



Department of Community Justice

MULTNOMAH COUNTY OREGON

Adult Services Division
Pretrial Services Program

1120 SW 3rd Avenue, Rm. 358
Portland, Oregon 97204
(503) 988-5042 Phone
(503) 988-4157 Fax
<http://www.co.multnomah.or.us/dcj/>

July 6, 2010

To : Bill Penny, District Manager

From: Don Trapp, Community Justice Manager

Re : Norming the Virginia Pretrial Risk Assessment for application in Multnomah County.

In 2009, the Pretrial Supervision Program (PSP) began utilizing the Virginia Pretrial Risk Assessment Instrument (VPRA) to assist in conducting release investigations. This instrument was developed to establish a standardized, validated risk tool that could be utilized across the state of Virginia. The rigorous methodology employed in the development of this instrument^[1] have made it a model for pretrial service agencies.

The instrument consists of 9 items that cover standard pretrial release criteria.^[2] The items are equally weighted (1 point), with the exception of one item (prior FTA's) that is worth 2 points. In the validation study, increasing scores corresponded to increasing failure rates (see Appendix A). Although the general criteria are almost identical across pretrial assessments, the weighting, the metrics within each item, and the inclusion/omission of items do vary across jurisdictions. In addition, there are differences in how the instrument is utilized. Specifically, PSP is the second tier in the pretrial process in Multnomah County, as all the defendants have previously been screened through another process and have already made a first court appearance. The Virginia tool is designed as the initial screen prior to the first court appearance. For the initial screen, some factors exist as sorting mechanisms such as: level of current offense

(felony/misdemeanor) and prior violent offenses. The purpose of this study is to examine how this tool performs with regard to these differences and specifically in Multnomah County.

Since implementing the VPRA, the Pretrial Supervision Program has been collecting data related to the performance of pretrial defendants with the purpose of conducting a norming process on the VPRA. This is the process to assure that an assessment instrument address regional difference in regards to: citing/arrest practices, booking and sentencing practices, and the differences with regard to the ethnic make-up of the region. The results of this process could recommend some modifications to the use of the tool specifically in Multnomah County. The purpose may be summarized in the following questions:

1. Does the Virginia Pretrial Instrument accurately assess risk of pretrial misconduct in Multnomah County?
2. Do the individual items on the instrument make valid discriminations between low-high risk defendants?
3. Does reliance on the Virginia Pretrial Instrument have any adverse impact on minority overrepresentation in Multnomah County?

Section 1: Descriptive Statistics:

Methodology:

The sample was 641 defendants who were referred to PSP for either release investigations or pretrial supervision during the period March – September 2009. The VPRA was completed when the case was received by PSP. Each case was then tracked to resolution (sentencing, FTA, Rearrest, Revocation).

The sample was predominantly male (80%), and the majority (58%) were defendants awaiting trial on felony charges. Figure 1.1 provides a summary of the descriptive statistics compared with the sample used in the Virginia study. The data indicate significant differences in almost every category except gender. These differences illustrate the need to examine the impact of this instrument on the population in Multnomah County.

Figure 1.1 Comparison of samples from Virginia and Multnomah County

Variable	Virginia Study (n = 1971)	Multnomah County (PSP) Study (n = 641)
Gender	78% male	80% male
Age	31.03 avg (10.15 SD)	37.27 avg (11.83SD)
Ethnicity—White	40%	62%
Ethnicity—Black	58%	28%
Residence—less than one year	39%	77%
Employment—during last two years	44%	13%
Current charge—felony	34%	58%
Charge type—drug	11%	23%
Charge type—person	23%	32%
Charge type—property	17%	29%

Figure 1.2 illustrates the release status showing that 40% of defendants in the sample were released pending trial. While detained defendants were all subject to PSP's release decision-making process, this was not true for all defendants who were released. Defendants released by the court, or on bail may not have had completed investigations.

Figure 1.2

RELEASE STATUS (n = 641)

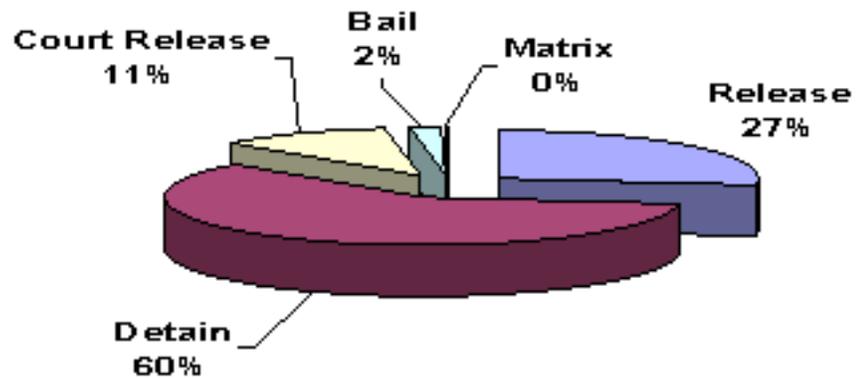


Figure 1.3 illustrates the release and detention rates by ethnicity and gender. Figure 1.4 shows the outcomes of release. Technical violations are included as a closure type as these defendants are returned to secure custody. While generally counted as a negative closure, they may be considered a proactive response to inappropriate behavior by the defendant, which could include contact with the victim or one of the state's witnesses. Release outcomes did not vary by release type. Figure 1.5 illustrates the variance in the average time to outcome. The data contrast with widely held views and Bureau of Justice Statistics Data [\[3\]](#)

Figure 1.3

Percentage Released and Detained by Ethnicity and Gender (n = 638)

