

**Research to Validate the New Zealand  
Police Youth Offending Risk  
Screening Tool (YORST) Phase II:  
Predictive Ability Analysis**

**A report prepared by**

**Dr Elaine Mossman  
of the Crime and Justice Research Centre**

**November 2011**



Published in 2012 by the  
New Zealand Police  
PO Box 3017  
Wellington  
New Zealand

© Crown Copyright

ISBN 978-0-477-10371-8

# Acknowledgements

The report was written by Elaine Mossman, with input from Venezia Kingi. Dr Sue Triggs provided peer review on the statistical analysis.

Melanie Atkinson, Bea Makwana and Alison Chetwin of New Zealand Police have project managed this research and provided valuable support and feedback. They have reviewed the draft report along with Basia Arnold and Su-Wuen Ong from the Ministry of Justice, Anna Duncan from New Zealand Police, Fleur McLaren from the Ministry of Social Development and Peter MacIntosh and Onny Gajadhar from Child Youth and Family. Stuart Morley of Child Youth Family provided great assistance in extracting time at large information for the YORST sample. All comments and suggestions were very much appreciated.

A final acknowledgement needs to be given to Associate Professor Anthony Thompson of Charles Sturt University for his independent review and valuable feedback in relation to this report.



# Contents

<b>Key findings.....</b>	<b>i</b>
<b>1 Introduction.....</b>	<b>1</b>
1.1 Background.....	1
1.1.1 Existing YORST.....	2
1.2 Methods to test the predictive ability of the YORST.....	4
1.2.1 Sub-group.....	7
1.2.2 Outcome variables.....	7
1.2.3 Predictor variables.....	8
1.2.4 Analysis.....	9
<b>2 Accuracy of existing YORST.....</b>	<b>11</b>
2.1 Sample characteristics.....	11
2.2 Percentage of cases of high risk offenders correctly predicted.....	12
2.3 Accuracy of the prediction across the range of total risk scores.....	12
2.4 Mean YORST scores of those re-apprehended compared to those not.....	14
2.5 Point bi-serial correlation.....	14
2.6 Receiver Operating Characteristics (ROC) analysis.....	14
2.7 Summary.....	17
<b>3 Item response analysis.....</b>	<b>19</b>
3.1 Item analysis – frequency of responses.....	19
3.2 Inter-correlations between items.....	19
3.3 Logistic Regression.....	22
3.4 Modifications to existing YORST.....	22
3.4.1 Removing redundant items.....	23
3.4.2 Weighting of items based on logistic regression model.....	23
3.4.3 Re-coding of items.....	23
3.5 Summary.....	29
<b>4 Logistic model building.....</b>	<b>31</b>
4.1 Identifying alternative predictor variables.....	31
4.2 Additional LOST predictors.....	32
4.2 Logistic model building.....	33
4.2.1 Comparison of construction and validation sample.....	35
4.2.2 Utility of the new model.....	36
4.2.3 Summary.....	36
<b>5 Conclusion.....</b>	<b>39</b>
<b>References.....</b>	<b>41</b>

<b>Appendices.....</b>	<b>43</b>
Appendix 1: The YORST .....	43
Appendix 2: Re-coding of existing YORST.....	45

## Tables

Table 2.1	Characteristics of re-apprehension offences.....	11
Table 2.2	Percentage of cases correctly predicted .....	12
Table 2.3	Percentage of cases re-apprehended across low, medium and high YORST total risk scores .....	13
Table 2.4	Difference in mean total YORST risk scores of those re-apprehended and those not re-apprehended .....	14
Table 2.5	ROC analysis for existing YORST .....	16
Table 3.1	Logistic Regression analysis of existing YORST items .....	22
Table 3.2	Logistic Regression analysis of existing YORST items .....	29
Table 4.1	Logistic Regression analysis of alternative variables and existing YORST items.....	34
Table 4.2	ROC analysis for existing YORST .....	35
Table 4.3	ROC analysis for existing YORST (validation sample) .....	36

## Figures

Figure 2.1	Existing YORST scores against percentage re-apprehended .....	13
Figure 2.2	ROC curve for the YORST total risk score (n=1965); AUC=.695 .....	15
Figure 3.1	Correlations with YORST items 1 to 14 with each other and re- apprehension outcome variables .....	21
Figure 3.1	Individual item responses against proportion of group re-apprehended (mean of apprehension outcome) .....	25
Figure 4.1	ROC curve for revised model (construction sample n=982); AUC=.738 .....	35

# Key findings

The Lifetime Offender Seriousness Tool (LOST) provided a summary of the criminal histories of 1965 youth who completed a Youth Offending Screening Tool (YORST) screen between July and October 2009. In the 12 months following the YORST screen 1303 (66.3%) had been re-apprehended for at least one new offence. The accuracy of the existing YORST to predict this recidivism behaviour was tested through a series of analyses.

