Nebraska Detention Risk Assessment Instrument (RAI) Validation Study

Zachary Hamilton, Ph.D.

Anne Hobbs, Ph.D., J.D.

Addison Kobie, M.A.

Sadie Brewer, M.S.&

Monica Miles-Steffens, M.P.A.

Nebraska Detention Risk Assessment Instrument (RAI) Validation Study

Final- Submitted September 25th, 2025

Zachary Hamilton, Ph.D.

Anne Hobbs, Ph.D., J.D.

Addison Kobie, M.A.

Sadie Brewer, M.S. &

Monica Miles-Steffens, M.P.A.

Table of Contents

Acknowledgements

Executive Summary

The Administrative Office of the Courts and Probation (AOCP) Juvenile Probation Services Division of Nebraska contracted with the Juvenile Justice Institute (JJI) and Nebraska Center for Justice Research (NCJR) to validate the state's juvenile detention screening instrument: the Nebraska Intake Risk Assessment Instrument (RAI). Juvenile Probation has been responsible for this function since the passage of LB790 in 1988. The RAI was last enhanced in 2013 with national technical assistance from the Annie E. Casey Foundation (AECF). The Juvenile Probation Services Division has embarked on two efforts to validate the tool since that time. The University of Nebraska at Omaha Juvenile Justice Institute produced an evaluation of the tool in 2016 that concluded that while the tool appeared to be predictive, the 45% override rate impacted the overall findings. This report can be found on the Nebraska Supreme Court website. In 2020 the Juvenile Probation Services Division contracted with an out of state entity but concluded without a final report of study findings. For the current study, the research team, in collaboration with the AOCP, created a plan to assess the validity of the tool through statistical analysis and provide system and process recommendations via a qualitative approach. Collectively the study focused on two research questions.

1. *What factors affect probation officer implementation and use of the RAI?*

   For this question, the research team focused on a qualitative evaluation assessing Probation Officers (PO) and managers' attitudes regarding the value of the tool, and the processes of utilizing a standard assessment (training & practical application). The team utilized focus groups and surveys to gather feedback on the assessment and utilization. Further, other constituent groups (judges, attorneys & law enforcement) were also invited to provide their perspectives on the RAI.

2. *Is the RAI a valid and predictive assessment instrument?*

   For this question, the research team conducted a quantitative analysis focused on validating the RAI through its ability to predict Failure to Appear (FTA) and/or have a new intake to probation termed Future System Involvement (FSI). This analysis was multifaceted, and the team looked at the predictive strength of the RAI considering risk levels, and demographic values.

Findings: Overall, there was a consensus that having a uniform statewide tool brings objectivity to the intake process. However, qualitative findings demonstrated a lack of confidence with the RAI's ability to identify high risk youth in which detention is warranted. Inconsistent use of the tool was also described, where probation officers and stakeholders identified confusion regarding proper scoring of items. Noted inconsistencies impact the RAI's reliability and, in turn, its predictive performance.

For the quantitative analysis the study sought to assess, for those youth that were released, predictors of FSIs and FTAs. Findings revealed less-than-optimal use of the tool. First, overrides of tool recommendation are very common, reducing its effectiveness. Further, our analysis revealed the tool predicted FSI and FTA at a rate lower than that does not reflect an effective prediction tool. Conclusively, study findings indicate that the RAI does not meet industry validity standards. Supplemental findings demonstrate that the RAI can be improved by adjusting scoring to better fit the tool to the Nebraska youth population.

Recommendations are provided and suggest that the AOCP begin the process of redeveloping and modifying the RAI to ensure that it is validated to the Nebraska population. In addition, new training and policies should be developed to ensure the assessment's fidelity and quality assurance. Finally, routine validation should be completed to ensure that the assessment tool is working as intended.

Introduction

For the last 40 years, risk and needs assessments (RNAs) have provided an evidence-based method to assess the risk of future justice system involvement. For many years these tools provided a standardized approach for categorizing the intensity of supervision for adults on probation and parole supervision. However, RNA utility expanded, providing an evaluation of youth on supervision, classification in placement settings, and early release from detention (Sullivan & Childs, 2021).

In a similar expansion, RNAs have been developed to assess the need for pretrial detention. For young people that have been taken into custody and detention is being requested by law enforcement, it is best practice to use an objective screening tool to determine if they should be released, released with supervision on an alternative to detention, or securely detained. A detention hearing is typically held within 24 to 48 hours for youth placed on an alternative, staff secure, or in secure detention. At this hearing, a judge determines whether to extend the placement recommendation made by probation intake, or make changes, such as release the youth to an alternative or their home. The judge often uses statutory criteria to assess the risk of release, or more specifically, if the youth is likely to return to court for future hearings and remain arrest-free while in the community. The major considerations for judges are the youth's risk of 'failure to appear' (FTA) for future court hearings and/or 'future system involvement' (FSI).

Risk assessments are used to provide a standardized method for assessing youths' FTA and FSI risk. Best practice outlines that assessments are to be created and maintained via several stages of development. These stages include, 1) create a list of potential assessment items, 2) survey the target population with potential items, 3) track and observe those that commit FTA and FSI outcomes, 4) remove items that do not predict assessment outcomes, 4) deploy the developed tool, and 5) validate. Notably, Step 5 is used to identify the accuracy of an assessment that has already been deployed and make improvements to tool scoring to increase said accuracy.

The current study aimed to evaluate and validate the RAI tool. Specifically, we sought to answer two study questions A) what factors affect Probation Officer (PO) implementation and use of the RAI? and B) is the RAI a valid and predictive assessment instrument? Using focus groups and interviews we gathered information on the RAI implementation and use. Next, gathering a large sample of youth (5,604) assessed via the RAI, we examined rates of FTAs, FSIs, and use of overrides. Further, we examined risk level categories: Low (release without restriction), Moderate (release with restriction/alternative), or High-Risk (detention). These findings are further broken down by gender, race/ethnicity, and probation district.

Background

While juvenile intake started in 1998 with a statute change in LB790, since 2001, pursuant to Nebraska statute, the Nebraska Administrative Office of Courts and Probation (AOCP) has held responsibility for administering a risk assessment which determines whether a youth should be placed in juvenile detention. The AOCP developed an initial standardized tool and administered statewide training for officers conducting intakes. The current version was modeled after a

detention screening assessment utilized in Santa Cruz, California, a Juvenile Detention Alternatives Initiative (JDAI) model site and based on national best practices of the Annie E. Casey Foundation (AECF) in 2013.

Assessing risk for Failure to Appear (FTA) and Future System Involvement (FSI) is a complex process. Pursuant to Nebraska statutes at the time this report was written, a juvenile should be placed in a secure detention facility for only two reasons: (1) "immediate and urgent necessity for the protection of such juvenile or the person or property of another or (2) if it appears that such juvenile is likely to flee the jurisdiction of the court" (Rev. Stat. § 43-251.01(5)). When a juvenile is presented for intake, a Nebraska probation officer gathers information and completes the RAI following Nebraska Juvenile Intake Protocol (n.d.). According to the protocol, the RAI is completed when a youth meets one of the following criteria. The juvenile:

  1. Has violated a law;
  2. Is uncontrollable and has violated the law;
  3. Has violated the law and is on probation;
  4. Has violated the law and is in Department of Health and Human Services (DHHS) custody;
  5. Is an in-state or out-of-state runaway. Until a determination is made regarding detention authorization or alternative placement, the juvenile remains in law enforcement custody.

In Nebraska, youth taken into custody by law enforcement and for whom detention is being requested, must be screened by a trained probation officer using the Nebraska Juvenile Intake Screening Risk Assessment (RAI). According to Nebraska statute §43-260, only trained probation officers may administer the RAI and conduct intake screenings. Probation officers interview the youth, parent/guardian, and law enforcement and score five areas of the tool: 1) most serious instant offense, 2) legal status, 3) risk of FTA and re-offense, 4) mitigating, and 5) aggravating factors. Youth scores fall into three recommendation categories of a) release without restriction, b) return to community with restrictions/ alternative to detention (ATD), or c) detention (staff secure or secure). Probation officers may override the recommendation, either upward or downward, and must have the override approved by their supervisor.

In Nebraska, a youth who is placed in detention or on an alternative, must have a detention hearing within 24 to 48 hours excluding judicial holidays and weekends. The judge overseeing a youth's detention hearing may consider RAI findings and statutory considerations. A final decision is made by the presiding judge on whether to continue the youth's current placement or make changes, such as release to a lower level of supervision pending their next court date.

Assessments are an evidence-based practice used across many decision points in the justice system including pre-trial detention (Taxman, 2018; Berk, 2019; Sullivan & Childs, 2022). Tool accuracy is crucial as an assessment can be a way to give standardized information on risk, such as probability to commit a new offense or Failure to Appear (FTA) (Taxman & Dezember, 2017). While accuracy is a larger discussion that includes many different components, predictive validity is considered primary (Berk, 2019; Sullivan & Childs, 2022). Predictive validity is how accurate a tool is in correctly identifying the outcome of interest and is often what drives accuracy (Berk, 2019). This standard is the basis of assessments and how tools are validated (Taxman, 2018).

When tools are deemed valid, they can help guide decisions and give opportunities for interventions by practitioners (Taxman, 2018). For example, for a pre-trial tool, those individuals who are considered low risk of failing to appear are candidates to be released without restrictions. This type of release has many net benefits for individuals and the agency including reducing supervision costs and limiting system contact. Therefore, a validated tool is highly beneficial for agencies, standardizing the assessment of risk for the population. Moreover, standardization is important in the decision-making process as it can reduce subjectivity between multiple actors (Taxman, 2018). Specifically, a standardized tool allows multiple actors from diverse backgrounds (i.e., Judges, Probation officers) to have similar information to inform decision-making across all youths. This standard information being shared can be used to improve accuracy in decision making while reducing potential bias (Sullivan & Childs, 2022; Taxman, 2018).

Purpose of the Study

The purpose of this study was to assess the utility and validity of the RAI as it relates to the juvenile intake process. To accomplish this task, a robust process and outcome evaluation was designed and deployed. Two primary research questions were established:

Research Question 1: What factors affect probation officer implementation and use of the Nebraska RAI? To answer this question, the JJI team conducted four focus groups in October and November with probation officers and management. We then added an additional five focus groups in December 2024 and January 2025, to ensure probation officers and management had adequate opportunity to share their views, and to include the perspective of attorneys, judges and law enforcement.

Research Question 2: Is the RAI a valid and predictive assessment instrument? To answer this question, the NCJR team evaluated whether youth who are released under RAI guidelines, later 1) fail to appear in court (FTA) and/or 2) have a new intake to probation (within 6 months after the RAI was conducted) for new law violations, which we term future system involvement (FSI).

This research question is multifaceted and requires several analytical phases. First, we developed samples appropriate to answer Research Question 2. As some subjects are detained, and thus not able to commit an FTA or FSI, we created a sample of subjects that included those that were detained. This first sample allows for an understanding of intake types and reasons, release types, RAI scoring and overrides. We examined these sample descriptives in relation to key factors such as gender, race/ethnicity, and district.