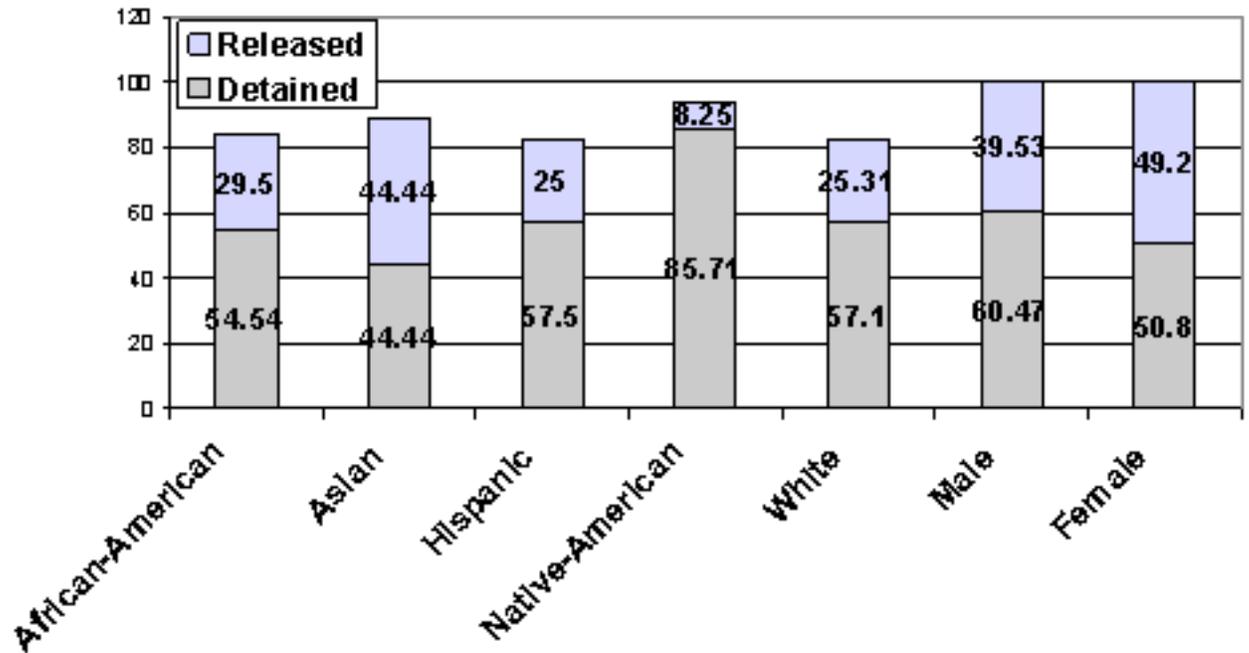


Figure 1.4

Release Outcomes (n = 256)

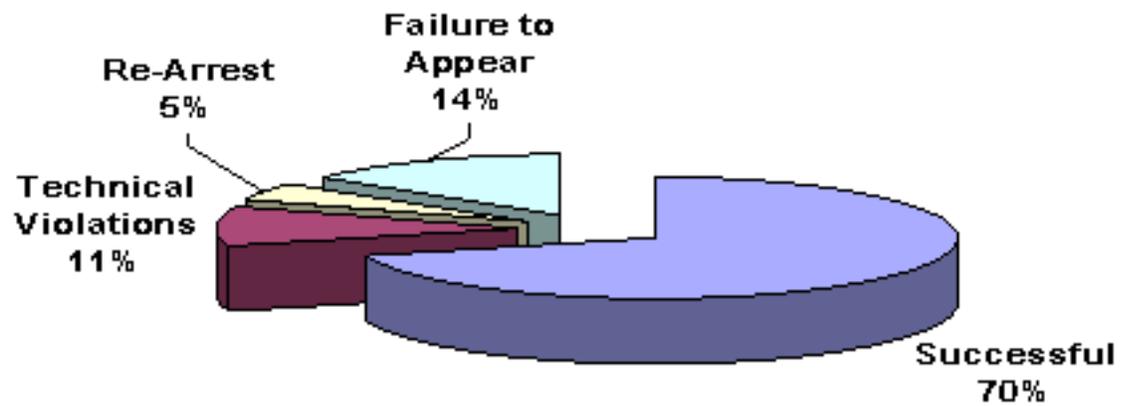
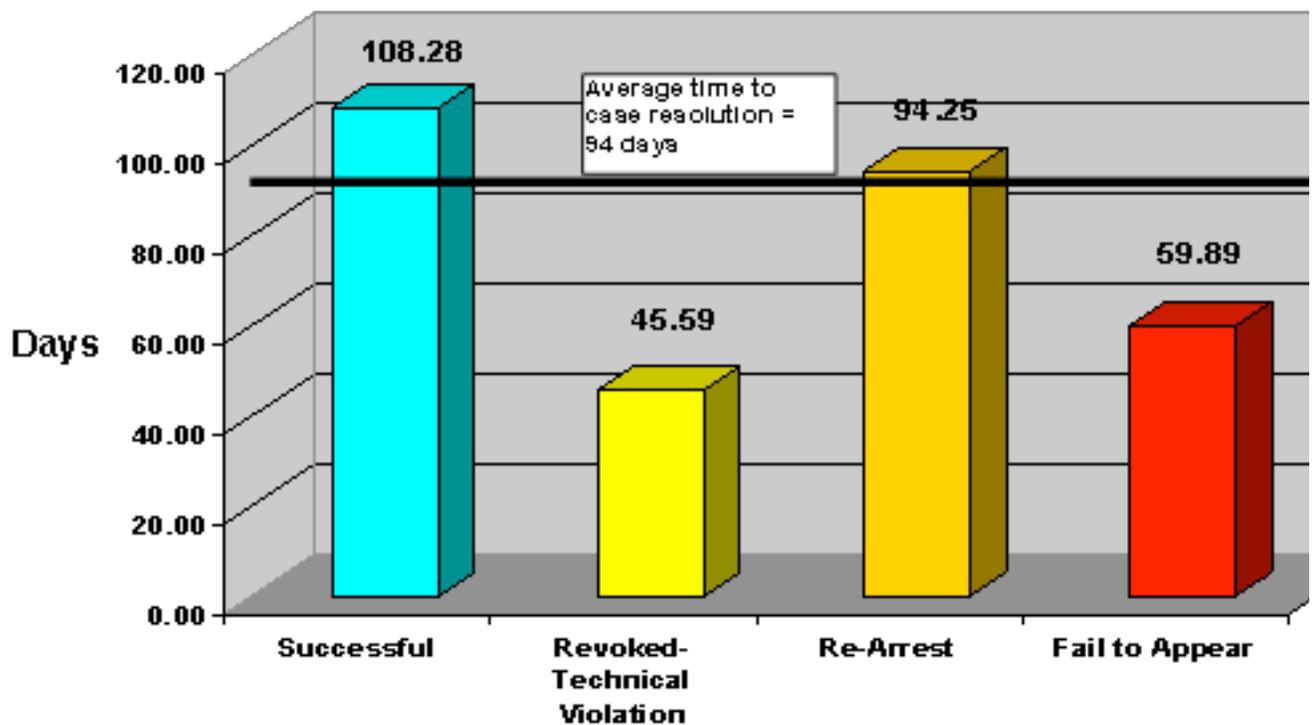


Figure 1.5

Average Time to Case Resolution (n = 256)



In terms of PSP's decision-making process, the analysis indicated that:

- PSP has a release rate of 33% (185 of 561)^[4]
- Of those released, 66% (122 of 185) were successful (70% in total released sample, 180 of 256)
- An additional 12% (23 of 185) were revoked for technical violations (11% in total released sample, 27 of 256)
- The remaining 22% (40 of 185) engaged in some form of pretrial misconduct (19% of total released sample, 49 of 256).

