

## Top Common Mistakes or Misunderstandings in Juvenile Risk Assessment

By: Rachel Larsen, Senior Attorney, NDAA & Susan Broderick, Program Director, NDAA

A trend is emerging in some juvenile courtrooms and state legislatures to uncritically accept the latest risk assessment tools and adopt them without a thorough understanding of the tools' strengths and limitations. This trend is so prevalent that there are reports that some judges are making decisions around juvenile dispositions entirely based on a number produced by a juvenile risk assessor without first determining anything about the capacity, quality, or validity of the risk assessment.

With this trend in mind, here is a list of four overlooked issues with risk assessments performed in juvenile court:

(1) General risk assessment tools can predict recidivism with moderate accuracy, but they cannot predict whether someone will re-offend with a serious or a minor offense.

Risk assessment tools are, on average, more accurate than non-tool outcomes based on experience, gut feelings, or even clinical-level psychology skills. That does not mean these tools are omniscient. In fact, they are far from it: a risk assessment tool operating at a 60% accuracy rate is considered acceptable and will not be thrown out. Do you know how well the risk assessment performs in your jurisdiction? If you don't, find the statistics keepers in your jurisdiction and ask for the numbers. After all, if one wants to follow scientifically valid best practices, it is necessary to know how accurate the assessment is in practice.

The other piece of information we need to know is how likely is it that someone will re-offend, according to the assessment that your jurisdiction uses. For example, what are the chances that someone classified as low risk will re-offend? What are the chances that someone classified as medium risk will re-offend? In one study of the YASI (Youth Assessment and Screening Instrument), the one-year new petition rate based on the pre-screen showed that low-risk offenders reoffended at 9.9%; medium-risk offenders reoffended at 24.9%; and high-risk offenders reoffended at 36.8%.<sup>1</sup> The definition of petition was any new referral to the court with new charges.

And this is where one runs into a limitation on understanding of how to translate the risk assessment prediction into courtroom practice. One might reasonably say about a juvenile offender in the low-risk category that they do not need to have a courtroom intervention and instead should be diverted. Unnecessary involvement in the criminal justice system is harmful to juveniles, and it should be avoided.<sup>2</sup> That is a well-respected tenant of risk assessment science. But, what if the juvenile in question is charged with murder? Should one look at the low-risk score and release them to the custody of their parents?

For multiple reasons, prosecutors and judges should not release low-risk juveniles based solely on a risk score if there are indications of danger to the public because of a very serious offense. The first is that

<sup>1</sup> Robinson, D. & Jones, N. (2017) The Validity of the Youth Assessment and Screening Instrument for Justice Involved Youth in Milwaukee County. Orbis Partners. Long-Term Validation of the Youth Assessment and Screening Instrument (YASI) in New York State Juvenile Probation (2007).

<https://www.criminaljustice.ny.gov/opca/pdfs/YASI-Long-Term-Validation-Report.pdf>

<sup>2</sup> Andrews, D.A. Bonta, Hoge. The Psychology of Criminal Conduct, 5th ed. Chapter 2 (New Providence, NJ: Anderson Publishing, 2010).

low risk may mean that the juvenile has a 10% risk of re-offending within one year, but the risk score, very importantly, does not indicate what type of re-offending will occur.

What the 10% rate of re-offending – for low-risk juveniles – means is that the juvenile has a roughly 10% chance of committing a crime and that the crime could be anything from murder to stealing a pack of chewing gum. General risk assessments are incapable of predicting violent offenses<sup>3</sup>. So, when one considers this hypothetical 10% chance of re-offending and understands that re-offense could be something minor or another murder, the low-risk score is less re-assuring in the context of someone already charged with murder.

Another reason that a low-risk score should not mean immediate release is that “[r]isk assessment tools should inform legal decision-making and offer additional grounds for decisions but should not replace legal decision-making.”<sup>4</sup> This is fundamental; a risk assessment score is not the only factor in legal decision making. Other factors, often not included in risk assessments, include the seriousness of the current offense, victim input, offender accountability, whether there are pending cases, the current supervision status, and acceptance of responsibility. I have yet to meet a doctor in forensic psychology who believes that the risk assessment score should trump the seriousness of the current offense. And yet, some courtrooms operate as if the risk score is the only input to be considered.

## **(2) Risk assessment instruments must be tested on the juvenile population in your state and be separately calibrated for males and females.**

I recently spoke with a jurisdiction that is, by statute, rolling out several mandated risk assessment instruments. One of the assessments is newly created and has never been validated, and there currently aren't plans to validate it. The jurisdiction is also using YASI, which is widely used and researched. I asked one of the state trainers for the YASI risk assessment how long they planned on waiting before they validated the YASI on the population of their state. The answer was that there is no need to validate the test because it is “self-validating.”

