Risk and needs assessment for juvenile delinquents

van der Put, C.E.

Publication date
2011

Citation for published version (APA):
van der Put, C. E. (2011). *Risk and needs assessment for juvenile delinquents*. [Thesis, fully internal, Universiteit van Amsterdam].

Chapter 2

Predictive validity of the *Washington State Juvenile Court Pre-Screen Assessment* in the Netherlands: The development of a new scoring system.

Van der Put, C.E., Stams, G.J.J.M., Deković, M., & Van der Laan, P.H.

Abstract

This study examined the predictive validity of the Washington State Juvenile Pre-Screen Assessment (WSJCPA) in the Netherlands. Previous research conducted in the United States showed the predictive validity of the WSJCPA to be modest, as is the case with the predictive validity of most other risk-assessment instruments for juveniles. Therefore, it was also examined whether the predictive validity of the WSJCPA can be improved by modifying the scoring procedure. The present study showed the predictive validity of the WSJCPA in the Netherlands to be modest too, with an 'area under the receiver-operating-characteristic curve' (AUC) of .63. Modifying the scoring procedure by means of CHAID analyses significantly improved the predictive validity to an AUC of .74. The modified scoring procedure is time-saving because only variables that uniquely contribute to the prediction of recidivism are included, which at the same time leads to a more accurate prediction of recidivism.

2.1 Introduction

Risk assessment provides insight into the level of the recidivism risk for juveniles who have come into contact with the law, and into the risk factors that determine the level of risk. Risk assessment is important for establishing the appropriate intensity of an intervention. If the recidivism risk is low, a low-intensity intervention – or even no intervention at all – is sufficient, whereas an intensive approach works better for high-risk juveniles (Andrews & Bonta, 2003; Andrews, Bonta, & Hoge, 1990). Matching the intensity of an intervention to the recidivism risk, also referred to as the 'Risk principle', is an important condition for the effectiveness of forensic interventions. If the intensity of an approach is not properly matched to the recidivism risk, its effect can be counterproductive. Both a too-light intervention for juveniles with a high recidivism risk and a too-intensive approach in the case of low-risk juveniles can increase the probability of recidivism, for example because juveniles come into contact with a target group of far more serious offenders, or because placement in a highly structured, restrictive programme has a disruptive effect on precisely those factors that underlie a low risk of recidivism (Andrews, 1995; Lösel, 1993; Lowenkamp & Latessa, 2005). Or, in the words of Andrews: 'If it ain't broke, don't fix it' (Andrews, 1995. p.54). Risk assessment is therefore an essential requirement if intervention is to be effective. The literature distinguishes between four generations in the development of risk-assessment instruments (Andrews & Bonta, 2003). The first generation of instruments consists of an unstructured judgement by a professional, based on knowledge, experience and intuition (clinical judgement). Second-generation instruments are structured actuarial risk-assessment instruments comprising mainly static risk factors, or immutable risk factors associated with recidivism, such as age at first offense and previous offenses. These instruments can predict the risk of reoffending, but do not provide information on how that risk can be reduced. Third-generation instruments are structured risk-assessment instruments incorporating dynamic as well as static risk factors. Dynamic risk factors can potentially be changed, such as problems at school and affiliation with delinquent friends. These instruments therefore provide insight into the possibilities for reducing the risk of recidivism. Finally, fourth-generation instruments comprise protective factors in addition to risk factors, and they also link the results to case management planning. Research has shown that structured risk-assessment instruments (second-, third-, and fourth-generation instruments) all perform considerably better than clinical assessment (e.g. Aegisdottir et al., 2006; Mossman, 1994; Trout & Bishop, 2002).

Despite the importance of risk assessment, there are currently no valid instruments available in the Netherlands for estimating the *general* risk of recidivism for a broad group of juveniles. The two principal validated instruments in the Netherlands that are

currently used in the juvenile justice system are the BARO (Basis Raadsonderzoek), which can be used to assess whether the juvenile has psychiatric problems (Doreleijers, Bijl, Veldt, & Van der Loosbroek, 1999) and the SAVRY (Structured Assessment of Violence Risk in Youth; Lodewijks, Doreleijers, De Ruiter, & De Wit-Grouls, 2006), which is used to assess the risk of violent recidivism among juveniles in a correctional institution. Since 2007, the Dutch probation service has used a Dutch translation of the Washington State Juvenile Court Pre-Screen Assessment (WSJCPA) for the purpose of risk assessment. The WSJCPA is a third-generation risk-assessment instrument that enables a relatively quick assessment of the risk of recidivism because it uses only the main predictors of recidivism (Barnoski, 2004a). The WSJCPA has been validated in the United States, but not yet in the Netherlands. The main aim of this study was therefore to examine the predictive validity of the WSJCPA in the Netherlands.