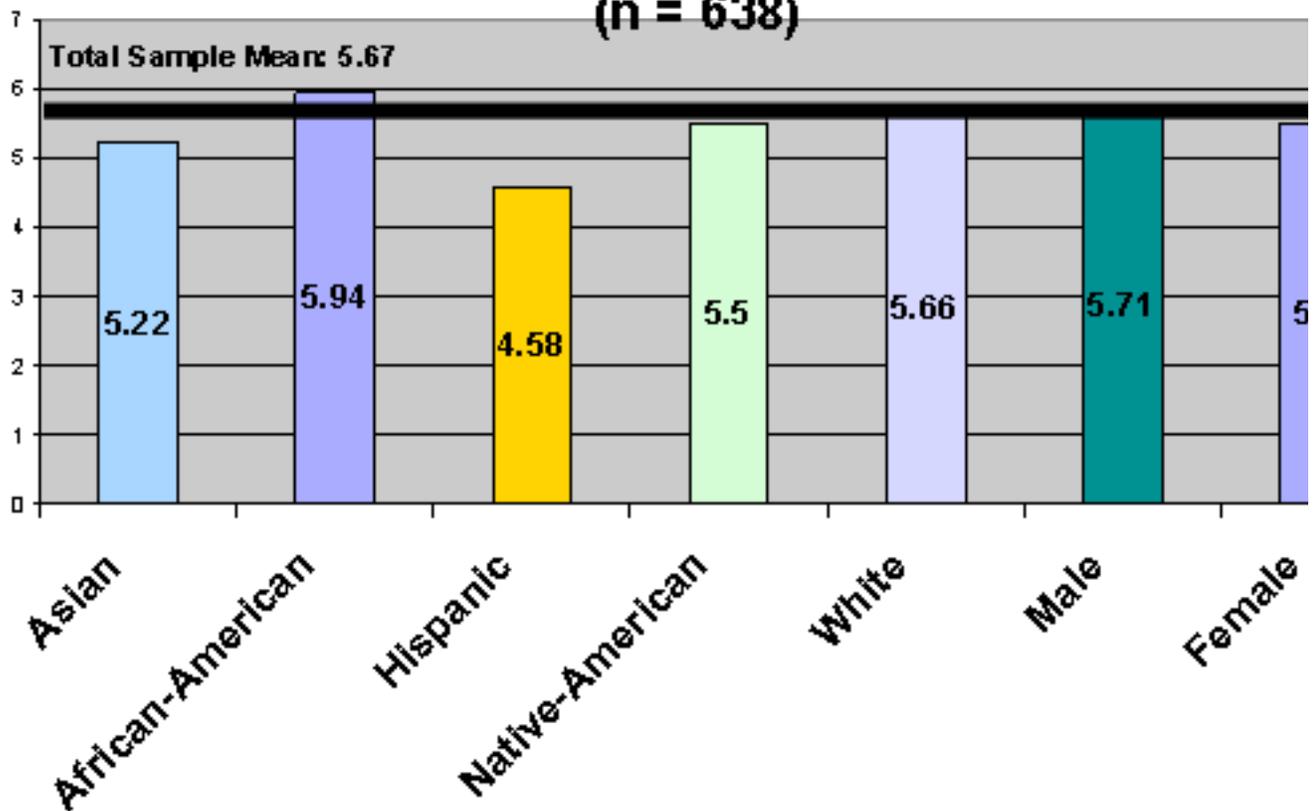
The correlation between the item scores and outcome was .31. As the succeeding graphics will illustrate, the relationship between VPRA score and outcome took on a parabolic distribution. This was probably due more to the selection process than the tool, as defendants who scored higher on the VPRA were less likely to be released. For example, only 3 of 22 defendants (13.6%) who scored 9, and 12 of 84 (14.28%) who scored an 8 were released. The total release rate for the sample was 39.9%.

For the total sample, the average risk score was 5.67 with a standard deviation of 1.91. The following graph (Figure 1.6) illustrates the VPRA scores by ethnicity. Significant differences were found with the Asian and Hispanic population. Scores for African-Americans, Native Americans, and Whites were roughly equivalent.

The VPRA scores were compared with scores from the Pre-trial Release Screening Instrument^[5] In the application of this tool, there are a number of administrative policies requiring either overrides or the awarding of points to move a defendant above the "release cutoff score". This would confound efforts to conduct a simple correlation between the two scores. Such a correlation was conducted yielding a Pearson's *r* of .208. When the overridden cases were removed, the correlation increased to .4. Additionally, the correlation of only the overridden cases with the VPRA was also .4.

Figure 1.6

Mean VPRA Scores by Ethnicity and Gender (n = 638)



Risk scores also varied by crime type (Figure 1.7). However, the average risk scores for felony and misdemeanor crimes were not significantly different. Of note is that Item 1 on the VPRA assigns weight (assumes increased risk) for pending felony offenses. Person-to-person and traffic crimes had the lowest average risk score and the lowest incidence of pretrial misconduct (24% and 13% respectively), while drug and property crimes had the highest average risk scores and incidence of misconduct (39% and 30% respectively).

The risk score did vary significantly for defendants with different outcomes (Figure 1.8). As indicated in the following graphic, successful defendants had the lowest average risk score (4.24), followed by technical violators (5.25). Defendants who were rearrested pending trial had the highest average risk score (6.16). The average scores of unsuccessful defendants did not vary from one another, but the differences between successful and unsuccessful defendants were significant ($p = .000$). The results indicate the predictive ability of the VPRA and that PSP staff are accurate and consistent in their use of the tool.