The RAI classification provides a designation of low, moderate, or high risk, where each risk level corresponds to a recommendation of 'release without restrictions', 'release to an alternative to detention (ATD) or to parent/guardian with restrictions', and 'secure detention', respectively. For those youth that are detained, they are not in the community and prevented from experiencing negative outcomes prior to their case disposition. When evaluating the predictive strength of the RAI tool, it requires youth to be in the community and thus possess the ability to be involved in FTA's and FSI's. To be described in greater detail in the PHASE II section, youth

are often taken into custody by law enforcement or probation who then administer the RAI. If a judge decides to keep a youth in detention prior to their case disposition, this generally occurs within 72 hours of intake.[1] Therefore, a second, reduced sample removes those youth that were detained for 72 hours or more and is used to evaluate the RAI score and risk level classification's ability to predict FTAs and FSIs.

We begin by outlining the qualitative research process employed to explore the experiences of probation officers and practitioners. This process allowed us to examine the factors influencing their work, as well as the challenges they encounter when implementing a risk assessment tool. These findings informed the RAI validation in Phase II.

---

[1] 72 hours was chosen as a conservative estimate of when youths are detained prior to case completion. While most hearings are done in the first 24 hours, this number also accounts for delays because of weekends and/or holidays that may affect that initial hearing.

PHASE I: Process Evaluation

Qualitative Methodology

JJI conducted nine total focus groups that included probation officers, probation management, law enforcement, attorneys, and judges. The first groups were conducted with probation staff. These participants were randomly selected from a list of active probation officers and managers supplied by the AOCP administrative staff. The aim of this approach was to invite probation officers and management representing all of Nebraska's judicial districts, while keeping focus groups small enough to allow all participants an opportunity to speak. All focus groups included semi-structured questions and were conducted via an online platform, lasting approximately 90 minutes. Law enforcement, attorneys, and judges who were unable to join a focus group were given an opportunity to respond to questions via an online survey.

The focus groups were transcribed, and the data were then analyzed using MAXQDA, a qualitative data analysis software designed to organize, code, and interpret large volumes of text. MAXQDA allows researchers to systematically analyze qualitative data by identifying patterns, themes, and relationships within the text. The primary researcher reviewed the transcripts, coding themes that emerged from the discussions. A second researcher (who had not been involved in interviews or transcription) then independently coded the transcripts, to examine consistency in coding. Segments were identified, coded, and categorized within MAXQDA, enabling the research team to systematically track recurring themes and explore insights across focus group responses. This approach ensured a rigorous and structured analysis of the qualitative data while maintaining the privacy and confidentiality of all participants.

Results

The focus groups included a total of 30 probation officers with an average of 4.3 years of experience conducting intakes. The officers' individual experience ranged from 0.83 to 13 years, reflecting a diverse mix of newer and more seasoned probation officers. On average, managers participating in the focus groups had 11 years of experience supervising officers and ranged from 1-to-20 years.

All of Nebraska's 12 Districts were well represented by probation officers, with 67% to 100% of those invited participating. AOCP managers were invited from all Nebraska judicial districts, but Districts 1 and 5 did not have management representation in the focus groups (Table 1).

Table 1. Focus Group Participation of Probation Officers and Managers by District

District	Probation Officers N (Invited)	%		Managers N (Invited)	%		Total %
1	2(3)	67%		0 (2)	0%		40%
2	2(2)	100%		2 (2)	100%		100%
3J	4(6)	67%		1(3)	33%		55%
4J	4(5)	80%		4(4)	100%		88%
5	2(3)	67%		0 (2)	0%		40%
6	3(3)	100%		2(2)	100%		100%
7	2(3)	67%		2(2)	100%		100%
8	2(3)	67%		1(1)	100%		75%
9	1(2)	50%		3 (3)	100%		80%
10	3(3)	100%		1 (2)	50%		80%
11	2(3)	67%		2(2)	100%		80%
12	3(3)	100%		1 (2)	50%		80%

A list of law enforcement professionals was provided by AOCP and were well represented in both focus groups and surveys. Judges were also recruited via email from the AOCP. Attorneys were recruited by the Nebraska State Bar Association. Law enforcement personnel had the most representation with 47 participants (Table 2).

Table 2. Method of Participation by Attorney and Law Enforcement

Interviewee Type	Focus Group	Survey
Prosecutor	1	1
County Attorney	0	2
Defense Attorney	6	3
Private Attorney	0	4
Law Enforcement	12	35
Judges	5	4
Other	0	1

Sarpy County had the greatest representation across law enforcement and attorneys, at 30.8% n=20), of juvenile justice personnel (not including probation officer and management). Lancaster County had the next highest participation levels, followed by persons representing multiple counties, Box Butte (9.2%) and Douglas County (9.2%). Table 3 provides county representation broken down by county.

Table 3. County Representation by Combined Attorney and Law Enforcement Participants

County Represented	Participants		Percent
Adams	2		3.1%
Box Butte	6		9.2%
Buffalo	1		1.5%
Cheyenne	1		6.2%
Dawes	1		1.5%
Douglas	4		6.2%
Gage	1		1.5%
Hall	1		1.5%
Lancaster	12		18.5%
Multiple	7		10.8%
Richardson	1		1.5%
Sarpy	20		30.8%
Scotts Bluff	3		4.6%
Missing	5		8.0%
Total	65		100 %

Thematic Analysis

A total of 1,056 statements were coded across focus groups and survey documents, providing a robust qualitative response. Multiple themes emerged, with three themes predominant in the discussions and surveys. Expectedly, these themes correlated to the semi-structured questions. The three most mentioned subject areas included: training (67%), the RAI tool (58%), and the lack of uniformity across the state (42%).

Frequency of Topic Codes

Topic Code	Percentage (Count)
Process Not Uniform Across the State (+)	42% (5)
Use Interview Guide	17% (2)
Rural Issues (+)	25% (3)
Process (+)	25% (3)
NPACS	33% (4)
Training	67% (8)
District trains on intake	25% (3)
Shadowing Intake	25% (3)
Ongoing training	17% (2)
Time since training	17% (2)
NPOT	8% (1)
Training Needs	17% (2)
Training Processes	17% (2)
District-Based Training Differences	17% (2)
Recommendations for training: capture scenerios	8% (1)
Tool	58% (7)
Tool Shortcomings/Needs	17% (2)
Scale Ranking	17% (2)
Supervisor Support	17% (2)
Physical location of intake	17% (2)
Overrides/Underrides	17% (2)
Recommendation - aggravating and mitigating poin...	17% (2)
Following Guidelines	17% (2)

Training

Probation officers spoke at length about intake training. Currently all new probation officers receive basic intake training, which is provided online. Additional new probation officer training is provided at the district level related to district specific intake considerations. They indicated that ongoing training is provided by the local district and administrative office. Detailed themes and recommendations are described below.

Statements Related to Training

Category	Value
Training	82% (9)
Shadowing Intake	27% (3)
Time since training	18% (2)
District trains on intake	27% (3)
Ongoing training	18% (2)
NPOT	9% (1)
Training Needs	18% (2)
Training Processes	18% (2)
District-Based Training Differences	18% (2)
Recommendations for training: capture sceneri...	18% (2)
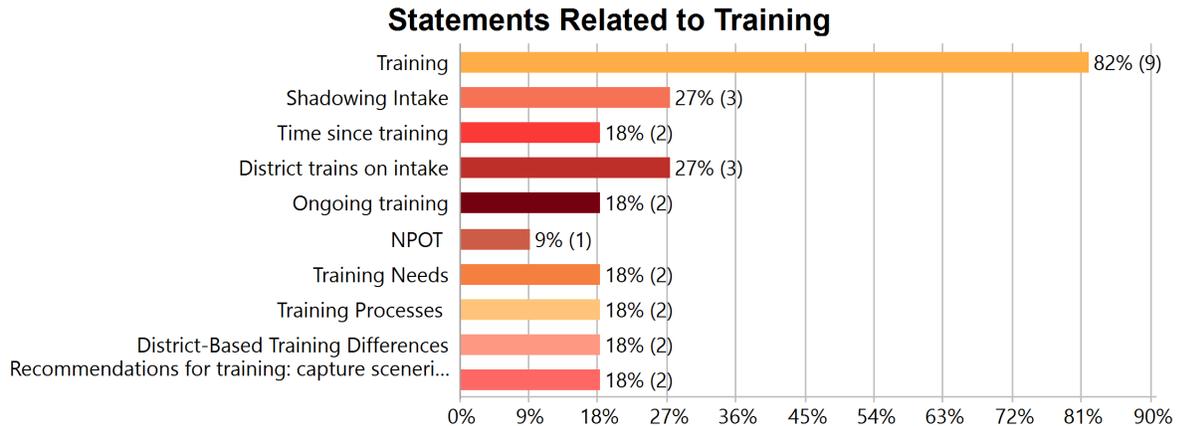
One of the most common discussion topics was that training completed at the district level is more in-depth and concrete. A frequent training topic included the benefits of district training where newer probation officers shadowed and then processed the intake with more experienced officers. By comparison, New Probation Officer Training (NPOT) was not the most frequent topic, but there was a consensus that *NPOT is more of an overview* and includes only a high-level description of the RAI. Probation officers indicated that much of the learning is hands-on and completed at the district level, which may explain some of the inconsistencies in the tool's use.

One newer officer stated, "I feel like it was recent for me since I haven't been here very long, and I can honestly say, I don't remember anything that I learned in NPOT about it. I mean, we probably went over it, but it was very vague, and it wasn't as in detail as it has been in my district." Another officer remarked that, "I believe every district almost has to do their own in-house training with intake, just because every district is so different, judges are different. So, you always have, you think almost every district does a follow up something after NPOT." Still another remarked: "it just is so district specific that they kind of say here's your exposure to it. Go back to your district, then figure out how it's done."

Some districts have taken steps to create their own specific training documents. While it might be expected that management would play a role in this process, in one district, probation officers indicated creating their own training materials. "I make an intake notebook for everybody that starts that has everything in it. Everything that they're going to need. When they get the notebook; I go through with them [explaining] what is [in] each of these forms."

Officers also indicated that booster training sessions are conducted, "we have them shadow for a while and then they get [to go out] and can do [the RAI]. And we also do a lot of intake refreshers here." It is clear administration also helps with training: "our supervisors train intake,
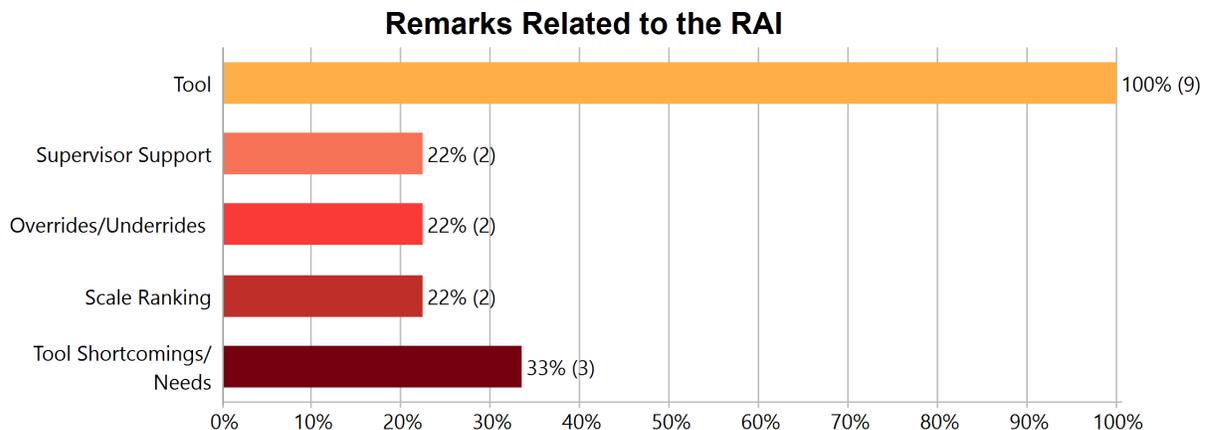
so we do boosters or whenever we get a particularly weird or unique case, we talk about it in our next juvenile staffing and just kind of go over it. So, if anybody else comes up with something strange in the next coming intakes, then they kind of know how to handle it."