Predictive validity indicates how effectively a risk-assessment instrument predicts recidivism. The 'area under the receiver-operating-characteristic curve' (AUC) is regarded as the most important measure for this (Rice & Harris, 2005). The AUC indicates which percentage of correct classifications the instrument will yield overall (Hanley & McNeil, 1982). With a value of 0.50, the instrument is no better at predicting than a random assessment. A value of 1.00 indicates a perfect positive prediction, and a value of 0.00 indicates a perfect negative prediction. A meta-analysis of the predictive validity of risk-assessment instruments for juveniles shows that the AUC varied from 0.53 to 0.78, with an average AUC of 0.64 (Schwalbe, 2007). This average AUC value corresponds to a medium effect size, according to Rice and Harris (2005), but is considered moderate according to generally accepted criteria, because AUC values are considered 'acceptable' only from .70 upwards, and 'high' from .75 upwards (Dolan & Doyle, 2000; Hosmer & Lemeshow, 1989; Shapiro, 1999).

According to the research conducted in the United States, the predictive validity of the WSJCPA was modest with an AUC of .64. The predictive validity of the PACT (Positive Achievement Change Tool), an instrument that is largely based on the WSJCPA, was recently studied and also proved to be moderate, with an AUC of .59 (Baglivio, 2009). The present study therefore also examined whether the predictive validity of the WSJCPA can be increased by using a different scoring procedure. This was studied using Chi-squared Automatic Interaction Detector (CHAID) analyses. CHAID is a classification technique for identifying interaction effects between independent variables, and hence combinations of risk factors that lead to a high or low probability of recidivism. This technique is therefore highly appropriate for gaining insight into offender profiles with, respectively, a high and low probability of recidivism (Thomas & Leese, 2003).

There has been a great deal of discussion in the literature on the question whether the

same risk-assessment instruments can be used in a gender-neutral approach, i.e. for both sexes at the same time (e.g. Baglivio, 2009; Covington, 2003; Daly, 1992; Funk, 1999; Holtfreter & Morash, 2003; Reisig, Holtfreter, & Morash, 2006; Simourd & Andrews). The predictive validity of the WSJCPA, therefore, was examined separately for boys and for girls. To summarise, the following research questions were examined:

1   What is the predictive validity of the WSJCPA in the Netherlands?
2   To what extent can the predictive validity be increased by using a different scoring procedure?
3   To what extent is predictive validity determined by gender?

2.2 Method

2.2.1   Participants
The sample comprised 520 juveniles, 19% girls and 81% boys, who had been referred for rehabilitation. The average age of the juveniles at the time of screening was 16 years (SD = 18.0).

2.2.2   Measures
Washington State Juvenile Court Pre-Screen Assessment (WSJCPA). The WSJCPA is a risk-assessment instrument developed and validated in the United States (Barnoski, 2004a). The WSJCPA comprises the most important predictors of recidivism from two domains: the criminal history domain and the social history domain (Barnoski, 2004b). The items from the criminal history domain are: 'age at first referral', 'misdemeanour referrals', 'felony referrals', 'weapon referrals', 'against-person misdemeanour referrals', 'against-person felony referrals', 'detention dispositions', 'Juvenile Rehabilitation Administration (JRA) dispositions', 'escapes' and 'failure-to-appear warrants'. The items from the social history domain are: 'male gender', 'school problems', 'peer relationships', 'out-of-home placements', 'runaway history', 'criminal family member', 'parental rule enforcement', 'alcohol/drug problem', 'victim of abuse', 'victim of neglect', 'mental health problem'.

The criminal history score is the sum of items from the criminal history domain and ranges from 0 to 31, and the social history score is the sum of items from the social history domain and varies from 0 to 18. The two scores, criminal and social history, are brought together in a matrix, determining the youth's overall risk to re-offend (see Table 1).

Table 1 WSJCPA Scoring Matrix

	Social history score		
Criminal history score	0 to 5	6 to 9	10 to 18
0 to 5	Low	Low	Moderate
6 to 8	Low	Moderate	High
9 to 11	Low	Moderate	High
12 to 31	Moderate	High	High

Data recidivism. Recidivism is defined as the occurrence of one or multiple new contact(s) with the law within two years. To measure recidivism, data from the Client-Tracking System Youth Crime (CVS-JC) of the Ministry of Justice were used.