Figure 1.7

AVERAGE RISK SCORE BY CRIME TYPE (n = 641)

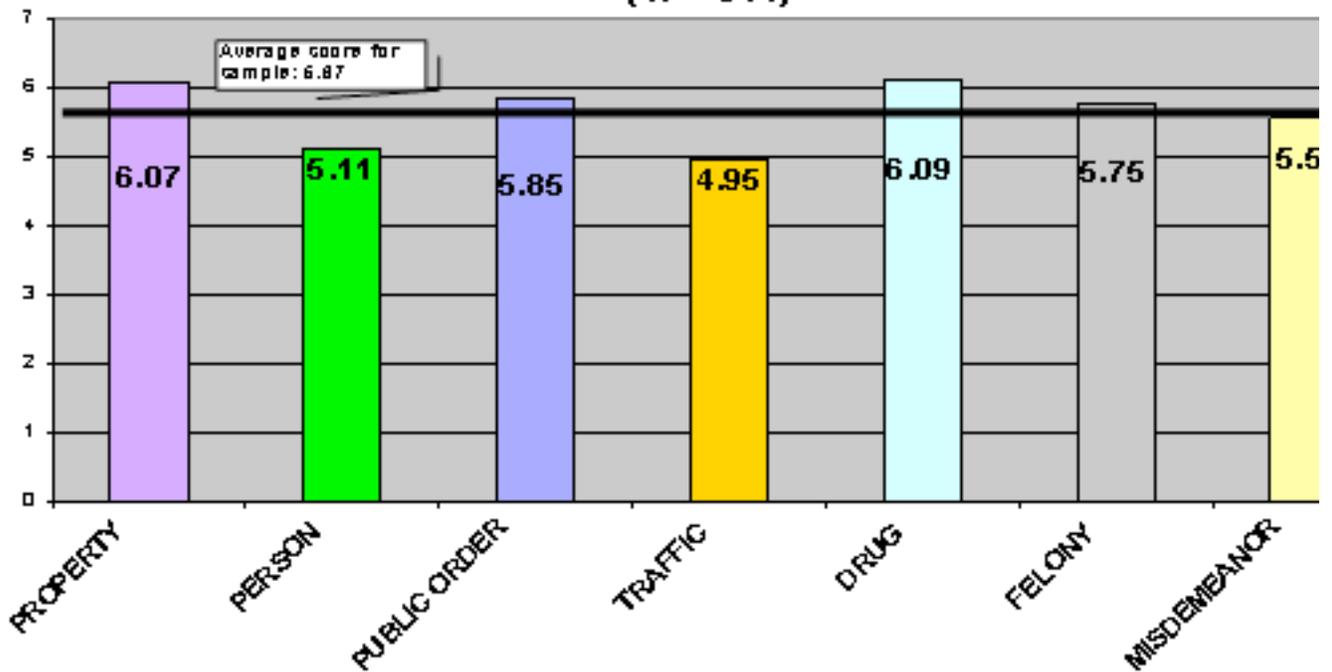
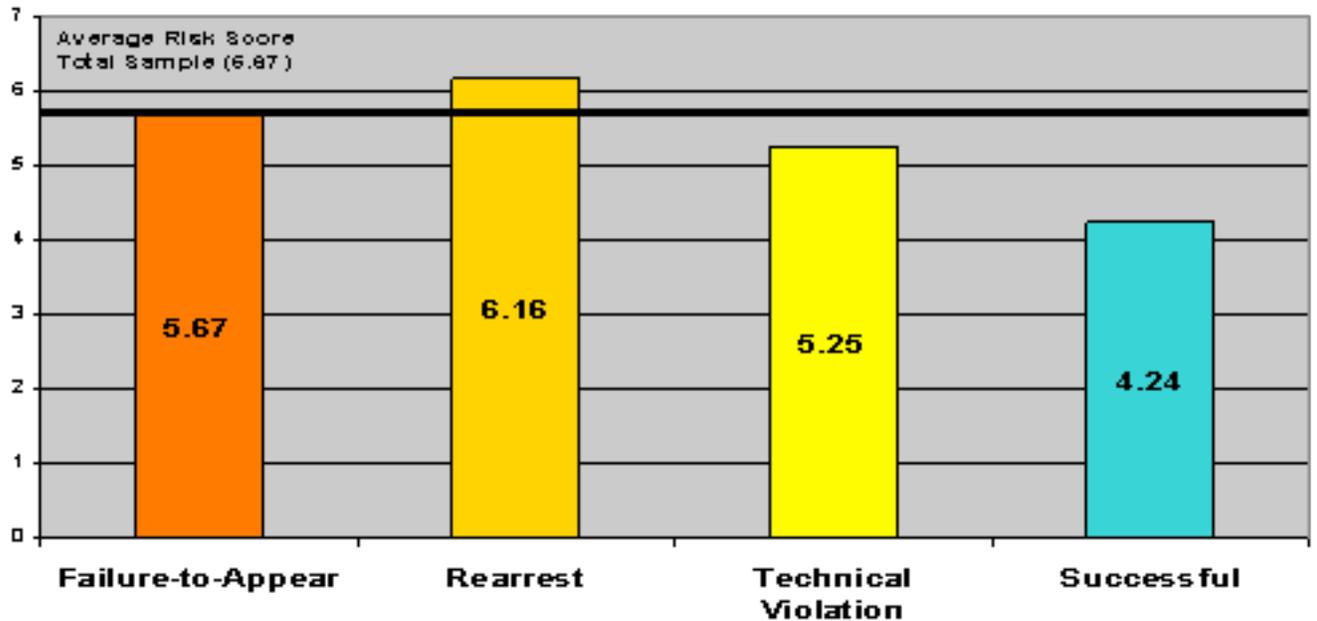