One recommendation that officers had was statewide scenario training, especially between districts that have similar issues. One officer remarked "I don't know what information is being shared with other districts and I agree that that might be helpful, especially possibly between Lincoln and Omaha, since I think we're dealing a lot of the times with some of the same type of situations."

The next theme is the specific challenges that probation officers encountered while administering the RAI. These challenges emerged from our analysis and provided valuable insights into the practical difficulties of implementing the RAI in the field.

The RAI Tool

There was broad concurrence in the benefit of utilizing a statewide assessment tool. Most participants agreed that an objective tool reduces bias (implicit or otherwise) and helps officers make decisions, especially "in the middle of the night." Probation officers stated that "it is something solid you can rely on." In other words, if a youth scores for a certain level of restriction, then officers felt they could rely on that: "it's just something else to rely on that for your decision."

Remarks Related to the RAI

Category	Percentage
Tool	100% (9)
Supervisor Support	22% (2)
Overrides/Underrides	22% (2)
Scale Ranking	22% (2)
Tool Shortcomings/Needs	33% (3)

Although probation officers and management clearly value having objective guidance and structure for intakes, officers also widely acknowledged areas of subjectivity and flexibility that allow room to score the tool based upon their own perception. Some referred to areas of the RAI generally, "But then there's still those areas that are - that can be skewed based on your perception as the officer. And so, I think that's one of the pieces that with the RAI, I think it's very important to have the positives in there and then (you know) and everything else. But that's where you could easily go from you know an ATD to staff secure, secure detention. And I think there's just too much of that for our officers to - there's too much wiggle room there."

Both AOCP managers and officers indicated times when the RAI doesn't fit the situation, "You

hire officers and you're asking them to critically think through a situation, and sometimes the RAI doesn't fit with the situation that's at hand - and a lot of our interview is, here's the charges and we score them."

Still others cited specific areas of the tool that they struggled to score. One area was scoring violent felonies, which came up repeatedly. When recalling their training, one officer noted, "I think the thing that I remember that has stuck with me the most from it really was how to gauge whether or not we would score something as a violent felony offense, and I go back to that quite a bit when I'm looking at the scoring. Just because somebody else considers it violent doesn't necessarily mean that it meets the scoring requirements to be considered that."

Another acknowledged that "sometimes it's our own interpretation as to if it's a violent felony or just a felony. Other felony offense versus a violent felony offense because one of them on there is terroristic threats or felony for sexual assault. Look, that's a violent one, but it's under 'other felony offense,' so that's something that we have questions about [that area.]."

Probation officers described interactions with attorneys and judges that centered on the distinction between felony and violent felony. This is especially salient because the points can determine the outcome (12 points versus 6 points). One focus group described how attorneys on each side could argue whether the point system was valid: "it used to be 30-minute discussions over nothing. It has just been going back and forth. So, I think they finally just realized it wasn't worth anybody's time because it didn't. They're just arguing about nothing."

Participants keyed in on specific situations, especially related to scoring violent felony offenses. One scenario that came up repeatedly, as an example that causes disagreement related to violent felony, includes younger youth and vehicles. "We had a 13-year-old driving a stolen vehicle down [a] street like 75 miles an hour. . . . It ended up being a felony. But I put violent, which they (defense attorney) disagreed with, because to me, a 13-year-old has no business driving. They (the youth) don't know how to drive. They haven't been to drivers ed. So, I mean, I justified it to detain them (the youth). There [are] other factors, but the attorney said, well, that's not a violent felony. He didn't hit anybody. He didn't cause any damage. I just interpret it differently."[2]

Despite officers' certainty, there were numerous examples of situations where probation officers were unsure how to clearly score the tool. Officers were transparent about their uncertainty in scoring offenses as a violent felony and did not seem to find direction from the scoring guide, "I would say the lack of clarity on certain sections. So, if you look at the scoring guide specifically, . . . you look at . . . violent felony compared to the regular felony or violent, violent misdemeanor is not as bad. But when you look at violent felony in itself, it has some other things that are very clear, like it would almost be helpful if there was a second, or a third category, under the felony. . . I don't know it. It's weird how it is because certain charges show up as a violent felony when it's actually not a violent felony but if you read the scoring guide, it's supposed to be." Another officer stated, "The way that it's worded, it will make things look like a violent felony that aren't, and then sometimes it feels like things that should be in there, aren't."

Weapons were another area of concern. Some probation officers requested specific changes to

---

[2] In this example it was inferred that defense counsel disagreed with the probation officer.

the tool, specifically consideration for the addition of points under aggravating factors for weapons: "the one thing that I would like to see that I don't think is properly accounted for is the involvement of a firearm or other dangerous weapon. I mean, I know it can be incorporated as part of the offense, whether it's a violent offense or not but if they have a juvenile that's stolen a firearm from a vehicle - I don't know if that is necessarily incorporated as part of this tool. I think it's an aggravating factor that should be included, that that's my input. That's my concern . . . that firearms aren't being taken into account or other dangerous weapons."

Mitigating and Aggravation
Officers spoke openly about the flexibility and subjectivity of scoring aggravating and mitigating factors, some describing discomfort with the lack of objectivity inherent in scoring mitigating and aggravating factors. **"**There is [a section] underneath the … aggravating factors where you can put 'other'. That can be pretty general cause I mean it says things like school attendance and access to victims' stuff like that can be all - No local ties, stuff like that can be added into that section. So that means like and it's one point. But that kind of leaves it open to a whole bunch of interpretation, right there." The officer concluded, "If I need that one point for whatever I want, so it's not being used on, you know it's to me, it's the weirdest section because it is very - very subjective." Another noted "I find myself really utilizing the aggravating or mitigating factors, the one that we can fill in ourselves because I feel like some of the aggravating factors that matter more to me." This type of subjectivity and flexibility undermines the consistency of the tool and stakeholder confidence.

Overrides and Underrides: Alternative Placement and System Influence
Another subtheme included influence from other system players (i.e., law enforcement) regarding overrides and underrides. Probation officers discussed how overrides occur because "law enforcement is unwilling to transport," or "there's no place to put them [the youth.]" One group focused on youth with gun charges, "I think one of the [situations] is most notably youth with gun charges. We don't have an alternative for them. They don't necessarily always score [out] to be detained but if they aren't detained on that charge, there's no . . . other options because a lot of alternatives (even out of homes) won't take youth that have those charges." Other interviews revealed reduced placement options in the years since the pandemic: "Since COVID, like there is no shelter beds, you know. There is nothing." We heard the concerns related to lack of detention alternatives over and over again: "there's no place to put them . . . Yeah, or nobody will take them [the youth]."

Placement and problems at home are clear factors in many RAI decisions. We heard this theme repeatedly from probation officers, law enforcement and judges, who all expressed concern with situations in the youth's home, or victims within the home. One probation officer indicated that "I would say like a huge chunk of our overrides are parents who are victims, and we realistically cannot put the youth back in the home, whether they (the youth) used weapons, destroyed the (house) - whatever it is, you know."

Probation officers also *underride* the tool due to lack of placement and transportation issues: "I think we do more underrides, and it has to do with the score for out of home and there's no way to transport them. So, then we have to underride it, cause they're gonna have to go [home]. I

mean, I think we do that way more than the overrides." Probation officers mentioned the lack of alternatives across districts and in different focus groups: "Finding a shelter bed or a bed of any kind if you need that. We have to underride to home on an EM a lot."

Youth already adjudicated to probation was a cause for underrides in some circumstances: "We also have instances where it might be kids that are currently on probation, they have a history, they score pretty high on the RAI - but we have a lot of experience with them where we know, you know, they don't need that. A lot of them will score for detention. We know they don't need detention. And so, I do see us underriding in those instances because they really have the services they need already, and we just know enough about them." Some officers use a slightly different process for youth already on probation: "Sometimes the interviews are a little bit shorter. If there are kids [we know] already. We know a lot of those answers."

Some probation officers expressed feeling pressure from law enforcement: "I think one of the hard things that we hear and I don't know if other districts hear this . . . is when law enforcement . . . they tell you what the citation is going to be, and you have the outcome and they might say, 'well, how do I need to change this to get a different outcome?'" Pressure from law enforcement seems to arise from misunderstanding of the tool: "We run into it a lot with the law enforcement not fully grasping what they can and cannot detain on, and the age comes into play a lot. I've had to actually go and explain to our law enforcement like "This is why we do this, and this is the reasoning for it. We're not trying to fight with you. This is what we want to do, but we're limited too."

Warrants and Jurisdiction Detainers

Warrants and detainers are a primary reason for overrides: "I think another point is depending on which jurisdiction [you are located] some judges expect you to override or go to detention." Another probation officer remarked that judicial preferences sometime drive outcomes, and the officer may be concerned with being found in contempt for not following a court order. Probation officers repeatedly raised the issue of warrants and court orders: "[There are] court orders that we have to follow, and the same with warrants in our district. We just know (like) we have to detain them, or our judges will not be happy. Like it's just a standard in our district if there's a warrant, you detain them. No questions about it."

"On out-of-state runaways [are another reason] you have to override - because if they don't have a warrant that their parent is coming from another state to get them, you have to hold them. . . [even] if they don't have any charges, they just got picked up because they are an out of state runaway. They're going to score for release without restriction, but you got to hold on to them till their parents can get there. So, you have to override."

Overrides also occur because the Nebraska detention center structure no longer matches the structure the tool was originally based upon, and this can cause either overrides or underrides. Officers described this as follows: "I think another one . . . [is that] we don't have a staff secure facility. So, if they score staff secure, we have to go override or underride except for the rare occasion that we can use JJC, but that's more the exception than the norm."

Process Not Uniform

Despite widespread agreement that the aim of using a statewide tool is consistency and fairness in decision-making, differing approaches, especially at the district level, were a predominant theme across all focus groups. The third most common theme, occurring in almost every focus group, was that the RAI process is not uniform across the state.