2.2.3 Procedure and analyses

For this study, the WSJCPA questionnaires, completed by the probation service in the period 2007 to the end of 2008, were analysed. The predictive validity of the WSJCPA was assessed using the 'area under the curve' (AUC). Chi-squared Automatic Interaction Detector (CHAID) analyses were used to study whether predictive validity can be increased by modifying the scoring procedure. To build the CHAID models, the total group of juveniles was randomly divided into two groups; 50% of the sample was used to build the models (construction sample) and 50% of the sample was used to validate the models (validation sample).

2.3 Results

2.3.1 Predictive validity of the WSJCPA

Table 2 shows the frequency distribution of the overall risk score and the average recidivism rate for each risk group.

Table 2 Overall risk score en recidivism rate per category

Overall risk score	N	%	Recidivism
Low	110	21%	32%
Moderate	215	41%	58%
High	195	38%	65%

In the present sample, 21% of the juveniles were low-risk, 41% were moderate risk and 38% were high-risk. The recidivism rate of the low-risk group was 32%, compared with 58% for the moderate-risk group and 65% for the high-risk group.

Table 3 shows the correlations with recidivism and the AUC values for the individual

items of the WSJCPA and for the criminal history score, the social history score and the overall risk score.

Table 3 Correlations with recidivism and AUC's of WSJCPA Items

|	r	AUC	95% C.I.
Criminal history domain			
Age at first referral	.25[*]	.641	(.594 - .689)
Misdemeanor referrals	.15[*]	.572	(.523 - .621)
Felony referrals	.29[*]	.659	(.612 - .706)
Weapon referrals	.05	.511	(.462 - .561)
Against-person misdemeanors	.12[*]	.530	(.481 - .580)
Against-person felony referrals	.14[*]	.570	(.520 - .620)
Detention dispositions	.15[*]	.569	(.519 - .620)
JRA dispositions	.06	.528	(.478 - .578)
Escapes	.02	.501	(.451 - .551)
Failure to appear warrants	.13[*]	.560	(.521 - .620)
Social history domain			
Male gender	.19[*]	.573	(.523 - .623)
School problems	.13[*]	.562	(.512 - .611)
Peer relationships	.14[*]	.570	(.521 - .620)
Out-of-home placements	.04	.510	(.460 - .560)
Runaway history	-.00	.500	(.450 - .550)
Criminal family member	.02	.504	(.454 - .554)
Parental rule enforcement	.08[*]	.541	(.491 - .591)
Alcohol/drug problem	.03	.506	(.456 – 556)
Victim of abuse	-.03	.510	(.461 - .560)
Victim of neglect	.04	.509	(.459 - .559)
Mental health problem	.02	.504	(.455 - .554)
Criminal history score	.32[*]	.694	(.648 - .740)
Social history score	.13[*]	.582	(.533 - .632)
Overall risk score	.24[*]	.625	(.576 - .674)

[*] $p < 0.01$

The strongest predictors of recidivism from the criminal history domain were: 'age at first referral', 'total number of felony referrals', 'total number of misdemeanour referrals' and 'detention dispositions'. Items from the social history domain that were strongly associated with recidivism were: 'male gender', 'school problems', and 'peer relationships'. In addition, the total criminal history score showed a considerably stronger correlation with recidivism than the total social history score. The AUC value for the overall risk score was .625. The AUC value for a number of items turned

out to be higher than the AUC value for the overall risk score, namely the items 'age at first referral', 'total number of felony referrals' and 'total criminal history score'.

2.3.2   Modified scoring procedure

CHAID analyses were used to examine whether predictive validity could be increased by modifying the scoring procedure. Separate CHAID analyses were performed for the items from the criminal history domain and for the items from the social history domain. Figure 1 shows the outcome of the analysis for the criminal history domain.