Figure 1.8

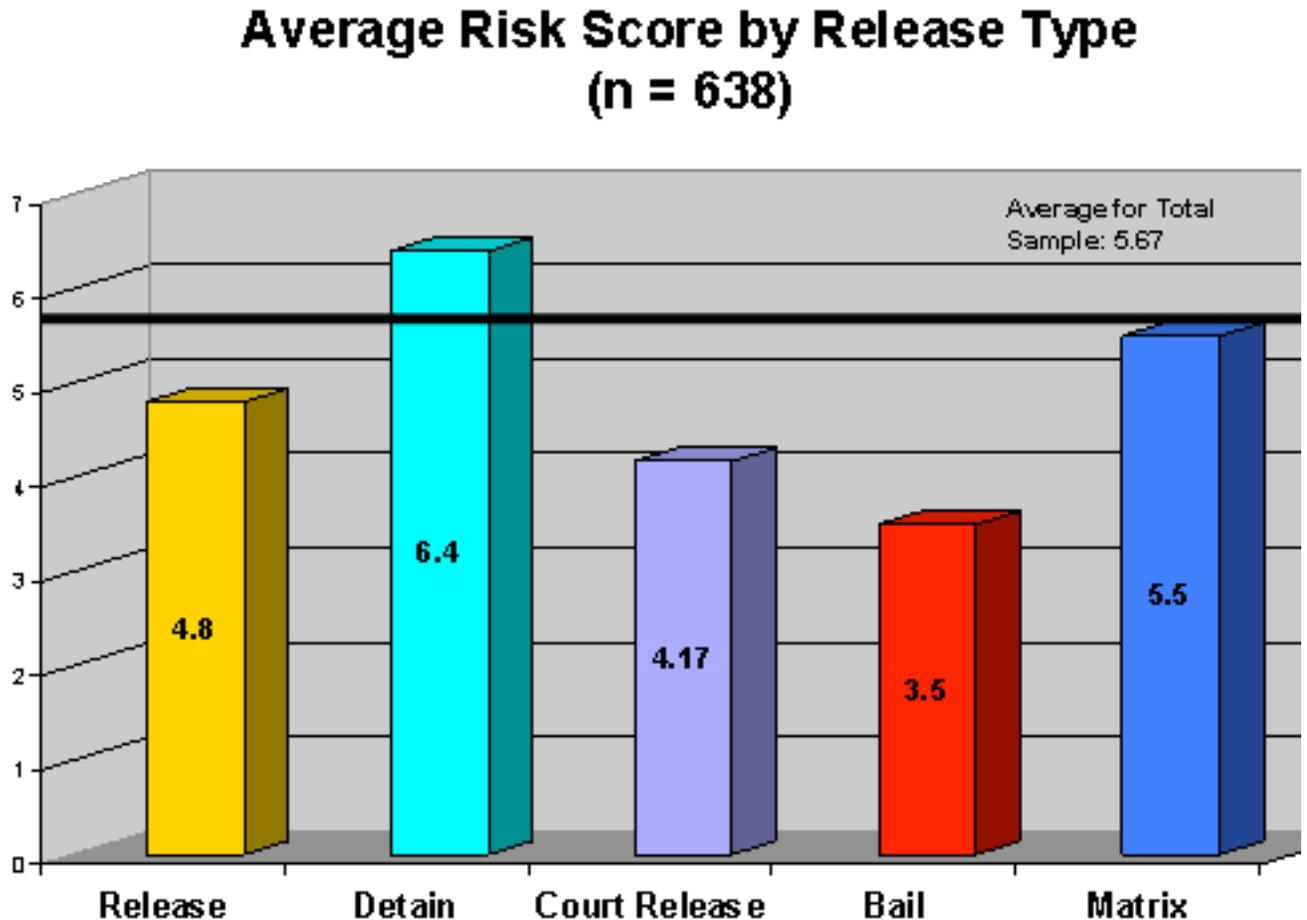
Average Risk Score by Outcome (n = 641)



Risk scores also varied significantly by release type as illustrated below (Figure 1.9). Defendants who were released onto supervision after posting bail had the lowest

average risk (3.5). Defendants who were matrix released or detained following a PSP investigation had the highest average risk (5.5 and 6.4 respectively).

Figure 1.9



The ability of the VPRA to drive release decision-making and predict pretrial misconduct is illustrated in the following graphs. Figure 1.10 illustrates the distribution of risk scores. While the distribution is approximately normal, the distribution of released subject is heavily skewed in that defendants with very higher scores were not and probably could not be released. Figure 1.11 illustrates the release and success rates by risk score, and demonstrates that PSP staff utilize the VPSA in their decision-making. Moreover, that the release decisions are appropriate. It should be noted that the sample employed here was not a stratified, random sample, as that kind of research could not be employed.

A clear trend is evident in terms of increasing incidents of misconduct with increasing scores. The overall correlation of the instrument to outcomes was $r^2 = .31$. Failure was not linked to ethnicity, type of release, or whether the pending charge was a felony or misdemeanor. The number of technical violations resulting in revocation provides an interesting dynamic to assessing the risk instrument. As in the other forms of pretrial misconduct, the challenge will be to develop supervision strategies to address

these behaviors, and their potential, allowing defendants to continue to remain in the community pending trial.