## Accuracy of existing YORST

Total YORST risk scores can range from 0 to 100. These scores have been categorised into low, medium or high risk based on the following cut-off points. The proportion of youth who were re-apprehended increased according to their YORST risk classification:

- Low risk (total YORST risk scores of 0 to 29), 32.7% re-apprehended
- Medium risk (total YORST risk scores of 30 to 69), 70.6% re-apprehended
- High risk = total YORST risk scores of 70 to 100, 83.5% re-apprehended.

The current cut-off points were better able to differentiate between low and medium risk offenders, than medium and high risk, suggesting a review of these cut-off points may be warranted in the future.

Overall the accuracy of the existing YORST compared favourably with other well established and typically more comprehensive risk assessment tools (e.g. UK's ASSET and YLC/CMI). It should be noted, however, that predicting recidivism in young offenders is not an easy task and none of the tools do exceptionally well at this.

- YORST total risk scores correctly predicted re-apprehension status for 68.2% of the sample. The ASSET predicted 67% over the same period.
- total YORST risk scores were significantly correlated with re-apprehension status ( $r=.34$ ). This is comparable to other well established tools which have correlation coefficients ranging from .28 through to .43.
- the overall accuracy of the YORST was found to be moderate to high (ROC analysis on the YORST produced an AUC of .695 for the entire sample and .703 for the Special Group of more reliable raters). Again the accuracy of the YORST compared well to other well established tools (for reconviction over 12 months, AUCs for the YLS/CMI ranged from .641 to .67, and .712 for the ASSET).

YORST total risk scores were significantly higher for those who were re-apprehended for any offence, for violent offences and those of medium seriousness or higher. The tool performed equally well for males and females, but was more accurate for New Zealand European compared to Māori and also more accurate for older offenders. As noted above, the existing YORST appeared better able to differentiate between low and medium risk offenders than it did between medium and high risk offenders.

## Item response analysis

Item response analysis was carried out on 14 individual items that make up the total YORST risk score to examine their characteristics and how well they perform in predicting recidivism.

All 14 items had statistically significant correlations with re-apprehension. The strongest predictors in order of their strength were item 1 – time since last came to police notice for offending, item 3 – highest level of previous intervention, item 7- influential peers known to police and item 2 – time since last came to notice for an incident.

Despite each of the 14 items on an individual basis being useful in differentiating those who had been re-apprehended, a high level of inter-correlation amongst items suggested some redundancy in items. This was confirmed with a logistical regression analysis which found, once each item's unique contribution was accounted for, only four items significantly contributing to the prediction of re-apprehension (items 1, 2, 3 and 7), with item 1 providing the greatest contribution. This means that if all that was required from the YORST was a prediction of recidivism, it could consist of just four items. However, this is not the case with the YORST which also aims to capture information on criminogenic need. Hence, although not required for predicting risk of recidivism, retention of other items enables the collection of useful case management information.

The item-response analysis provided important information on the potential role of different items and how they might be modified to enhance their primary objective, be it predicting recidivism or identifying areas of criminogenic need.

A number of modifications to the existing YORST were trialled to see if it was possible to improve its predictive ability. All modifications showed some improvement compared with the accuracy of the existing YORST. Modifications included:

- the removal of redundant items (those items offering no additional predictive value once the strongest predictors had been accounted for) which achieved an AUC of .723, and .787 for the Special Group of raters
- applying weightings to items based on the logistic regression model which achieved an AUC = .729, and .788 for the Special Group of raters
- re-coding of items based on responses to each item and their relationship with re-apprehension. ROC of the re-coded items achieved an AUC of .735, and .774 for the Special Group of raters.

This re-coding was assisted by examining plots of the response options for each item against the proportion of the group that had been re-apprehended for each option. This enabled for the first time an objective assessment of whether response options for items were useful in differentiating levels of risk. These graphs, together with the identification of items that are significant predictors of re-apprehension, provide valuable data for reviewing and revising the existing YORST.



### **Logistic model building**

Logistic model building was used to explore if an alternative combination, formation or transformation of predictor variables could better predict recidivism. An alternative model was developed using a construction sample and then tested on a validation sample.