Figure 1 Results of CHAID analysis: Criminal history items
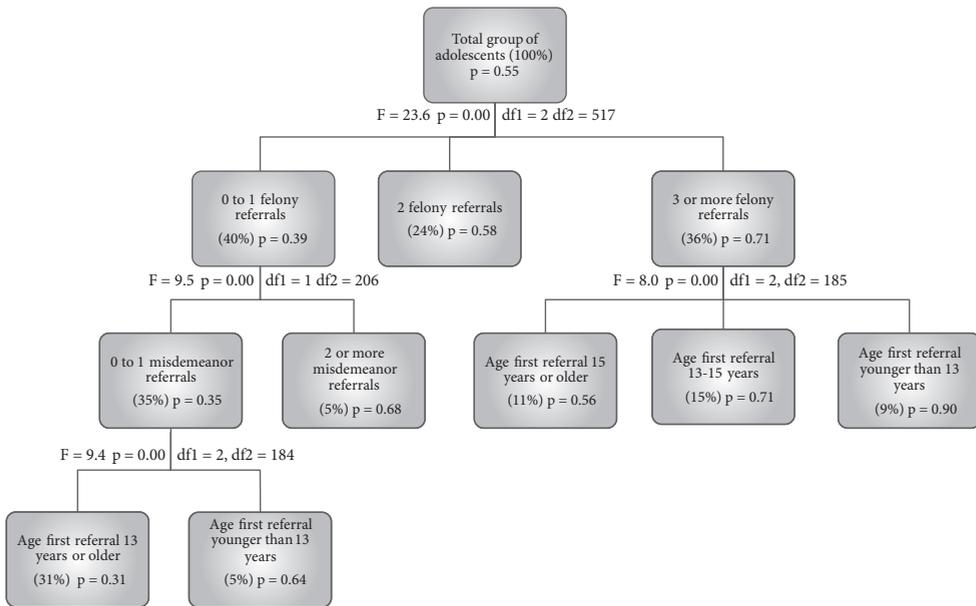
As input for the analysis all items from the criminal history domain were included as independent variables. Only three variables turned out to make a significant and unique contribution to predicting recidivism, namely the variables 'total number of felony referrals', 'total number of misdemeanour referrals' and 'age at first referral'. On the basis of the main predictor of recidivism ('total number of felony referrals') the total group of juveniles (100% with a recidivism probability of .55) was split into three groups: a group of juveniles with 0 or 1 felony referrals (40% of the juveniles with a recidivism probability of .39), a group with 2 felony referrals (24% of the juveniles with a recidivism probability of .58) and a group with 3 or more felony referrals (36% of the juveniles with a recidivism probability of .71). The groups were then di-

vided again according to the variables 'total number of misdemeanour referrals' and 'age at first referral'.

The total group of juveniles was thus divided into seven different risk groups, and the recidivism probability for the groups varied from .31 to .90. This CHAID classification has an AUC of .706 (.662 - .751), which is not significantly higher than the AUC of the original criminal history score (AUC = .694). Table 4 shows the risk groups and the related recidivism risks.

Table 4 Description of the criminal history risk profiles

	Criminal history risk profiles	%	*p*
1.	0-1 felony referrals, 0-1 misdemeanor referrals, age at first referral is 13 years or older	31%	.31
2.	3 or more felony referrals, age at first referral is 15 years or older	11%	.56
3.	2 felony referrals	24%	.58
4.	0-1 felony referrals, 0-1 misdemeanor referrals, age at first referral younger than 13 years	5%	.64
5.	0-1 felony referrals, 2 or more misdemeanor referrals	5%	.68
6.	3 or more felony referrals, age at first referral is 13 to 15 years	15%	.71
7.	3 or more felony referrals, age at first referral is younger than 13 years	9%	.90

The same analysis was performed for the items from the social history domain (see Figure 2). As input for the model all items from the social history domain were included as independent variables. The main predictor of recidivism was the variable 'gender'. The total group was divided as follows based on this variable: a group of girls (19% of the total group with a recidivism probability of .35) and a group of boys (82% of the total group with a recidivism probability of .59). The group of girls was then divided again into two subgroups based on the variable 'Parental rule enforcement' and the group of boys was divided again into four subgroups based on the variables 'peer relationships', 'problem behavior at school' and 'school results'. The total group of juveniles was thus divided into six different risk groups and the recidivism probability for the groups varied from .24 to .81. This CHAID classification has an AUC of .662 (.615 - .708), which is significantly higher than the AUC for the original social history score (AUC = .582). Table 5 shows the risk groups and the related recidivism risks.