There is no such thing as a “self-validating” risk assessment. Each state has a unique population of individuals that vary due to the state's unique culture, population subsets, and other factors. To give an idea of how this works, in a 2014 study, the LS-CMI (Level of Service Case Management Inventory) risk assessment which was developed in Canada functioned with almost fifty percent less accuracy in the U.S.<sup>5</sup> You might wonder how that can be the case when the principles of risk assessment don't vary no matter which country or state they are used in. Well, it turns out that when it comes to risk assessments, a test validated in Canada cannot be used in the U.S. without first studying its accuracy and adjusting its cut off points accordingly. The same is true for each state. The differences between states may not be 50%, but they are not negligible either. Yet, the states charged with validating the risk

<sup>3</sup> Min Yang, Stephen C.P. Wong, and Jeremy Coid, “The Efficacy of Violence Prediction: A Meta-Analytic Comparison of Nine Risk Assessment Tools,” *Psychological Bulletin*, vol. 136, no. 5, 2010, p. 757.

<sup>4</sup> Vincent, G.S., Guy, L.S., Grisso, T. (2012) [Risk Assessment in Juvenile Justice](https://njjn.org/uploads/digitallibrary/Risk%20Assessment%20in%20Juvenile%20Justice%20A%20Guidebook%20for%20Implementation.pdf).  
<https://njjn.org/uploads/digitallibrary/Risk Assessment in Juvenile Justice A Guidebook for Implementation.pdf>

<sup>5</sup> Olver, M.E., Stockdale, K.C., & Wormith, J.S. (2014). Thirty years of research on the level of service scales: a meta-analytic examination of predictive accuracy and sources of variability. *Psychological assessment*, 26 1, 156-76.

assessment (not an inexpensive endeavor) often do not know that for the test to perform most accurately, it must be validated on that state's population.

Once validated on the state's population, the need for validation does not end. Population characteristics, and thus risks, change as the population changes over time. Imagine, for example, giving the same risk assessment to a 1970s population of adolescents and a population living in 2020. The test for each group would need to be adjusted for that group. The phenomenon holds with shorter periods of time: if your jurisdiction adopted its risk assessment tool a decade ago and hasn't validated it for the current adolescent population, a new validation study needs to be done.

Another issue occurs when risk assessment instruments treat male and female adolescents the same. Males and females are different and respond differently to various forms of treatment and supervision. Yet when it comes to risk assessment, some tools treat them the same. Recidivism studies consistently show that females are less involved in criminal behavior, are less likely to commit violent crimes, and are less likely to recidivate after being placed on probation or parole. Thus, if the test isn't calibrated to these differences, it is not accurate.

**(3) Those that deliver risk assessments must be trained annually, certified to deliver the risk assessment, and their competence must be reviewed regularly. (Training alone is not enough.)**

To determine if the staff is sufficiently skilled enough to conduct the risk assessment, jurisdictions must measure the reliability of the individual test administrators. The best practice for doing this is called inter-rater reliability testing.

In this type of testing, all assessors are given the same example of an adolescent and asked to score his risk. For risk administrators whose results are too far outside of the correct risk score, retraining will be needed. This is a matter of quality control. We need to know that no matter which administrator gives the test, the results will be consistent. And yet, this is not being done in the field, and best practices can only be implemented when stakeholders, and in particular, prosecutors understand that it is a best practice to do this type of training and assessment. Go ahead and ask your risk assessment administrator whether they have received inter-rater reliability training and assessment. If they haven't, the lack of best practices needs to be brought up inside and outside of court.

Intensive initial certification as well as yearly training reboots are also necessary for accurate results. Complicated risk assessments like YASI and the LS require a well-structured interview and a review of all relevant case file data as well as contact with secondary sources of information. Complicated risk assessments often have over 30 inputs with several sub-scales reflecting varying spheres of risk levels. With such instruments it is more difficult to achieve the minimal levels of reliability without periodic auditing by qualified staff and yearly refresher training. While this finding may not apply to all jurisdictions, one 2015 study showed that only one in 69 officers administered Risk/Needs Assessment (RNA) correctly.<sup>6</sup> Without highly trained staff, the use of these instruments is not recommended.

<sup>6</sup> Jill Viglione; Danielle S. Rudes; Faye S. Taxman. *Criminal Justice and Behavior Volume: 42 Issue: 3 Dated: March 2015 Pages: 263-285.*