Topics Related to Consistent Application of the RAI

Topic	Value
Process Not Uniform Across the State (+)	75% (6)
Rural Issues (+)	38% (3)
Supervisor Support	25% (2)
Process (+)	38% (3)
Physical location of intake	25% (2)
Use Interview Guide	25% (2)
NPACS	50% (4)
Following Guidelines	25% (2)
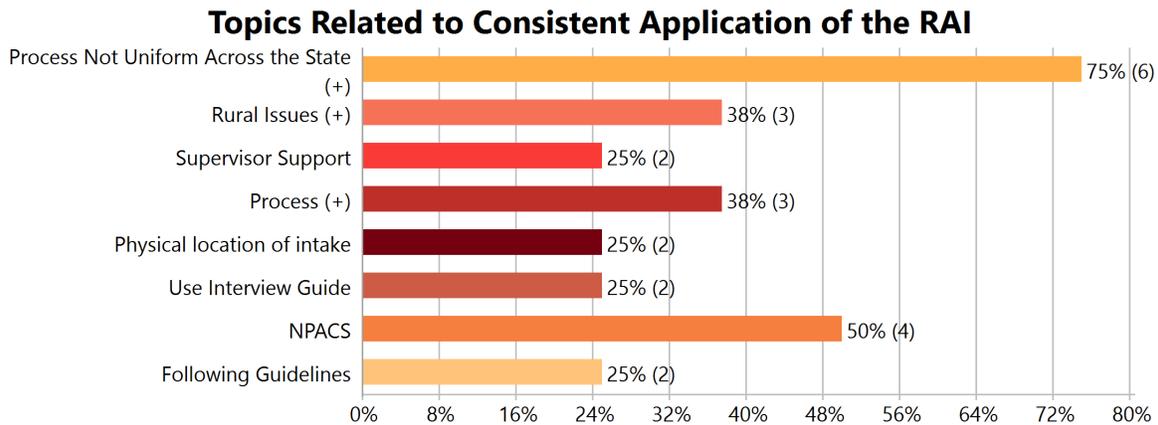
Some probation officers were transparent on this topic: "I get assigned one of the youth, one of my colleagues gets the other youth. Different probation officers score [the RAI] differently for the same exact offense. They [the youth] are both involved in the same stolen car, driving through the city. One is scored as a violent felony offense. One is scored as a felony offense makes a huge difference in terms of the bottom line is that kid going to meet the criteria for the tool to be detained." Probation officers noted specific areas they struggled with: "I just struggle with taking away a point just because they didn't get caught doing something until, they were 16."   Another area of inconsistency involves violations: "I think some of the challenges, each district kind of does things differently as it relates to probation violations."

Some officers follow the RAI Interview Guide closely, while others do not. One officer explained, "I don't. I don't rely on it all the time. I just it's there, but it's not. It depends on the situation . . . I sound like I'm in NPOT. It depends on the situation and 'go back to your district, though.'"

Some probation officers were surprised to hear of different approaches:" It's really surprising to hear kind of just the differences. I don't know if you're all surprised. Maybe you talk all the time to each other, but I'm like, well, it's different." Judges, who often see the intake tool from multiple probation officers, also felt that "there are definitely issues with consistency from one officer to another."

In other cases, the differences are not related to scoring but are about the process. For example, whether intakes are always conducted face to face, or the order in which the interview is conducted. One probation officer described, "Each officer I've noticed does it a little bit differently. Whenever I do mine, I try to get the law enforcement officer's statement first, just so I have a little bit of knowledge of what's going on. If the parents are there, then I'll hit the parents. If the parents aren't there, then I'll go to the youth. Those last two are just kind of

interchangeable."

JJI coded roughly 53 statements where the officer indicated their process differed from others across the state. These included: warrants, violations, face-to-face intakes, notifying a supervisor, length of the interview, and release without restriction. Probation violations were also an area where officers indicated slightly different approaches to intake: "I think some of the challenges, each district kind of does things differently as it relates to probation violations." Another officer directly expressed her surprise at hearing how different the districts approached things: "It's really surprising to hear kind of just the differences. I don't know if you're all surprised."

Process and Systemic Issues

Comments related to the assessment process were not simply about intake itself. Time constraints and transportation were the two most common sources of law enforcement frustration. One supervisor spoke to the amount of time the process can take, "I mean our night shift will bring a kid in at 2:00 (a.m.), our day shift will take the kid over at 6:00 in the morning and a lot of times when I'm going to lunch, that same kid is still sitting downstairs." Another officer indicated, "More often than not, the officer is made to wait a long period of time only to be made to transport the violent juvenile back home with an ankle monitor at best. The tool and juvenile justice system as a whole is and always have been severely flawed and needs revamped."

On the survey an officer notes: "The intake process is frustrating since the [facility] forces the officer to stay for the duration of the intake process. I've had several times where I was there for 3 or 4 hours all for the juvenile to not be detained. Whether they are detained or not, a road patrol officer should not be trapped inside the facility when there are other calls for service or legitimate lifesaving calls." We heard similar comments during focus groups with law enforcement: "We've had people doing assessments, not even show up just saying no, they won't lodge them on that and they didn't know all the facts - and I know a lot of the deputies and I know some of the [local] police officers because they've talked to me about the things that get very frustrated with how this whole thing goes." Although law enforcement clearly expressed frustration, transfer of custody is outlined in Neb. Rev Stat. sec. 43-250 (1)(c) and further dictates that "the peace officer shall implement the probation officer's decision to release or to detain and place the juvenile."

When a youth had access to, or had a weapon, law enforcement appeared most concerned and/ or frustrated by the process: "Well, [the probation officer will say] they [the youth] didn't use the gun, they just had the gun on them - that's very frustrating on my end."

Judges indicated an additional process consideration. Specifically, they felt like the summary was more helpful that the RAI itself: "Also getting the kind of the summary, the juvenile intake summary, which I think is a lot more helpful cause it will talk about the efforts that you know when they've talked to the what the police are saying, what interview with the kid? Not about the offense, but about circumstances and also talk to parents." While most judges understood they had access to the summary, some did not know that that could be included.

Justice System Collaboration and Opportunities for Improvement

Across almost all juvenile justice system professionals interviewed, there was genuine support for the concept of a uniform, statewide tool. Attorneys and judges cited the benefits of having a uniform tool to create consistency with decision making.

While some law enforcement agents endorsed the use of a tool, "It takes opinion and feelings out of the decision," and "it takes the subjective nature out of the decision," others indicated they did not value the tool or process, "There is [not] anything valuable that I can think of. It's really a waste of time. They never detain." Law enforcement officers, chiefs and sheriffs communicated high levels of frustration related to both the tool and the process. "Based on conversations with my [law enforcement] officers here, I think they get really frustrated. They feel like a kid here has to just about murder somebody to be taken into custody."

The topic of scoring violent felonies and the lack of clarity came up again during focus groups with law enforcement and attorneys. Judges who participated in focus groups specifically expressed concern regarding probation officers needing to interpret law. As one judge stated, "what we're asking our probation officers to do many times [is] to be lawyers and make a determination as to whether an offense is going to be a violent felony, 'another felony', a violent misdemeanor, or a misdemeanor."

Summary

Focus groups and survey data played a crucial role in understanding the findings that emerged during Phase II of this research study. By gathering qualitative insights from professionals who utilize the RAI tool and process on a daily basis, researchers were able to identify key themes, challenges, and patterns related to the study's objectives. These data sources provided a comprehensive understanding of the issues at hand, ensuring that our results were grounded in both lived experience and best practices statistically.

PHASE II: RAI Validation

Phase II of the RAI validation provided quantitative findings and was completed in several stages. First, we assessed sample descriptives, examining youth demographics and identifying the proportions of individuals with each RAI item response. Next, we examined the proportions of youth recommended and released to the community. We further examined override recommendations, again providing breakdowns by district and key demographic indicators. Third, the predictive strength of the RAI assessment was examined, including youth risk levels, and additional considerations for differential prediction by race/ethnicity and gender. Finally, we developed an updated version of the RAI as a "proof of concept", proposing a method to recalibrate item scoring to improve RAI predictive strength.

Methods

To assemble the needed measures for the RAI validation we worked with the AOCP to determine subjects' eligibility and obtain study data. First, we identified youth that were arrested in Nebraska for whom law enforcement requested a detention screening and were then assessed via the RAI at intake. A juvenile probation officer, trained in the provision of the RAI, administers the tool by reviewing the youth's offense history and conducting an interview. The tool items are scored, which determines the youth's release recommendation, which can include a) release without restriction (less than 5 points), b) return to community with conditions or alternative to detention (ATD) (6 to 9 points), c) staff secure (10 to 11 points), and d) secure detention (12 or more points).

A probation officer may override the release recommendation either upward or downward. The youth's completed RAI, including the override recommendation, is provided to the judge at a detention hearing for those youth who are placed on an ATD or in detention following intake. At the detention hearing, the judge considers whether to continue the placement or make a release recommendation. Nebraska statute 43-253(3) states, "A juvenile placed in an alternative to detention, but not in detention, may waive this hearing through counsel." Youth that are not detained are released to the community and returned to a guardian and/or ATD. Those that have conditions imposed may be supervised by a PO, with release orders held until a youth's case is disposed.[3] Youth that are detained might remain there until disposition; however, probation officers often work to find a less restrictive option when appropriate. Judges may review secure detention placement through case termination.

If a youth is released to the community and is cited with a new offense, probation will restart the intake process, complete a new RAI, and a "new law violation" is recorded. We define this outcome as *future system involvement (FSI)*. If a youth misses a court appearance, a judge may issue a warrant, where the youth is returned to custody. This event is documented as an '*FTA*'. The intent of the RAI is to assess FTA and FSI risk during the period between intake and adjudication. While most cases are adjudicated within 45 days, some may extend to six months (or 183 days). Therefore, we used a six-month follow-up to assess both FSIs and FTAs. Notably, a youth may receive an RAI and be held until their detention hearing up to 72 hours. Thus, we

---

[3] It should be noted that unless ordered by a judge, this hold is not an automatic process.

used a 72-hour window as an indication that a youth was released and not held beyond their detention hearing.

Data Frame

Working with subject matter experts from the AOCP, we were provided a sample of youth that completed an intake screening between July of 2019 and June of 2024, roughly five fiscal years of cases. This larger sample was used to examine intake types, reasons, RAI recommendations, overrides, and release types. Each of these measures was recorded via the RAI and provided to researchers.

To isolate the predictive performance of the RAI tool, as used by judges and staff considering release from detention, we needed to reduce the sample size. Specifically, to determine study eligibility for this portion of the evaluation, we removed youth that were detained for greater than 72 hours as these youth were detained and thus not able to commit FTAs or FSIs. Youth retained in this *reduced sample* were released and followed for six months to identify if an FTA or FSI occurred. This sample was used to validate the predictive performance of the RAI.

However, the *full sample* was still retained, and RAI performance was again analyzed. For any youth detained over 72 hours, we tracked FSI and FTA for the six months following release[4]. While the intent of the RAI is to predict FTA and FSI outcomes in the community prior to a youth's case disposition, the larger sample provides an opportunity to examine outcomes for all youth during their first six months in the community, regardless of release type. In total 3,875 youth were eligible to validate the predictive performance of the RAI, and an additional 1,729 youth were detained for more than 72 hours were included in a full sample, for a total of 5,604 study cases[5].