Figure 1.10

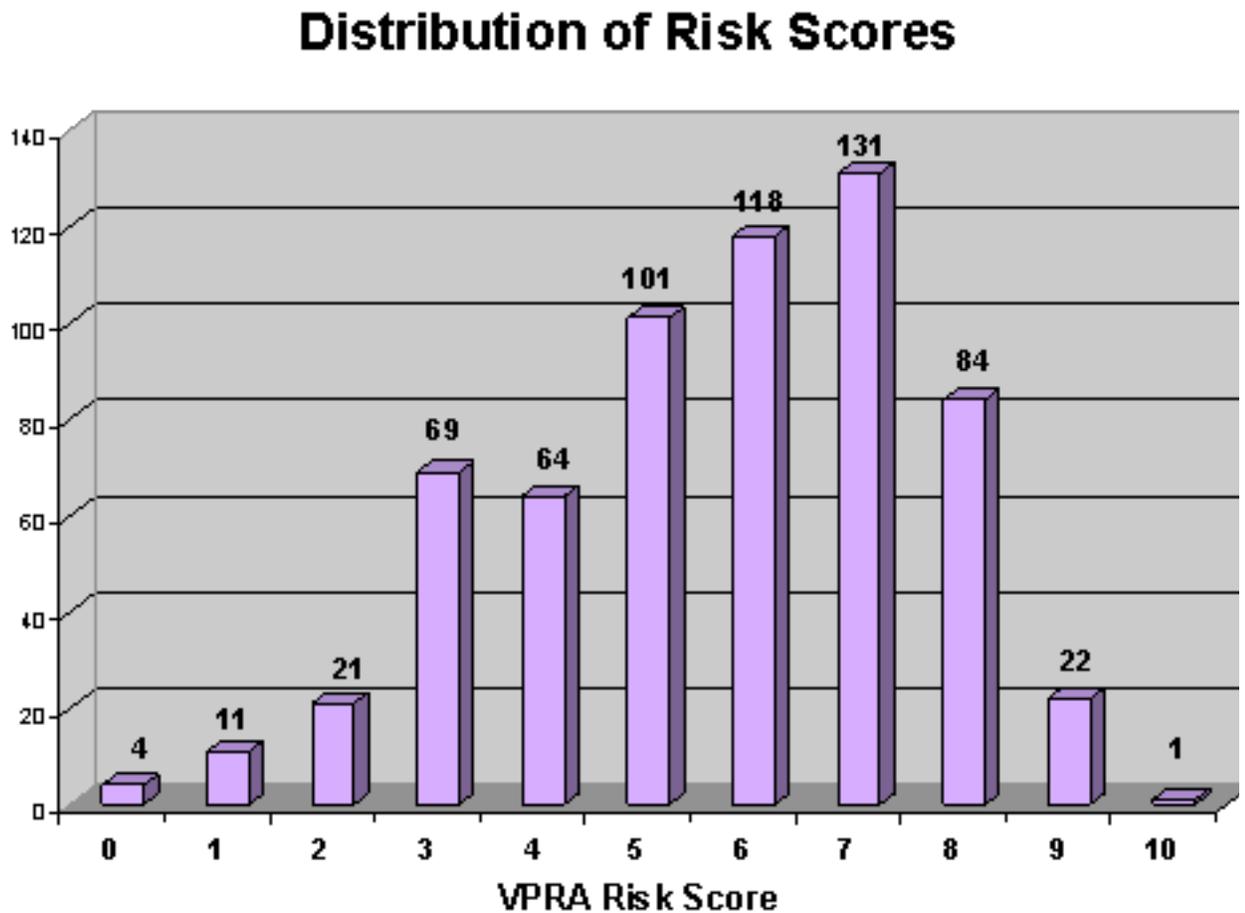


Figure 11 illustrates the failure rate by type across risk scores. In this graph, the incidents of technical violations are seen in lower-risk cases (scores 2 – 3), then increasing through scores 5 – 8. While not technically pretrial misconduct, these violations do present supervision issues both in terms of responding to non-compliance and addressing public safety issues such as contact with the victim or the state’s witnesses.

Figure 1.11

SUCCESS AND DETENTION RATES BY RISK SCORE

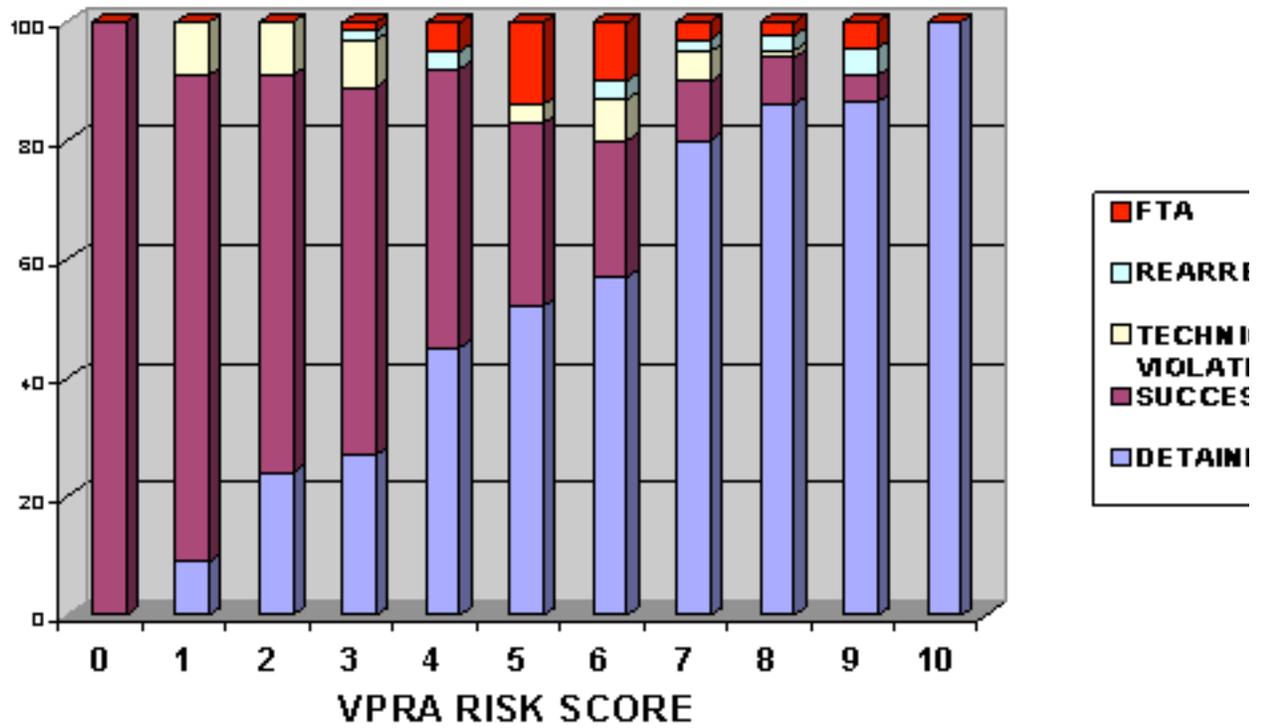
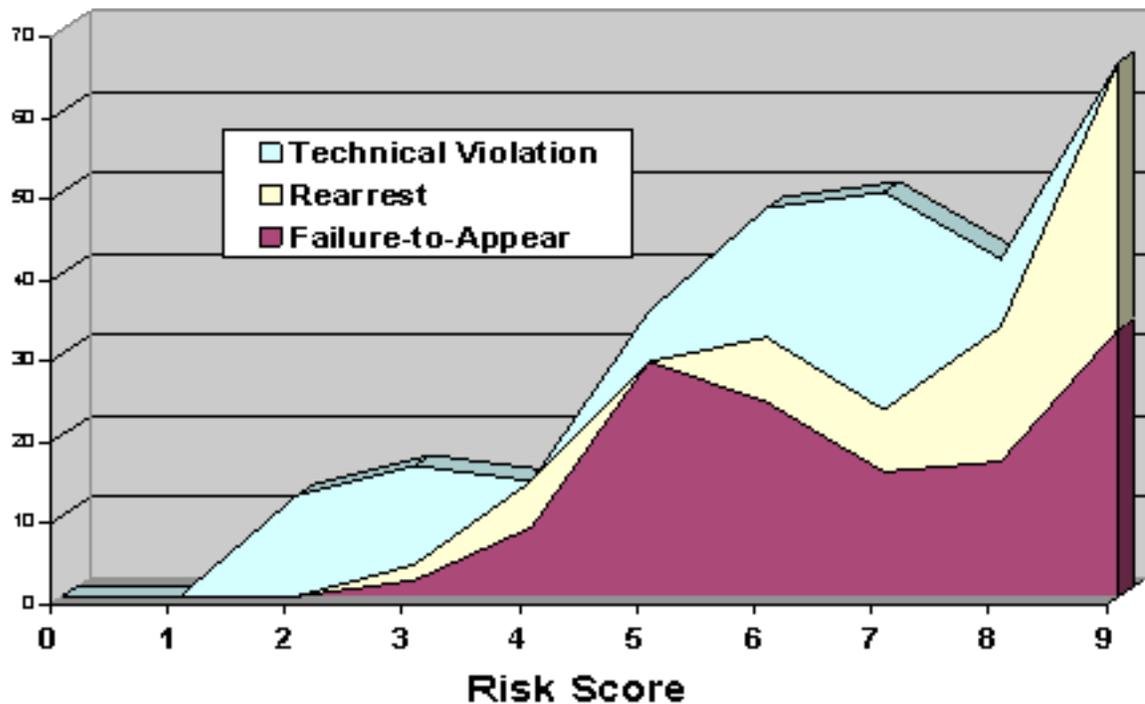