Additional variables found to be strongly correlated with apprehension included:

- number of prior offences (count of offences recode)
- number of prior violent offences
- log of the sum of all prior seriousness scores
- log of the single most prior seriousness score
- length of prior offending career

These were entered into a logistic regression analysis along with demographic variables and YORST items already found to be predictive. Just five variables were found to be significant predictors (gender, the number of prior offences, and YORST item 1 (re-coded) and items 2 and 7).

The new model based on these variables showed a slight improvement on the initial model that only used the original YORST items. The new model had comparable ROC results to other modified versions of the YORST (AUC=.748; and .826 for the Special Group) and appeared more accurate in its ability to predict re-apprehension status across different sub-groups. The model remained a good predictor of re-apprehension when re-tested on a new validation sample.

International best practice recommends the use of both static and dynamic factors in assessing risk of recidivism. This is so that assessments can inform case management as well identify level of risk. A limitation of the new model is the over-reliance on static factors. While providing a slightly more accurate picture of risk, the use of these items alone does little to inform case management and the over-reliance on static factors means the tool will not be sensitive to any changes in a young person's level of risk over time.

### **Conclusion**

Analysis presented in this report has revealed that modifications to the existing YORST through the re-coding of items, applying weights to predictive items or including additional predictor variables were able to improve slightly its predictive ability. However, of particular value has been the identification of items which are most important in predicting recidivism. Identification of these items has important implications for the practical application of the YORST.

The YORST has a number of objectives:

- to identify high risk offenders

- to identify areas of criminogenic need that can be targeted through interventions such as alternative action plans or youth development programmes is also of high importance
- to collect standardised information
- to enable information sharing among agencies
- to be able to assess the impact of interventions by measuring change in risk level.

Having established which items are most useful in predicting recidivism, items can now be reviewed and their role and potential utility in relation to the above objectives reconsidered. For example, it could be that items more useful for identification of criminogenic needs can be redesigned to better capture this information. It would also be good to review how dynamic factors that are known to predict risk of recidivism can be better measured so their inclusion in the YORST may contribute more significantly to the prediction of recidivism.

The current screening process should also be reviewed in the light of these findings. It will be useful to consider who should be given a YORST screen and whether they should be screened using all 14 items, or whether there may be advantages in screening a greater proportion of child and young offenders with a smaller number of items. Those identified as high risk could then receive a fuller risk and need screening and/or assessment.

Predicting recidivism in young offenders is not an easy task and none of the currently available risk assessment tools do exceptionally well at this. The predictive ability of the YORST was as good, and in some cases better than other well established and more comprehensive risk assessment tools. This is a good achievement for New Zealand Police, particularly considering the comparative brevity of the YORST screen with just 14 items.

# 1 Introduction

The Crime and Justice Research Centre was commissioned to carry out phases II and III of a three year programme of research that aims to assess the validity, reliability and predictive capability of the New Zealand Police Youth Offending Risk Screening Tool (YORST). A copy of the YORST instrument appears in Appendix 1.

Phase I was completed in August 2010. This consisted of a literature review and analysis of the quality of existing data (e.g. accuracy, rate of completion and adherence to completion criteria). The objectives of Phase II were to:

- consider whether any of the tasks partially completed in Phase I can be completed
- identify ways to improve the quality of data collected by YORST by thorough field observations and key informant interviews
- understand the primary purpose of YORST and its other objectives and how YORST information is currently being used
- assess the ability of existing YORST items and the total YORST score to predict recidivism.

This report presents key findings from the last sub-component of Phase II listed above (testing the predictive ability of the existing YORST). This includes the testing of the accuracy of the existing total YORST score to predict recidivism, an analysis of the relative contribution and utility of individual YORST items, and an exploration of whether an alternative combination, formation or transformation of items or additional variables are better able to predict recidivism.

Once Phase II has been completed, Phase III will test the concurrent validity and inter-rater reliability and predictive validity of the finalised YORST.

## 1.1 Background

A range of youth justice decisions are made on the basis of judgements around a young person's likelihood of re-offending together with their areas of criminogenic need. When offending by a young person first comes to the attention of the police, judgements on risk and needs inform decisions around what action to take, whether a custodial remand is appropriate, and what, if any, intervention or referrals should be made. A primary concern in making these decisions is public safety, how can the public be best protected from immediate risk of harm from young offenders, and also through reducing recidivism by providing appropriate interventions to youth. However, if these judgements on risk and need are made through informal and unsystematic procedures, this can lead to inconsistency and bias in the decision process. As a result, standardised risk and need assessment and/or screening of offenders is now well recognised as best practice (Andrews et al., 2006; The Werry Centre, 2009; Vincent et al., 2009).