Predictive Performance

The focus of the validation was the examination of the tool's predictive strength, focusing on two distinct elements. First, we examine the RAI score's ability to predict FTAs and FSIs. As indicated, the RAI includes values associated with individual item responses, which are then summed to create a continuous risk score. To examine the continuous risk score, the industry standard metric to assess predictive validity and strength is the Area Under the Curve (AUC), which balances four critical statistics – false positives, false negatives, true positives and true negatives (Berk et al., 2021). Simply put, the AUC identifies the rate at which an individual is identified as higher risk and commits the event (true positive), lower risk and does not (true negative), higher risk and does not (false positive), and lower risk and does commit the event (false negative). This AUC ranges from 0 to 1 with a value of 0.5 representing random chance prediction. The AUC also represents an effect size, where values from 0.5 to 0.55 are considered 'negligible' (or near random), 0.56 to 0.63 'weak', 0.64 to 0.70 as 'moderate', and greater than 0.71 as 'strong' effects (Rice & Harris, 2005). Further, the industry has set a standard for

---

[4] While youth that are detained have a limited ability to commit an FTA, this additional sample is used to gauge the RAI's risk of committing an FSI upon release. We note that, for the full sample, the six-month follow-up does not include time spent in detention.

[5] We note that the sample did not include subjects that were detained.

predictive accuracy, where a value of 0.64 or higher is considered a minimum acceptable value for tool use (Desmarais et al., 2022)[6]. AUCs were computed for each outcome and were further broken down by gender and race/ethnicity[7].

The second element of predictive performance is the risk level category (RLC). As indicated, the RAI has three distinct categories – Low, Moderate, High, and Very High. We note, for the current study we combined High and Very High-Risk categories. Each category should identify progressively greater rates of the outcomes – FTA and FSI. To identify the magnitude of the RAI's prediction effect, we computed a logistic regression and used RLCs as an ordinal predictor. Two models were computed, one for each outcome, which were further broken down by gender and race/ethnicity. Odds ratios (ORs) are computed for each model, where the rate of the outcome observed for the Moderate and High-Risk categories are compared to Low. ORs possess an anchor point at a value of 1 that equates to equal odds, or "no effect", where values of 1.1 to 1.5 (or 0.67 to 0.9) are a small effect, values 1.5 to 2.5 (or 0.4 to 0.67) are a moderate effect, and values greater than 2.5 (or < 0.4) represent a large effect.

Sample

As described, most of the analyses are completed twice, once with a sample of eligible youth and again with the full sample of youth that experienced four or more days of detention. Prior to creating the sample division, we examined detention durations. To aid in the examination we created an ordinal distribution of detention days. According to AOCP most youth cases are disposed within 45 days, therefore, we created an ordinal measure for those that spent 0 to 3, 4 to 45, and 46 or more days in detention. These detention stays were further broken down by gender and race/ethnicity. We note that race is categorized as youth responding as White, Black, or any "Other" race category, while the ethnicity Hispanic is coded separate from race; therefore, youth could identify as Hispanic and one of the three indicated race categories. Descriptive statistics by detention stay days are provided in Table 4. While there was some unexpected variation for the 'Other' race category, generally 68% to 71% of youth were either not detained or detained and released within 72 hours, roughly 21% were detained for 4 to 45 days, and 10% were detained for more than 45 days.

---

[6] We note that this report provided a standard based on post-conviction risk assessments, where it is expected that detention screening tools are likely to have lower AUC values on average.

[7] We note that the duration of days following a youth's detention hearing and case disposition varies. While we opted to use a fixed, six-month follow-up period for its ease of interpretation, time-to-event (i.e. survival) models were also computed, providing near identical findings.

Table 4. Detention Days (N=5,604)

	Total (N=5,604)	Female (n=1,606)	Male (n=3,998)	White (n=1,940)	Black (n=1,745)	Other (n=550)	Hispanic (n=1,369)
Detention Days							
0-3	69%	68%	70%	71%	68%	61%	71%
4-45	21%	24%	20%	21%	21%	28%	20%
46+	10%	8%	10%	8%	11%	11%	9%

We then examined sample descriptives of the full sample. Descriptive statistics are provided in Table 5. Most youth in the sample were between 15 and 16 years old, with little variation observed by gender or race/ethnicity. Males represent 71% of the sample, with 35% identifying as White, 31% Black, and 10% classified as 'Other'[8]. Further, 25% of youth identified as Hispanic.

Next, we examined RAI item responses. Regarding Most Serious Offense, there were substantial variations by gender and race/ethnicity subgroup, with females committing fewer Violent Felonies (8%) compared to males (25%) and were committed by a greater proportion of Black youth (30%) compared to White (13%). Generally, males have greater rates of felonies (34%) compared to females (24%), where females possessed a greater rate of misdemeanors (38%), and status offenses (12%) compared to males (24% & 12%, respectively).

Regarding youths' legal status, most intakes were completed for a pending court case (45%), followed by current probation terms greater than six months (29%), and those less than six months (19%). Regarding youth's risk of FTA and re-offense indicators, most youth did not have an escape (99%), prior FTA (80%), pending law enforcement referral or citation (76%), or a previous arrest or citation for a new offense (87%). When examining mitigating factors, a minority of youth indicated a 'family member or caretaker able to assume responsibility' (14%), and most indicated 'first arrest at 16 or older' (88%), and 'no arrest or citations within the last year' (85%). Notably, there was little variation in responses by gender or race/ethnicity for the areas of the RAI.

When examining aggregating factors, 45% of youth indicated runaway behavior. Roughly 10% of youth indicated two or more violent offenses in the last year and 31% identified multiple offenses at intake. Yet, 49% of youth indicated 'no arrests or citations' within the last year. Notably, fewer male and Black youth indicated 'runaway behavior' as a mitigating factor (37% & 38%, respectively).

Next, we examined RAI risk levels and outcomes. Specifically, 28% were scored as Low, 36% Moderate, and 36% High Risk. A greater proportion of male (42%) and Black (43%) youth were identified as High-Risk, where females are more likely classified Low-Risk (36%). Regarding outcomes, 4% committed an FTA and 18% indicated FSI, yet these outcomes did not vary

---

[8] We note that the 'Other' category represents a mix of individuals identifying mostly as Native American, Asian, and Pacific Islander.

substantially by gender or race/ethnicity. It should be noted, 96% of the sample appeared in court as directed and were not issued a FTA, and 82% remained free from further citation, or did not commit an FSI.

Table 5. Sample Descriptives (N=3,869)

	Total% (N=3,869)	Female% (n=1,086)	Male% (n=2,789)	White% (n=1,758)	Black% (n=1,224)	Other% (n=893)	Hispanic% (n=977)
Age at Intake							
18	3	5	2	3	3	3	3
17	25	24	26	26	25	24	25
16	25	23	26	25	25	25	25
15	20	20	20	20	20	21	20
<15	27	28	27	27	27	27	27
Most Serious Offense							
Violent Felony	20	8	25	14	29	20	20
Violent Misd.	10	12	9	11	7	11	9
Felony	31	24	34	31	32	30	29
Misdemeanor	28	38	24	29	26	27	30
Probation Vio.	3	5	3	5	1	4	4
Status Offense	6	12	4	8	4	6	6
Legal Status							
Pending Court	5	5	6	6	4	5	5
Active Probation Reentry YRTC	25	27	24	26	26	22	23
Current Probation <6 mon.	21	23	21	22	20	20	21
Current Probation >6 mon.	49	46	50	46	50	52	51
Current Diversion	5	4	5	6	4	6	5
FTA/Re-offense Risk							
Escape 24 hours	1	1	1	1	1	1	1
FTA < 12 mon.							
0	80	76	82	85	75	79	78
1	15	18	14	11	19	16	17

2	4	5	3	3	4	4	4
3+	1	2	1	1	2	1	2
Referral/Citation Pending							
0	76	75	77	74	79	76	76
1	16	19	15	17	15	18	17
2	5	4	5	5	4	5	6
3+	3	2	3	4	2	2	2
No prior	87	88	87	86	88	87	86
Mitigating Factors							
Family/ Caretaker Responsible	14	16	14	17	8	17	15
First Arrest 16+	88	90	88	87	90	89	88
No arrest/citation in last year	85	86	84	85	83	87	86
Other	88	88	90	90	85	91	91
Aggravating Factors							
Runaway	45	66	37	49	39	46	47
2+ Vio. Offenses in Last Year	10	10	10	11	8	11	8
Multiple Offenses Current Intake	31	22	34	28	33	34	29
Other	49	49	49	49	47	51	49
RAI RLC							
Low	28	36	24	31	24	25	26
Moderate	36	40	35	38	34	36	37
High	36	24	42	31	42	39	37
Outcomes							
FTA	4	4	4	3	6	4	4
FSI	18	14	19	16	20	17	16

Intake and Release Reasons

The primary purpose of implementing a risk assessment is to provide standardization for determining release and supervision recommendations across a population. For the RAI this population spans the State of Nebraska and intakes for youth in 12 separate probation districts. Using the full sample of youth receiving an RAI (N=5,604), we describe the intake type by

probation district. Findings are presented in Table 6a. We note that the primary intake types are those currently on probation and those with a pending court case (37%), followed by those that do not have an active case at intake (22%), and those on diversion/administrative probation (4%). When examining variations, Districts 11 (45%) and 8 (43%) possessed the highest percentage of cases currently on probation, Districts 9 (52%), 3J (47%), 12 (47%) and 10 (46%) were observed to have the highest percentage with pending court cases, and District 1 (35%), 5 (29%), and 6 (28%) had the most 'not active' cases at intake.

Table 6a. Intake Type by District (with Detentions) (N=5,604)

District	% of Cases	Currently on Probation %	Pending Court %	Not Active at Intake %	Diversion/ Admin Probation %
1	3	31	31	35	2
2	9	40	41	17	2
3J	17	32	47	17	4
4J	37	41	30	26	2
5	4	40	27	29	4
6	7	38	30	28	4
7	5	38	39	18	5
8	1	43	26	26	4
9	8	25	52	20	4
10	3	37	46	14	3
11	2	45	35	18	2
12	4	33	47	16	4
Total	**100**	**37**	**37**	**22**	**4**

Next, we examined youth intake reasons for the full intake sample (N=5,604). Findings are presented in Table 6b. We note that the primary intake reason was a new law violation (53%), followed by those brought in on a warrant (41%), probation violation (5%), and runaway (2%). When examining variations, Districts 1 (80%), 8 (43%), and 5 (71%) possessed the highest percentage of new law violation cases, Districts 3J (51%), 4J (50%), and 2 (48%) have the highest percentage of cases brought in on a warrant, Districts 7 (26%) and 12 (14%) have the highest rate of probation violators, and Districts 11 (12%), 10 (10%), and 8 (10%) indicated the highest rate of runaways at intake.

Table 6b. Intake Reason by District (with detentions) (N=5,604)

District	New Law Violation %	Warrant %	Probation Violation %	Runaway %
1	80	9	6	5
2	43	48	9	0
3J	42	51	6	1
4J	49	50	0	1
5	71	20	7	2
6	61	25	7	7
7	43	31	26	1
8	83	8	0	10
9	58	38	2	2
10	62	26	3	10
11	68	16	4	12
12	67	14	14	5
Total	**53**	**41**	**5**	**2**

Following the completion of the RAI, a recommendation for detention or release is made, where findings indicated notable variations by district. Again, using the full sample (N=5,604) we examined RAI recommendations by district, with sample descriptives provided in Table 6c. The primary recommendation indicated was 'detention' (41%), followed by ATD or return to parent with conditions (38%), and release without restrictions (21%). Notably, RAI recommendations for detention were most prevalent for Districts 7 (62%) and 3J (50%). RAIs recommending ATDs and 'return to parent with conditions' were most common for Districts 10 (54%), 2 (50%), and 9 (49). Finally, RAIs recommending releases without restrictions were common for District 8 (45%), 6 (43%), and 1 (34%).