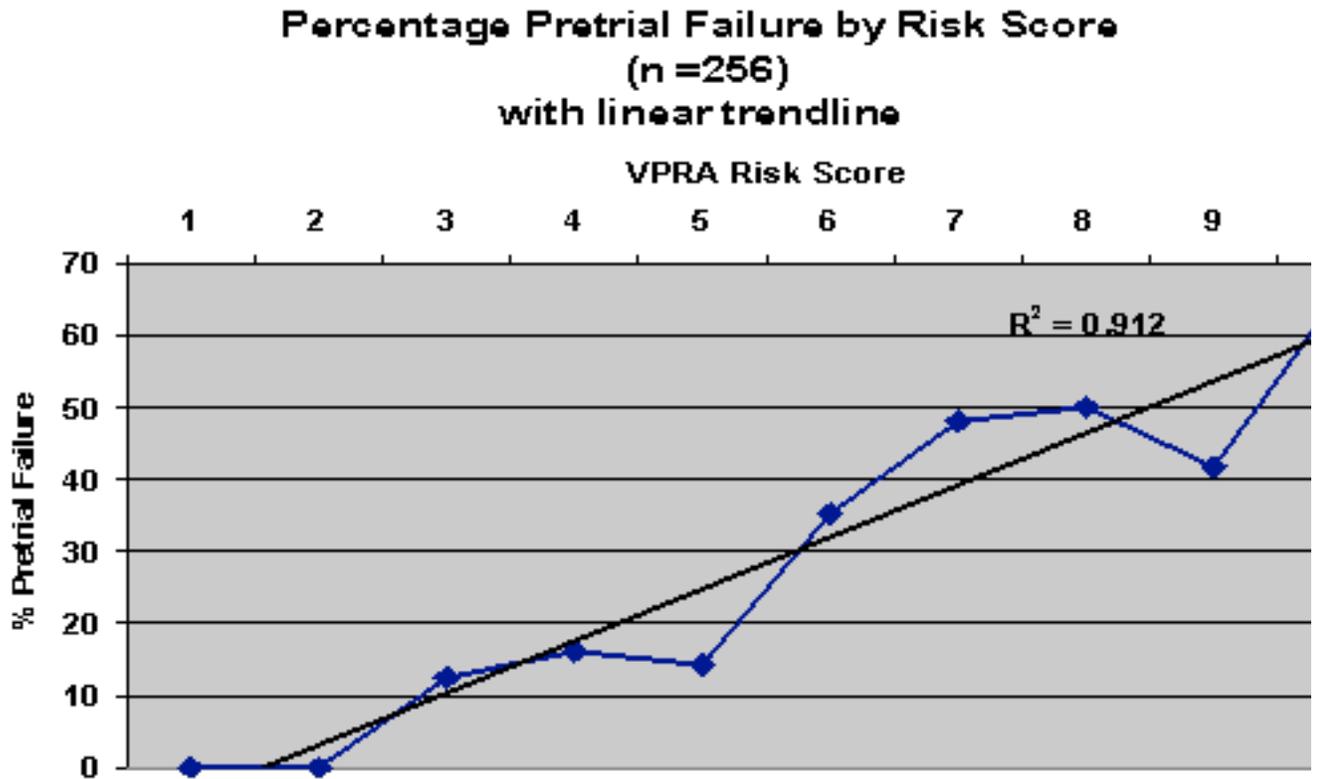
Figure 1.12

Failure Rate by Type of Failure and Risk Score



In summary, the data indicate that PSP staff utilize the VPRA to guide release decisions. The VPRA appears to make valid discriminations between low and high-risk defendants. Additionally, the predictive power of the VPRA extends to all areas of pretrial misconduct, specifically technical violations. This is an area of pretrial misconduct not normally addressed in pretrial assessments, although it remains a significant issue in terms of supervising pretrial defendants in the community. Table 1.13 illustrates the percentage of pretrial failure (including FTA, rearrest, and technical violations) by risk score. The trendline shows that for each point on the VPRA, there is a corresponding 7.2% increase in the incidence of pretrial misconduct. The associated r^2 equals .91.

Table 1.13



Section 2: Examining the VPRA Items for predictive power and discriminability

The items of the VPRA were examined with respect to their correlation with the outcome (success/failure) utilizing a 2 X 4 contingency table and ethnicity. A Chi-Square test of independence was conducted along with Cramer’s Phi (Φ) to assess the amount of variance explained. In addition, a measure of item discrimination, the *point-biserial correlation*^[6] was calculated to assess the degree to which an item discriminates between high and low scoring defendants. A summary of the results are detailed in Table 2.1

Table 2.1

Item Number	% Yes Total	% Yes Unsuccessful	χ^2	P	R_{pbis}
1	63.3	32.7	2.29	.514 d.f. = 3	.13
2	27.7	42.3	10.5	.015	.44
3	5.5	50.0	4.47	.215	.26
4	71.1	34.6	7.95	.047	.55
5	42.5	46.8	26.9	.000	.42 (adjusted) ^[7]
6	16.4	28.6	3.86	.276	.32
7	60.15	33.76	5.19	.158	.278
8	79.3	31.5	2.79	.425	.20

9	60.54	36.12	10.00	.019	.394
----------	--------------	--------------	--------------	-------------	-------------

Bolded items are significant.

The results indicate that 5 of the 9 items on the VPRA are essentially independent of the outcome of pretrial supervision. The remaining 4 items, do have an association and shared variance, and 5 items provide satisfactory levels of discrimination between high and low scoring defendants.

Item #1(Is current offense a felony?): Item # 1 and outcome are independent. While the scoring would suggest that defendants not currently charged with felonies pose less risk of pretrial misconduct than those charged with felonies, the data indicate otherwise. In this sample, 74.4% of defendants charged with felonies were successful on pretrial supervision compared with only 63.5% of defendants charged with misdemeanors.

A recalibration of the VA risk score, omitting Item #1, was completed and correlated with the outcome measure. The results indicated that the coefficient of correlation decreased from .30 to .20, suggesting that the covariance between Item 1 and rest of the items contribute to the overall efficacy of the instrument. The point-biserial correlation of .13 indicates that the item is a poor discriminator (the generally accepted range in between .3 and .7).

