Journal of the American Academy of Psychiatry and the Law* 36, no. 1 (2008): 282–285.
- ²⁸ Andrews and Bonta, *The Psychology of Criminal Conduct*; Randy Borum, "Improving The Clinical Practice of Violence Risk Assessment: Technology, Guidelines and Training," *American Psychologist* 51, no. 9 (1996): 945–956; Joel A. Dvoskin and Kirk Heilbrun, "Risk Assessment and Release Decision-Making: Toward Resolving the Great Debate," *Journal of the American Academy of Psychiatry and the Law* 29, no. 1 (2001): 6–10; Stephen D. Hart, "Actuarial Risk Assessment," *Sexual Abuse: A Journal of Research and Treatment* 15, no. 4 (2003): 383-88; Robert D. Hoge and Don A. Andrews, *Evaluation of Risk for Violence in Juveniles* (New York: Oxford University Press, 2010).

-
- ²⁹ Seena Fazel et al., "Use of Risk Assessment Instruments to Predict Violence and Antisocial Behaviour in 73 Samples Involving 24,827 People: Systematic Review and Meta-analysis," *BMJ: British Medical Journal* 345, no. 1 (2012): 19.
- ³⁰ Jack Fitzgerald and Steven Cox, *Research Methods and Statistics in Criminal Justice*, 3rd ed. (Belmont, CA: Wadsworth Thomson, 2002).
- ³¹ James Bonta, Moira Law, and Karl Hanson, "The Prediction of Criminal and Violent Recidivism Among Mentally Disordered Offenders: A Meta-analysis," *Psychological Bulletin* 123, no. 2 (1998): 123–142.
- ³² Derrick and Skinner, *Risk Assessment*.
- ³³ Alexander M. Holsinger, Christopher T. Lowenkamp, and Edward J. Latessa, "Validating the LSI-R on a Sample of Jail Inmates," *The Journal of Offender Monitoring* 17, no. 1 (2004): 8–9; Christopher T. Lowenkamp, Brian Lovins, and Edward J. Latessa, "Validating the Level of Service Inventory—Revised and the Level of Service Inventory: Screening Version With a Sample of Probationers," *The Prison Journal* 89, no. 2 (2009): 192–204.
- ³⁴ Edward Latessa et al., *Creation and Validation of the Ohio Risk Assessment System: Final Report* (Cincinnati, OH: University of Cincinnati Center for Criminal Justice Research, 2009), [available at ocjs.ohio.gov/ORAS_FinalReport.pdf](http://ocjs.ohio.gov/ORAS_FinalReport.pdf); Stephanie A. Evans, "Gender Disparity in the Prediction of Recidivism: The Accuracy of the LSI-R Modified" (master's thesis, University of Alabama, 2009).
- ³⁵ Holly A. Miller, "A Dynamic Assessment of Offender Risk, Needs, and Strengths in a Sample of Pre-Release General Offenders," *Behavioral Sciences and the Law* 24, no. 6 (2006): 767–782.
- ³⁶ David Farabee and Sheldon Zhang, *COMPAS Validation Study: First Annual Report* (Los Angeles: Department of Corrections and Rehabilitation, 2007), available at cdcr.ca.gov/Adult_Research_Branch/Research_Documents/COMPAS_Validation_Dec_2007.pdf; Tim Brennan, William Dieterich, and Beate Ehret, "Evaluating the Predictive Validity of the COMPAS Risk and Needs Assessment System," *Criminal Justice and Behavior* 36, no. 1 (2009): 21–40; David Farabee et al., *COMPAS Validation Study: Final Report* (Los Angeles: California Department of Corrections and Rehabilitation, 2010), [available at cdcr.ca.gov/Adult_Research_Branch/Research_Documents/COMPAS_Final_Report_08-11-10.pdf](http://cdcr.ca.gov/Adult_Research_Branch/Research_Documents/COMPAS_Final_Report_08-11-10.pdf).
- ³⁷ Ojmarrh Mitchell and Doris Layton Mackenzie, "Disconfirmation of the Predictive Validity of the Self-Appraisal Questionnaire in a Sample of High-Risk Drug Offenders," *Criminal Justice and Behavior* 33, no. 4 (2006): 449–466.
- ³⁸ See Jay P. Singh, Sarah L. Desmarais, and Richard A. Van Dorn, "Measurement of Predictive Validity in Studies of Risk Assessment Instruments: A Second-Order Systematic Review," *Behavioral Sciences and the Law* 31, no. 1 (2013): 55–73.
- ³⁹ Jay P. Singh, "Predictive Validity Performance Indicators in Violence Risk Assessment: A Methodological Primer," *Behavioral Sciences and the Law* 31, no. 1 (2013): 8–22.
- ⁴⁰ Kevin S. Douglas, Jennifer L. Skeem, and Elizabeth Nicholson, "Research Methods in Violence Risk Assessment," in *Research Methods in Forensic Psychology*, Barry Rosenfeld and Steven D. Penrod (Hoboken, NJ: John Wiley & Sons, Inc., 2011), 333.

⁴¹ Kristy Holtfreter and Rhonda Cupp, "Gender and Risk Assessment the Empirical Status of the LSI-R for Women," *Journal of Contemporary Criminal Justice* 23, no. 4 (2007): 363–382; Anne-Marie R. Leistico et al., "A Large-Scale Meta-Analysis Relating the Hare Measures of Psychopathy to Antisocial Conduct," *Law and Human Behavior* 32, no. 1 (2008): 28–45.