Table 6c. RAI Recommended Release Type by District (with Detentions) (N=5,604)

District	Detain %	ATD or Return to Parent %	Release w/out Restriction %
1	28	35	34
2	44	50	6
3J	50	36	13
4J	47	32	22
5	37	39	24
6	17	40	43
7	62	20	18
8	41	14	45
9	42	49	10
10	18	54	29
11	35	44	21
12	44	43	13
Total	**41**	**38**	**21**

Overrides

While a primary purpose of an assessment is to standardize recommendations, unless directed by policy, overrides to the RAI recommendation can undermine the intent of the assessment. However, some overrides are necessary, particularly when additional and important information is not captured via the tool. Yet, best practice has indicated that assessment overrides should not occur in more than 10% of cases (Vincent et al., 2012). Specifically, probation officers may recommend an override to a higher level (i.e., release with restrictions to detention), or to a lower level (i.e., ATD to release without restrictions). In Table 7 we compare override percentages by district. We find that 60% of RAI recommendations were not overridden, with 33% of scores overridden to a higher level, and the remaining 7% to a lower level. Districts 10 (87%) 1 (80%), and 9 (80%) have the highest percentage of cases in which the RAI recommendation is followed; however, we note that even District 10 still has an override rate that exceeds best practice recommendations (13%). Further, Districts 7 (67%), 3J (60%), and 2 (47%) have the highest rates with all districts demonstrating substantial rates of overrides to a 'higher level' than recommended via the RAI.

Table 7. Overrides by District (with Detentions) (N=5,604)

District	Lower Level %	No Override %	Higher Level %
1	2	80	18
2	8	52	41
3J	4	40	55
4J	10	62	28
5	3	64	33
6	6	66	29
7	2	34	65
8	11	55	34
9	7	80	13
10	2	87	12
11	8	71	22
12	7	76	17
Total	**7**	**60**	**33**

Next, using the full sample (N=5,604) we examined the most serious type of offense. This analysis was completed to identify cases in which the RAI score is most likely to be overridden. Findings are provided in Table 8. Notably, cases overridden to a lower level are some of the most serious offenses – felony (36%) and violent felony offenses (50%). When examining overrides to a higher level, violent felonies were the least likely (5%), while the most likely to have a RAI score overridden to a higher level were misdemeanors (38%), followed by the least serious offenses – status, probation violations, and other offenses (21%).

Table 8. Most Serious Offense by Overrides (with detentions) (N=5,604)

Most Serious Offense	Lower Level %	No Override %	Higher Level %
Status/Other/PV	1	8	21
Misdemeanors	8	27	38
Violent Misdemeanors	5	9	17
Felony Offense	50	33	20
Violent Felony Offense	36	23	5

In Table 9 we examine release type when a probation officer override was provided (N=2,258). In 67% of cases in which an override was recommended, the youth was detained, followed by ATD and conditional release to parent (28%), and release without restriction (5%). Override recommendations in which a youth was detained occurred at the highest rate for Districts 7 (90%) and 3J (82%).

Table 9. Release type for overrides by district (with detentions) (N=2,258)

District	Release W/out Restriction %	ATD or Return to Parent %	Detain %
1	3	36	61
2	1	33	66
3J	1	17	82
4J	11	28	62
5	1	38	61
6	2	50	48
7	1	10	90
8	8	25	65
9	1	48	50
10	0	52	48
11	8	61	32
12	2	35	63
Total	**5**	**28**	**67**

To provide an overview of the RAI release and override analyses, a summary is presented in Table 10. We note that probation officers recommend detention in 52% of cases, followed by ATD or condition release to parent (34%), followed by releases without restriction (14%). However, when the RAI scores a youth as Low, a 'release without restriction recommendation' is only retained by probation officers in 50% of the cases and overridden to detention 21% of the time and overridden to ATD or condition release to parent in 28% of cases. When an RAI score is classified as Moderate, a release recommendation (i.e., ATD or return to parent) is only followed in 60% of the cases, where a probation officer recommends an override to detention for Moderate cases 36% of the time. Finally, an RAI that scores a youth as High was overridden to a release (or a lower level) in only 13% of the cases. Further, of the 5,604 cases, agreement between probation officers' release recommendations and RAI classification were observed roughly 68% of the time.[9]

Table 10. RAI Risk Level by Release Type (with detentions) (N=5,604)

RAI RLC	Release w/out Restriction % (frequency)	ATD or Return to Parent % (frequency)	Detain % (frequency)
Mod	4 (84)	60 (1,278)	36 (777)
High	2 (28)	11 (232)	88 (1,828)
Total	**14 (803)**	**34 (1899)**	**52 (2,895)**

Overall, these findings describe the types and reasons for intakes, release types, and overrides by district. We find that overrides are common and pervasive across districts. Probation officers request an override of the RAI recommendation in 32% of the cases. Finally, probation officers

---

[9] This percentage was computed by adding the diagonal frequencies and dividing by the total ((690+1,278+1,828) / 5,604).

were more likely to request a higher-level placement, with overrides to detention being the most prevalent.

RAI Predictive Performance

Next, we examined the accuracy of the RAI in predicting FSIs and FTAs. Using the reduced sample (N=3,875), or those detained for 0 to 72 hours, we computed AUC statistics to examine the predictive strength of the RAI assessment. In Table 11a we provide the AUCs for the sample, broken down by gender and race/ethnicity subsamples. All AUC values were revealed to be 0.5 or below, indicating that the RAI's prediction is slightly less than random. This finding held regardless of gender and race/ethnicity. These AUC values indicate the RAI does not meet established predictive validity criteria, as values are below the industry standard for justice population use (AUC = 0.64).[10]

Next, we examined the rate of FTAs and FSIs by Risk Level Category (RLC). Like the AUC findings, the RLC FSI and FTA rates did not increase in a common 'stairstep pattern' with each progressive risk level. In fact, those youth scoring high risk, and released to the community, were observed to have lower rates of FTAs and FSIs than Moderate and Low-Risk youth. This finding was relatively consistent across gender and race/ethnicity categories. Further, odds ratio (OR) findings demonstrated a lack of substantial increase and/or inversion[11] of the predictive direction of risk level categories on FTAs and FSI.

Table 11a. AUC and RLCs by Outcome (N=3,875)

	Total (AUC)	Female (AUC)	Male (AUC)	White (AUC)	Black (AUC)	Hispanic (AUC)	Other (AUC)
RAI Total Score							
FTA	0.43	0.45	0.42	0.45	0.43	0.37	0.38
New Offense	0.47	0.50	0.43	0.46	0.46	0.47	0.49
	Total % (OR)	Female% (OR)	Male % (OR)	White % (OR)	Black % (OR)	Other % (OR)	Hispanic% (OR)
RAI RLCs FTA							
Low	5	5	5	4	6	7	6
Moderate	5 (1.08)	6 (1.30)	5 (0.98)	4 (1.20)	8 (1.20)	5 (0.64)	6 (1.17)
High	3 (0.05)	3 (0.54)	3 (0.53)	3 (0.51)	4 (0.51)	2 (0.26)	2 (0.95)
RAI RLCs FSI							
Low	17	12	20	17	21	19	15
Moderate	18 (1.09)	14 (1.18)	20 (1.02)	19 (1.09)	24 (1.09)	19 (0.96)	17 (0.96)
High	15 (0.88)	13 (1.09)	16 (0.75)	15 (0.18)	16 (0.21)	13 (0.24)	14 (0.37)

---

[10] We note that our comparison of AUC values by gender and race/ethnicity provides an examination of predictive validity for subgroups. Additional analyses were completed (i.e. Cleary Method) to assess slope and intercept prediction bias. However, due to the RAI's lack of prediction strength, bias assessments were inconclusive.
[11] We note 'inversion' is observed when lower risk level categories predict a greater rate of FSI or FTA than higher category levels.

While the predictive performance results of those released within 72 hours is less-than-optimal, it is possible that the tool's performance is better for those detained for longer periods, are higher risk and, in turn, more likely to commit FSIs and FTAs upon return to the community. Therefore, using the full sample (N=5,604), which contains youth detained for greater than 72 hours, we again examined the predictive performance of the RAI assessment. In Table 11b we provide the AUCs for the sample, broken down by gender and race/ethnicity subsamples.

Unfortunately, the performance is roughly the same when assessing the full sample. Again, AUC values were observed to be below 0.5, indicating that the RAI's prediction is slightly less than random. While the female AUC values were above 0.5 for both outcomes (0.52 & 0.53, respectively), we note that, regardless of gender and race/ethnicity, the AUC values indicate the RAI is *not a valid prediction tool for FSIs and FTAs*, where the observed performance is far less than industry standard for use with a justice population.

Next, we examined the rate of FTAs and FSIs by RLC. Like the previous findings, the RLC rates did not increase in a common stairstep pattern. Again, those youth scoring High-Risk were observed to have lower rates of FTAs and FSIs than those classified as Moderate and some Low-Risk youth populations. This finding was relatively consistent across gender and race/ethnicity categories. Further, odds ratio (OR) findings demonstrated a lack of substantial increase and/or inversion of the predictive direction of risk level categories on FTAs and FSIs.

These findings are notable in that they include those that were detained, some for a substantial duration. Our results indicate that, following release, those youth that were detained exhibit lower rates of FTAs and FSIs than their Moderate and Low-Risk counterparts. Given the similarity of findings with the reduced sample, these results indicate that the RAI is a consistent, poor predictor of FTA and FSI risk.

Table 11b. AUC and RLCs by outcome (N=5,604)

	Total (AUC)	Female (AUC)	Male (AUC)	White (AUC)	Black (AUC)	Other (AUC)	Hispanic (AUC)
RAI Total Score							
FTA	0.49	0.52	0.48	0.45	0.43	0.39	0.40
FSI	0.49	0.53	0.48	0.48	0.48	0.51	0.50
	Total % (OR)	Female % (OR)	Male % (OR)	White % (OR)	Black % (OR)	Other % (OR)	Hispanic% (OR)
RAI RLCs FTA							
Low	5	5	5	4	6	7	6
Moderate	6 (1.08)	6 (1.30)	5 (0.99)	4 (1.22)	7 (1.15)	6 (0.81)	6 (1.03)
High	3 (0.54)	3 (0.54)	3 (0.53)	2 (0.60)	4 (0.06)	3 (0.32)	3 (0.41)
RAI RLCs FSI							
Low	17	12	20	16	19	16	15
Moderate	18 (1.09)	14 (1.18)	20 (1.02)	16 (1.04)	21 (1.13)	17 (0.56)	17 (1.16)
High	15 (0.88)	13 (1.09)	16 (0.24)	14 (0.18)	16 (0.80)	15 (0.94)	15 (1.04)

Upon request from the AOCP, we further examined the predictive validity of the RAI using cases in which no override was requested. The concept of this analysis was to identify if those cases in which *the probation officer agreed with the RAI classification* more accurately predicted FTAs and FSIs. We did not find predictive improvement when the override cases were removed. Specifically, only the AUC for White-FTA increased and exceeded random chance (AUC = 0.51). AUC findings, broken down by gender and race/ethnicity, are provided in Table 12a separately for subjects that were not detained and/or those not detained for more than 72 hours, and compared to AUCs for all subjects.