Figure 2   Results of CHAID analysis: Social history items
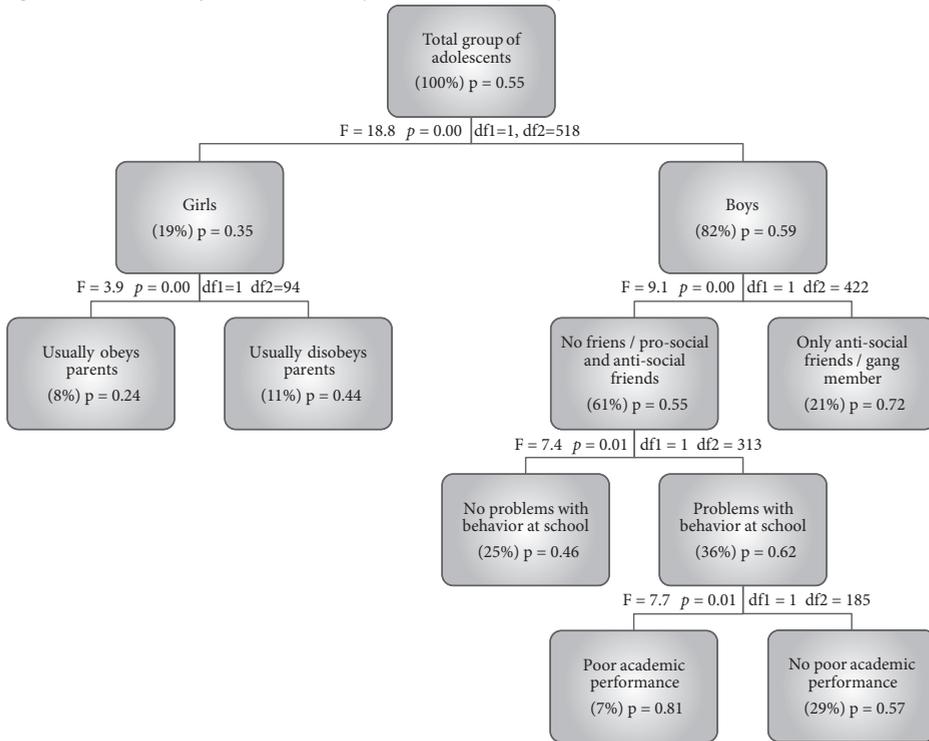
The total group of juveniles was divided into six different risk groups, and the recidivism probability of the groups varied from .17 to .90 (see Table 6). The AUC for the overall risk score was 0.738 (.695 - .780), which was significantly higher than the AUC of the original overall risk score (AUC=.625).

Table 6 Overall risk profiles

	Overall risk profiles	%	*p*
1.	Very low	11%	.17
2.	Low	20%	.39
3.	Moderate	27%	.50
4.	High	20%	.71
5.	High	12%	.76
6.	Very high	9%	.90

Table 5 Description of the social history risk profiles

	Social history risk profiles	%	p
1.	Girls, usually obeys parents	8%	.24
2.	Girls, usually disobeys parents	11%	.44
3.	Boys, no friends or both pro-social and anti-social friends, no problems with behavior at school	25%	.46
4.	Boys, no friends or both pro-social and anti-social friends, problems with behavior at school, no poor academic performance	29%	.57
5.	Boys, only anti-social friends or gang member	21%	.72
6.	Boys, no friends or both pro-social and anti-social friends, problems with behavior at school, poor academic performance	7%	.81

A further CHAID analysis was then performed in which the CHAID score for the criminal history domain (see Table 4) and the CHAID score for the social history domain (see Table 5) were combined to give a single overall risk score. This analysis used two variables as independent variables (the CHAID-score for the criminal history domain and the CHAID-score for the social history domain) in order to determine how these two scores should be combined to achieve the best prediction of recidivism (see Figure 3).

Figure 3   Results of CHAID analysis: Combining Criminal Risk score and Social Risk score
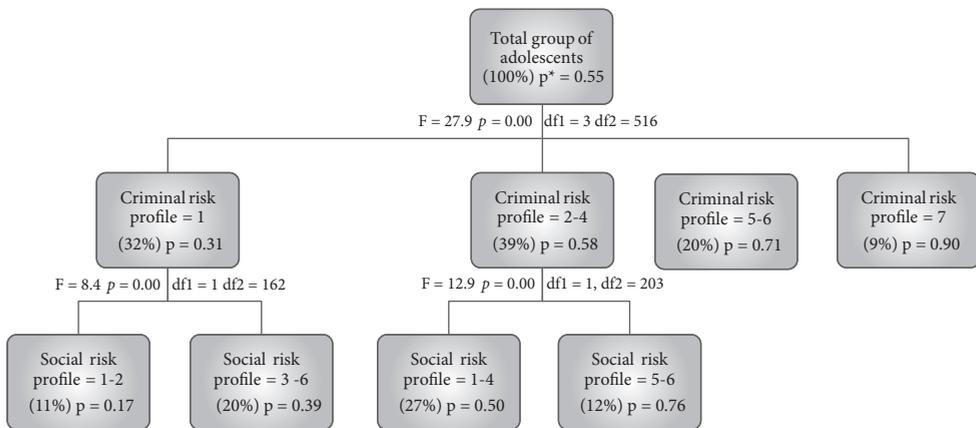
2.3.3   Gender differences in predictive validity
Table 7 shows the AUC values of the WSJCPA and the modified scoring procedure, separately for boys and girls. It is shown that predictive validity does not differ significantly between boys and girls. This applies to the WSJCPA as well as the modified scoring procedure.