Item # 2 (Were there one or more pending charges at time of arrest?):

The results indicate that there is a statistical relationship between Item #2 and outcomes. The percentage scoring “yes” in the sample of those released and successful (22.7.0) is significantly less than in the percentage of those released and unsuccessful (42.3). The point biserial correlation of .44 indicates the item is a good discriminator as well as a good predictor.

Item #3 (Was there an outstanding warrant in another locality at time of arrest?):

The results indicate that Item # 3 is essentially independent of outcome. Of the defendants who scored “yes”, 50.0% were successful. However, only 5.5% of the released population scored “yes. The point biserial correlation of .26 indicates the item is approaching an acceptable level of discriminability, but fails to meet the threshold of significance.

Item # 4 (Were there one or more prior convictions?): The results indicate that Item #4 has a statistical relationship with outcome. A higher percentage of unsuccessful defendants scored “yes” than successful defendants. and ethnicity. have a dependency, although the shared variance is small. Of the defendant’s who were released, 71.1% had a prior conviction. However, only 34.6% of those who engaged in pretrial misconduct scored “yes” on this item, indicating that the statistical link to outcome is negative. The point biserial correlation indicates that item #4 is a good discriminator.

Item # 5 (Were there two or more prior FTA’s?): Item #5 is significantly linked to outcome with over 10% shared variance. The item is also a valid discriminator.

Item # 6 (Were there two or more violent convictions?):

Item #6 and outcome are independent. The total frequency of affirmative responses to this item is low, and for two categories of misconduct, the frequencies fall below the accepted standard (5) for use of the Chi-square statistic. However, even when these variables are collapsed they remain independent. The point-biserial correlation is also below significance, indicating the item is not a good discriminator.

Item # 7 (At current residence less than one year?): Item # 7 and outcome are independent. Of the defendant’s who scored “yes” on this item, only 34% were unsuccessful. The variances observed in each category of outcome are insignificant. The item also fails to discriminate at a significant level.

Item # 8 (Employed continuously for the past two years?): Item # 8 and outcome are independent. While it appears that a higher percentage of defendants who are rearrested or fail to appear score “yes” than others, the difference lacks significance. The high response rate found in the sample renders this item a poor discriminator.

Item # 9 (Was there a history of drug abuse?): Item #9 and outcome share a dependency, albeit slight. The differences in affirmative responses for unsuccessful dependents is significantly higher than those who were successful. The item essentially approaches the threshold of acceptable discriminability

Section 3: Impact on minority overrepresentation

The VPRA items were examined with regard to ethnicity using a 2 X 5 contingency table with ethnicity and the item score as the variables. For the purpose of this analysis, disparate impact will be defined as an item that:

- Is independent of outcome
- Fails to adequately discriminate between low and high risk scores
- Shares variance with ethnicity

An item that is linked to outcome may be considered predictive of risk, statistical associations with other factors (age, ethnicity, gender) may provide interpretive value. However, an item that is only linked to ethnicity should be examined more closely.

Table 3.1

Item Number	X ²	P
1	16.6	.002 (d.f. = 4)
2	2.06	.725
3	3.6	.463
4	23.6	.000
5 ^[8]	13.3	.010
6	9.46	.051
7	10.6	.031
8	10.9	.027
9	4.12	.390

Bold items are significant

The results indicate that 3 of the 9 items are independent of ethnicity. While 6 of the items share variance with ethnicity, the effects varied widely between items.

Item #1(Is current offense a felony?): Item # 1 and ethnicity have a dependency. The data indicate significant differences in terms of the frequency of positive responses by ethnicity. For example, 67% of African-American defendants had pending felony charges compared to 54% for white and 22% for Asian.

Item # 2 and #3: Results not significant

Item #4 (Were there one or more prior convictions?): Item # 4 and ethnicity have a dependency, although the data indicate the source is limited to the low frequency of affirmative responses for Hispanic defendants (56%) as compared to 85% for white defendants.

Item # 5 (Were there two or more prior FTA's?): Item # 5 and ethnicity have a dependency. The frequencies of affirmative responses are quite varied. Hispanic defendants had the lowest frequency of affirmative responses (36.58) followed by Native Americans (50%). Both of these were below the overall frequency. African Americans and Asians had the highest levels (66% and 67% respectively), both well above the overall frequency of 56%.

Item # 6 (Were there two or more violent convictions?): Item # 6 and ethnicity have a dependency. The effect is small and must be viewed with respect to the low frequencies in several cells. African American defendants had the highest frequency (31%) compared to white defendants (23%) and Asian defendants (11%).

Item # 7 (At current residence less than one year?): Item # 7 and ethnicity have a dependency. The data indicate that white and Hispanic defendants are the least stable in terms of residence (81% and 75% scoring yes respectively) with Asian defendants reportedly the most stable (55% scoring yes).

Item #8 (Employed continuously for the past two years?): Item 8 and ethnicity have a dependency. The variability between the categories was significant, with white defendants reporting the least engagement in employment (90% yes) and Hispanic defendants the most (78% scoring yes).

Item # 9 (Was there a history of drug abuse?): Results not significant

Analysis

The purpose of the analysis was to determine how each item: contributed to the prediction of pretrial risk, facilitated discriminations between high and low risk defendants, and did not adversely impact defendants based on ethnicity. The data indicate that the items varied along these criteria:

- Correlation to outcome (5 out 9 items were significantly linked to outcome);
- Ability to discriminate between high and low scale scores (5 out of 9 items met the threshold of discriminability)
- Adverse impact based on ethnicity (6 of 9 items were significantly linked to ethnicity).

Based on the results, the items can be organized into three categories (see Table 3.2): valid, invalid, and poor performing. The Valid Items include items 2,4,5, and 9. These are appropriately linked with outcome and are good discriminators. The linkage that two of them (Items # 4 and #5) have with ethnicity is a concern; however it is a concern that rests in how this dynamic is organized by ethnicity as opposed to concerns of disparate impact. Ethnicity is complex as a variable and may covary with subcategories such employment, education, and housing. These items and their associated metrics and business rules should remain intact.

Table 3.2

Item	Predictability	Discriminability	Linked with Ethnicity
1. Felony arrest	No	No	Yes

2. Pending Charges	Yes	Yes	No
3. Outstanding Warrants	No	No	No
4. Criminal History	Yes	Yes	Yes
5. Prior FTA's	Yes	Yes	Yes
6. Prior Violent Crimes	No	Yes	Yes