Table 12a. Predictive validity for cases POs did not request an override (N=3,339)

	Total (AUC)	Female (AUC)	Male (AUC)	White (AUC)	Black (AUC)	Other (AUC)	Hispanic (AUC)
Not detained *(72hrs)*							
FTA	0.40	0.43	0.39	0.51	0.42	0.37	0.35
FSI	0.44	0.48	0.42	0.42	0.37	0.48	0.48
All subjects							
FTA	0.40	0.44	0.40	0.50	0.37	0.30	0.37
FSI	0.45	0.50	0.42	0.43	0.43	0.48	0.47

To follow through with the concept of overrides having an impact on predictive validity, we next examined cases in which probation officers requested an override. We note that when a probation officer requests an override, the RAI score of the subject does not change, only the recommended release type. Table 12b provides AUC findings for subjects in which probation officers felt the RAI score was not accurate and requested a different release type. As shown, the

AUC values are larger by comparison to all subjects presented previously. Specifically, the Total, Female, Male, Black, and Hispanic youth AUC values exceeded 0.50 (or random chance). Further, for Black youth the values for both outcomes, and for the samples that included and excluded youth detained for 72 or more hours, the AUC findings exceeded a small effect size (AUC > 0.55).

These findings are counter intuitive, in that they suggest that for cases in which the probation officer seeks to override the tool's recommendation, the RAI is more accurate than the probation officer's recommendation. Another unique finding indicated that probation officers were more often overriding Black youth's scores leading to overclassification, as the RAI appears to be more accurate for cases in which the POs requested an override for Black youth. Overall, while the findings suggest that the RAI is not a valid predictor of FSIs and FTAs, when probation officers decide to override the tool, they choose to do so on cases in which the RAI is most accurate. This finding suggests that human intervention with the RAI recommendation makes an already poor prediction slightly worse.

Table 12b. Predictive Validity for Cases POs did Request an Override (N=2,258)

	Total (AUC)	Female (AUC)	Male (AUC)	White (AUC)	Black (AUC)	Hispanic (AUC)	Other (AUC)
Not detained *(72hrs)*							
FTA	0.52	0.51	0.51	0.48	0.59	0.32	0.51
FSI	0.53	0.55	0.51	0.46	0.58	0.45	0.52
All subjects							
FTA	0.49	0.51	0.48	0.48	0.57	0.42	0.45
FSI	0.56	0.57	0.54	0.53	0.58	0.53	0.54

Updated RAI: Proof of Concept

As part of the scope of work (SOW) we sought to make improvements to the RAI and provide recommendations for future use. Based on prior assessment development efforts and recent findings, assessments are demonstrated to be more accurate when developed using local data (Hamilton et al., 2022). Specifically, when assessment items are statistically modeled using agency-collected data and records, a more accurate prediction model is produced (Duwe, 2024).

Using prior established methods, we developed an updated version of the RAI, recalibrating item scoring to improve the tool's predictive strength (Hamilton et al., 2020). We outline newly created models as a *proof of concept,* demonstrating that improvements can be made but that the models developed should not simply replace the current RAI without further consideration.

We note that when computing a prediction model with a single sample, instability of the prediction and overfit can become an issue. More specifically, a model that is constructed and validated on the same sample will demonstrate performance metrics that are artificially high. Therefore, we completed a cross-validation procedure by randomly splitting the sample in two, with a 75% construction sample (n=2,907) and 25% validation sample (n=969). We used a customized ridge regression analysis, that selects (or removes) items, through an iterative modeling process (i.e., one at a time), until all items included in the model positively improve the

model AUC value[12].

To stabilize the selection of items and weights we use a burgeoning validation procedure, k-fold. The k-fold creates a second layer of cross-validation. The method randomly splits that sample into 10 parts, running construction models on 9 parts and validating the prediction on the 10th. This process is completed 10 times, with replacement, producing an 'average' model of items and weights observed across multiple iterations of the construction folds (see Kohavi, 1995). Predictive metrics from each of the 10 test parts are also averaged across the validation folds to ensure stable estimates of risk model performance. We note that, even though k-fold is a cross-validation procedure that splits the sample into construction and validation parts, these processes are only completed on the 75% of subjects we identified as the construction sample, where performance metrics 9i.e. AUC values) are only computed for models reproduced in the 25% validation sample.

The final model produces selected items and coefficients that are rational numbers or 'decimal numbers', which are then converted to integers or 'whole numbers' by multiplying each by 100. The value assigned to each item is its 'weight', which is multiplied by the raw item value. For example, 'number of prior FTAs' has four values 0, 1, 2, and 3. If the model coefficients assign the weight of 5, this item would be coded 0, 5, 10, 15 for the associated responses. Items that are not selected are coded 0 for all responses.

After item weights have been established in the 75% construction sample, models scoring is computed for subjects in the 25% validation sample, which was held out to assess model performance. AUC values are computed to assess the predictive validity of the final models. The AUC values computed in the validation sample represent the new tool's estimated predictive accuracy, or the performance to be expected in future samples if the tool were to be deployed.

We included all RAI items to potentially be selected in the updated tool development. Two additional measures were also added to the item pool. Male was coded '1' for males and '0' for females. We also included youth's age at intake, which was coded '0' for 18, '1' for 17, '2' for 16, '3' for 15, and '4' for youth younger than 15.

As demonstrated in recent findings (Hamilton et al., 2023), some items are more predictive for FTAs, while others provide stronger weights for FSIs. Therefore, detention screening tools are more effective if models are developed to select and weight items for the specific outcomes they intend to predict. To increase the prediction accuracy of the updated RAI models, we computed two models, one that selects and weights items to predict FTAs and another to predict FSIs.

Updated RAI Model Findings
Items and response weights are provided in Table 13, with three columns of weights – current RAI, updated FTA, and updated FSI scoring. Findings not only demonstrate the need to weight items using local agency data, but also the importance of creating separate models for each

---

[12] We note that ridge regression is used as it is demonstrated to be ideal for data sets with high dimensionality, or those in which there is a large pool of potential predictor variables (Kigerl et al., 2022). Further, ridge is a type of 'penalized' regression model that is designed to reduce the impact of extreme weights that have a greater probability of occurring in a typical logistic regression.

outcome. Regarding local weighting, we observe how different the item scoring is by comparison to the current RAI. While the original RAI weights could be described as 'theoretical', by using Nebraska youth data to recalibrate the response values, the tool scoring is more aligned with local practice, which, in turn, will improve prediction.

As for the use of two scoring models, when comparing the variation in item weights between the two columns, we use each item as predictors for FTAs versus FSIs. For example, the item 'FTA < 12 mon.' scores for FTA but not for the FSI model and the item 'Referral/Citation Pending' scores for FSI but not FTA model. Moreover, there are several items that have stronger weights for FTA than FSI, and vice versa.

Table 13. Updated RAI FTA & FSI Scoring Models (N=2,907)

Item	RAI	FTA	FSI
Male	NA	10	45
Age at Intake			
18	NA	0	0
17	NA	5	40
16	NA	10	80
15	NA	15	120
<15	NA	20	160
Most Serious Offense			
Violent Felony	12	20	25
Violent Misdemeanor	4	20	25
Felony	6	15	25
Misdemeanor	2	25	25
Probation Violation	0	50	10
Status Offense	0	35	20
Legal Status			
Pending Court	4	30	3
Active Probation Reentry YRTC	4	25	30
Current Probation <6 mon.	3	3	10
Current Probation >6 mon.	2	12	2
Current Diversion	1	0	0
FTA/Re-Offense Risk			
Escape 24 hours	2	15	30
FTA < 12 mon.			
0	0	0	0
1	1	35	0
2	2	70	0

3+	3	105	0
Referral/Citation Pending			
0	0	0	0
1	1	0	15
2	2	0	30
3+	3	0	45
No prior	0	0	-15
Mitigating Factors			
Family/Caretaker Responsible	-1	-50	-43
First Arrest 16+	-1	-5	-5
No arrest/citation in last year	-1	-5	-20
Other	-1	-6	-1
Aggravating Factors			
Runaway	1	40	30
2+ Violent. Offenses in Last Year	1	5	30
Multiple Offenses Current Intake	1	5	13
Other	1	1	1

Using the 25% validation data set we computed the predictive validity estimates. Findings are presented in Table 14. AUC values were computed for both a) those detained for less than 72 hours (N=3,875) and b) the full sample (N=5,604). AUC values range from weak (0.54) to strong (0.75), with most identifying a moderate prediction strength (>0.64) (see Rice & Harris, 2005). Considered a 'proof of concept', this model development exercise demonstrates the improvement that can be gained by simply recalibrating the existing RAI items. To be described in greater detail, additional adjustments and item modifications may further improve FTA and FSI prediction.

Table 14. New Model AUCs by Gender and Race/Ethnicity

	Total (AUC)	Female (AUC)	Male (AUC)	White (AUC)	Black (AUC)	Other (AUC)	Hispanic (AUC)
Not detained *(<72hrs)*							
FTA	0.63	0.54	0.65	0.63	0.65	0.60	0.62
FSI	0.67	0.71	0.66	0.66	0.68	0.65	0.68
All subjects							
FTA	0.66	0.64	0.68	0.68	0.65	0.75	0.67
FSI	0.65	0.69	0.64	0.68	0.65	0.64	0.66

Conclusions

Within the juvenile justice system, the use of risk assessments is common practice, providing a standardized method for evaluating the likelihood of negative outcomes. Originating in adult probation, over the last forty years these tools have expanded and are now used in nearly every part of the adult and juvenile system, including pretrial detention. Since 1988, Nebraska Juvenile Probation has been responsible for the detention decision function, utilizing multiple generations of tools since that time. Several efforts have been made, with limited success, to validate the tool until this report.

In 2024 AOCP contracted with JJI and NCJR to complete a validation, consisting of both qualitative and quantitative portions. Specifically, we sought answers to two study questions 1) *What factors affect probation officer implementation and use of the Nebraska Risk Assessment Instrument?* and 2) *Is the RAI a valid and predictive assessment instrument?* To address the first study question, JJI researchers conducted interviews and focus groups with probation officers and managers to assess their attitudes about the value, processes, and use of the RAI.

Qualitative analyses revealed interesting, yet consistent findings. Specifically, while probation officers understood the need to provide a standardized assessment, they identified several items and areas of subjectivity that contributed to the inaccuracy of the RAI. Further training is conducted primarily at the local level, creating the opportunity for district variations in scoring and use. Moreover, probation officers indicated the pressure often provided by law enforcement to override RAI results and provide a detention recommendation. Ultimately our focus groups and interviews identified that the provision of the RAI is not a uniform process, where items, training, and use of the tool vary greatly. This combination of characteristics is common for poorly performing tools that have experienced practitioner drift and ultimately have lost stakeholder support (Burrell, 2016). Through the examination of these process evaluation findings, we then sought to determine if data collected by AOCP supported the perceptions.