Table 7 AUC values seperately for boys and girls

	WSJCPA overall risk score	WSJCPA adapted score
Total group	.625 (.576 - .674)	.738 (.695- .780)
Boys	.605 (.549 - .660)	.715 (.666 - .764)
Girls	.636 (.520 - .753)	.731 (.623 - .840)

2.4 Discussion

This study examined the predictive validity of the Washington State Juvenile Pre-Screen Assessment (WSJCPA) in the Netherlands. The predictive validity of the WSJCPA was measured in previous studies conducted in the United States and proved to be moderate (Baglivio, 2009; Barnoski, 2004b). A meta-analysis showed that the predictive validity of the vast majority of other risk-assessment instruments for juveniles was also moderate (Schwalbe, 2007). The study therefore also examined to what extent the predictive validity of the WSJCPA can be increased by using a modified scoring procedure.

First, we examined the relation between the individual items and recidivism. The strongest predictors of recidivism from the criminal history domain were: 'age at first referral', 'total number of felony referrals', 'total number of misdemeanour referrals' and 'detention dispositions'. The strongest predictors of recidivism from the social history domain were: 'male gender', 'school problems', 'peer relationships' and 'parental rule enforcement'. These findings are in accordance with the validation study from the Unites States, which also showed that these items are the strongest predictors of recidivism (Barnoski, 2004b).

However, in the Barnoski study a number of items from the social history domain (i.e. 'runaway history', 'parental rule enforcement' and 'alcohol/drug problem') were found to have a stronger correlation with recidivism than in the Dutch study. Consequently, the correlation of the total social history score in the Barnoski study (r = .22) is significantly higher than in Dutch research (r = .13) (Fisher's z = 2.08, p = .04). The correlation of the total criminal history is, in contrast, significantly higher in the Dutch study (r = .31) than in the Barnoski study (r = .20) (Fisher's z = 2.89, p = .00). The predictive validity of the WSJCPA in the Netherlands proved to be moderate, with an

AUC of .63, and, therefore, does not differ significantly from the predictive validity in the other validation studies (Baglivio, 2009; Barnoski, 2004b).

The second aim of this study was to examine whether the predictive validity of the WSJCPA can be increased by modifying the scoring procedure. CHAID analyses were used for this purpose. Before performing the analyses, the total sample was divided into a construction sample and a validation sample. The AUC value of the modified scoring procedure was .74, which was significantly higher than the AUC of the original overall risk score. In addition, this AUC value meets the generally accepted criterion for an acceptable prediction (Dolan & Doyle, 2000; Hosmer & Lemeshow, 1989; Shapiro, 1999).

The third aim was to examine whether predictive validity varied by gender. The predictive validity of both the WSJCPA and the modified scoring procedure did not differ significantly between boys and girls, so the prediction is equally in boys and girls. The results of this study have a number of important implications for policy and clinical practice. In the first place, the results show that the AUC of the WSJCPA can be significantly improved by modifying the scoring procedure, which means a more accurate prediction of recidivism. Second, considerably fewer variables are needed to calculate the overall risk score. Due to the strong correlations between the individual items, only 7 of the 21 items make a unique contribution to predicting recidivism, and it is therefore sufficient to measure these 7 items only. A final advantage is that aspects that are difficult to assess (e.g. sexual abuse, mental health problems, detention history of parents) can be omitted without reducing predictive capacity, because these items do not make a unique contribution compared to the other variables in terms of predicting recidivism. To summarise, the modified scoring procedure leads to an increase in predictive validity as well as a considerable time saving when completing the WSJCPA, making the instrument readily applicable for a broad group of juveniles. Given the moderate predictive validity of the majority of risk-assessment instruments for youth, a recommendation is to examine whether the predictive validity of other instruments can be improved as well by using an alternative scoring procedure. In this way it might be possible to make a considerable improvement in the predictive capacity of risk-assessment instruments for juveniles and hence in the effective matching of the intensity of treatments to the level of risk of the juveniles.