Our validation findings utilized over five years of data, collecting risk scores and classification levels, FTAs and FSI outcomes, and overrides for a large sample of Nebraska youth assessed via the RAI. Findings revealed that the RAI tool was not a valid predictor of FSI or FTA, with predictive performance statistics indicating near random and often less-than-random chance. Possibly because of perceived performance issues, the RAI recommendations were often overridden and commonly used discretion to recommend youth for detention. Additional examinations indicated that the tool's performance did not vary substantially by gender, race/ethnicity, or if the youth was released, detained, or the RAI recommendation was overridden. Overall, the RAI is identified to be underperforming, and well under the field's established criteria for use with a justice population (see Desmaris et al., 2022).

As described, tools may fail for many reasons, but the best performing tools are those that are created with local data and where items are selected and weighted to produce optimal results for agency outcomes (Duwe, 2024). Given the large data set already collected using assessment development methods, we provided a 'proof of concept' by creating an updated version of the RAI. This version used more advanced development methods, selecting and weighting items for two models, one predicting FSI and another for FTA. Findings demonstrated how AOCP can

improve the RAI's performance, potentially raising predictive metrics above national standards.

Limitations

There are two notable limitations of the current study. The first is notably with the participation of individuals in the focus groups. While researchers attempted to randomly choose participants, this was not always achievable. For some districts there was simply no way to randomly select as there were only a few juvenile probation officers conducting intakes. Further, for two districts AOCP indicated that they wanted a larger sample of individuals, meaning that 3J and 4J were not randomly assigned. Moreover, there were instances where individuals from a district that were randomly assigned pointed researchers to another individual who they indicated had more information about the process. In these cases, researchers went with the substituted individual. Finally, there were a few individuals that were noted by AOCP that were important key staff that needed to be in a focus group. In this case, there may be selection bias in the focus groups as not all individuals were randomly assigned.

Second, the RAI measures are primarily measures of a youth's offense history, which have well-known accuracy limitations and may possess greater race and gender bias than tools that attempt to collect a greater number of youths needs items. Further, our six-month evaluation period includes a period where additional assessments (i.e. the YLS/CMI) are completed, and programming and interventions are often completed. These processes produce unmeasured moderating effects that may also impact RAI's prediction. Further, the outcome variable was limited to those who had an FSI or FTA within the juvenile justice system, meaning that those who had entered the adult system may not have been captured.

Finally, our proof of concept created an update to the RAI by recalibrating existing item weights to improve the tool's prediction. These items provide limited content, primarily focused on the current and prior offending record, which include items with well-known prediction accuracy and bias issues (i.e., seriousness of current offense). Further, our qualitative findings indicated that district application of items vary. Therefore, it is anticipated that greater improvements to an updated model will be found by using an updated or new set of items that are not affiliated with the current RAI scoring system and may be further examined for neutrality of prediction.

Recommendations

Following our examination of the background, use, and accuracy of the RAI, we developed a set of recommendations we hope will help guide AOCP next steps and use of their detention screening tool.

Develop a New Detention Screening Tool

1. **Take Advantage of Current Resources:** The current RAI was developed in conjunction with Annie E. Casey Foundation, which was part of a larger project effort in which the Foundation offered a set of common items used in other detention tools around the country. This initial tool was built with little data or knowledge of effective predictors for

Nebraska youth. With the understanding that tools are more accurate when created with local data, we recommend using data collected as part of RAI, and other routinely collected information, be used to develop a tool for Nebraska youth. This tool could be developed in conjunction with UNO or independently. Ultimately, our proof of concept demonstrated that a tool can be created that is optimized for Nebraska youth detention decisions.

2. **Specify Models to Agency Needs:** Noted in our validation findings, RAI focus group participants are concerned with youth risks of both FSIs and FTAs. As risk assessment development has advanced, research has identified that, while there is substantial overlap in the predictors of outcomes, some are stronger predictors of some, but not all, outcomes. As demonstrated, developing models for specific outcomes improves the predictive accuracy of developed tools. We recommend outlining outcomes and creating scoring models for each, with the potential to expand outcomes predicted (i.e., violent FSI).

3. **Reducing Assessment Labor**: Due to the expanded use of assessment tools, probation officers are often asked to complete and score multiple assessments and spend several hours completing an intake assessment. We recommend a deeper examination of assessment systems used by the state, with attempts to consolidate efforts, where responses to items collected are automatically scored or require less time to collect. Further, software applications can be developed to transfer data from one assessment to another, and even pull administrative records (i.e., offense histories) from case management systems to reduce assessment labor and increase scoring reliability.

4. **Establish Quality Assurance Practices and Plan Validation:** To ensure the effectiveness of a tool, quality assurance processes need to be completed to ensure the assessment is being completed as intended. Implementation research has outlined processes such as randomized case audits, semi-annual interrater reliability analyses, and routine validations are recommended practices to be implemented routinely with timing dictated by policy. Further, when assessment administrators (i.e., intake officers) are held accountable for scoring accuracy and limitations are placed on overrides, greater tool use and accuracy is observed.

Training Recommendations

1. **Statewide Standardization and District Representation:** Training should emphasize both accurate scoring and procedural consistency. To ensure alignment across districts, each district should continue to appoint a representative to serve on a statewide subcommittee responsible for developing unified policy and training guidelines. This committee should promote standardized practices and ensure all districts are equitably represented.

2. **Enhanced Initial Training for New Probation Officers:** Return to an in-person training for NPOT that includes structured, scenario-based learning tailored to real-world challenges encountered during intake and detention decision-making.

3. **Use of Real-Life Scenarios:** Continue to incorporate complex, real-life case scenarios into statewide booster sessions or communities of practice to deepen learning and improve decision-making consistency across districts.

4. **Improved Integration Between State and Local Training Efforts:** To strengthen the coordination between New Probation Officer Training (NPOT) and other state-level training and local booster sessions, a subcommittee could help design consistent scenario-based training modules and quarterly test a new scenario to enhance inter-rater reliability.

5. **Ongoing Stakeholder Engagement and Training:** Provide regular, targeted training for all stakeholders involved in the detention process, with a particular focus on law enforcement, to assist with custodial issues and exchange of information. It is also important to build and maintain relationships through participation in local and statewide venues such as law enforcement conferences, training centers, and police academies to ensure shared understanding and alignment with the tool's purpose.

Systemic Recommendations for Improving the Intake Process

1. **Warrants:** The current intake process is hindered by a lack of clarity and standardization regarding warrants. It is recommended that the AOCP examine the types of warrants being issued and their implications for intake decisions. Specifically, frequency of not being able to implement the screening decision should be examined, with the goal of developing clear guidance for how different warrant types should be considered during intake. Additionally, a workgroup could examine discrepancies between statutory guidance and current practice to promote greater alignment and fairness.

2. **Discrepancies in Statute Versus Practice for Warrants:** There appears to be misalignment between what Nebraska statute permits/requires and how warrant-related decisions are implemented in practice. A legal and procedural review should be conducted to identify these inconsistencies and ensure that intake decisions reflect statutory intent. Clear guidance and training materials should then be developed to ensure that all stakeholders—including judges, probation officers, and law enforcement—understand the legal framework and how to apply it appropriately.

3. **Evaluation and Strengthening Alternatives to Detention (ATDs):** A comprehensive review should be conducted to assess the availability, quality, and utilization of ATDs across jurisdictions. If gaps exist in specific geographic areas or service types, targeted investments and support may be needed to expand access. The review should also explore reasons for underutilization, such as eligibility restrictions, lack of awareness among practitioners, or stakeholder perceptions regarding program effectiveness.

4. **Mental Health and Crisis Response Alternatives:** One significant gap identified in the intake process is the lack of accessible, community-based mental health or crisis intervention services for families. The youth may have caused substantial harm within the family, making it difficult for them to return home without intentional intervention.

General Recommendations Related to Collaboration with Key Stakeholders

1. The AOCP should compile data on the average length of time required to complete an intake, number of intakes per month, and other relevant data to share with law enforcement partners. This may assist stakeholders in observing an overall picture versus a focus on specific cases.
2. Probation officers should meet with both law enforcement administration and members of the judiciary on a regular basis to discuss opportunities to improve the processes.
3. Local probation staff and law enforcement should collaboratively bring system issues (i.e., transportation & ATD needs) to their local community planning collaboratives to foster broader response and support.

Acronyms

1. AOCP – Administrative Office of the Courts and Probation
2. ATD - Alternatives to Detention
3. AUC – Area Under the Curve
4. FSI – Future System Involvement
5. FTA – Failure to Appear
6. JJI - Juvenile Justice Institute
7. NCJR – Nebraska Center for Justice Research
8. PO – Probation Officer
9. RAI – Risk Assessment Instrument

Works Cited

Berk, R. (2019). Accuracy and Fairness for Juvenile Justice Risk Assessments: Accuracy and Fairness for Juvenile Justice Risk Assessments. Journal of Empirical Legal Studies, 16(1), 175–194.

Berk, R., Heidari, H., Jabbari, S., Kearns, M., & Roth, A. (2021). Fairness in Criminal Justice Risk Assessments: The State of the Art. Sociological Methods & Research, 50(1), 3–44. https://doi.org/10.1177/0049124118782533

Burrell, W. D. (2016). Risk and needs assessment in probation and parole: The persistent gap between promise and practice. In *Handbook on Risk and Need Assessment* (pp. 39-64). Routledge.

Desmarais, S. L., D'Amora, D. A., & Tavárez, L. P. (2022). Advancing Fairness and Transparency: National Guidelines for Post-conviction Risks and Needs Assessment. US Department of Justice, Office of Justice Programs, Bureau of   Justice Assistance.

Duwe, G. (2024). Evaluating bias, shrinkage and the home-field advantage: Results from a revalidation of the MnSTARR 2.0. *Corrections*, *9*(1), 20-42.

Hamilton, Z., Kigerl, A., & Kowalski, M. (2020). Optimizing juvenile assessment performance. CrimRxiv.

Hamilton, Z., Kigerl, A., & Kowalski, M. (2022). Prediction is Local: The Benefits of Risk Assessment Optimization. Justice Quarterly, 39(4), 722–744.

Kigerl, A., Hamilton, Z., Kowalski, M., & Mei, X. (2022). The great methods bake-off: Comparing performance of machine learning algorithms. *Journal of Criminal Justice*, 82, 101946.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI, 14(2), 1137–1145.

Neb. Rev. Stat. § 43-253(3) (2024)

Neb. Rev Stat. sec. 43-250 (1)(c) (2024)

Moore, S., & Hobbs, A. (2015). Analysis of the Nebraska Intake Risk Assessment Instrument 2015. *University of Nebraska Juvenile Justice Institute*.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior, 29(*5), 615–620

Sullivan, C. J., & Childs, K. K. (2021). J*uvenile risk and needs assessment: Theory, research, policy, and practice*. Routledge.

Taxman, F. S. (2018). The Partially Clothed Emperor: Evidence-Based Practices. Journal of Contemporary Criminal Justice, 34(1), 97–114

Taxman, F. S., & Dezember, A. (2017). The value and importance of risk and need assessment (RNA) in corrections & sentencing. In F. S. Taxman (Ed.), Handbook on risk and need assessment: Theory and practice (pp. 1–20). New York, NY: Routledge.

Vincent, G. M., Guy, L. S., & Grisso, T. (2012). *Risk assessment in juvenile justice: a guidebook for implementation.*