The value standardised risk assessment of young offenders in New Zealand has been highlighted in several government reports (Ministry of Justice and Ministry of Social Development, 2002a; Ministry of Justice, 2002) and was a recommendation of the 2002 Youth Offending Strategy (Ministry of Justice and Ministry of Social Development, 2002b). As a result, in 2003 New Zealand Police began work to develop a risk screening tool.

- In 2003 a tool known as the Adolescent Risk Needs Inventory (ARNI) was developed. This was an extensive assessment instrument of approx 30 pages.
- The 30 page approach was rejected by police as being impractical for Youth Aid Officers to use and at the end of 2003 a group of experienced Youth Aid Officers were assembled to develop a shorter version. This shorter version was known as the ARNI by some and the Risk Screening Tool by others. The tool was circulated and staff were informed that as the initial period of use would be regarded as a 'trial period' there was some flexibility around completion.
- In July 2006 a review of the Risk Screening Tool was carried out by Dr Melanie Atkinson to seek feedback about its current use, views on its usefulness, and suggestions for improvements. Responses from Districts indicated that there was no standard model of practice. There appeared to be no application in some areas and some quite sophisticated attempts in others. Some Districts made adaptations to the tool itself. Feedback from those areas that were using the tool was positive and Youth Aid Sergeants identified a number of benefits. A review of literature supported the content of the Risk Screening Tool (i.e. risk factors included).
- A focus group of Youth Aid and Youth Development staff met to assist in the refinement/improvement of the Risk Screening Tool. They looked at four other risk and needs assessment tools, comparing them to the Risk Screening Tool and also discussed the suggestions for change that had come out of the review from those Youth Aid and Youth Development staff who had been using the tool.
- The tool was revised and renamed the YORST. Piloting of his revised tool took place in the Waikato and Bay of Plenty Districts between June and December 2007.

### **1.1.1 Existing YORST**

The revised YORST was rolled out nationally on 1st of July 2009. A series of workshops were held around the country to introduce the YORST and provide training on how to use it. These consisted of a one day joint training session attended by all Police Youth Aid and Youth Development staff and CYF Youth Justice staff. The training provided information about risk factors for youth offending, how and when to apply the tool, and the processes required for police to share the tool with CYF. Policy guidelines were also developed and distributed that outlined the completion instructions for the YORST (New Zealand Police, 2009a).

There are a number of potential uses of the YORST. Recent New Zealand Police documents have outlined five intended outcomes which are to (New Zealand Police, 2009b):

- screen for levels of risk in young people

- screen for areas of risk in young people
- enable Child Youth and Family to ensure that further assessment and interventions are targeted for young people who offend
- enable the investigation of the impact of police interventions on risk of re-offending
- indicate the likelihood (or risk) of the young person re-offending so that appropriate decisions can be made about the intervention pathways for individuals

### **Content of the YORST**

The YORST is intended to be a brief risk screening tool rather than a comprehensive assessment (a copy of the YORST appears in Appendix 1). It contains 14 items that enquire about the prevalence of factors which increase the likelihood of a child or young person re-offending.

The items are focused on the following areas of the child or young person's life:

- offending factors
- peer group factors
- education/employment factors
- care and protection history
- alcohol and drug use
- family factors.

Six items are automatically generated from the New Zealand Police National Intelligence Application (NIA); the remainder are based on ratings given by the individual completing the YORST. Some items require secondary sources to be accessed (e.g. information gained from speaking with the young person, their family or whānau, their teacher or trainer, and CYF).

The YORST identifies both dynamic and static factors and so can be used in both a diagnostic and prognostic way.<sup>1</sup> The total score ranges from 0-100 and is intended to be predictive of the likelihood of future re-offending. These total scores are translated into the following categories of risk:

- (i) Low risk = 0-29%
- (ii) Medium risk = 30-69%
- (iii) High risk = 70-100%

A separate score is also produced for the nine dynamic factors. This score is more sensitive to change. With repeat applications of the YORST, this dynamic score can provide an indication of whether the child or young person's risk of re-offending is increasing or decreasing.

---

<sup>1</sup> Dynamic factors are those which can be potentially changed, for example anti-social attitudes, negative peer associations, truancy or abuse of alcohol or drugs.

## Minimum completion criteria

The YORST policy guidelines outlined the minimum criteria for completing a YORST (New Zealand Police, 2009b). These are presented in Figure 1.1.

The guidelines stipulated that Sergeants or Senior Sergeants in charge of Youth Aid or Youth Services were able to set other requirements over and above the minimum criteria - depending on area priorities and issues being focused on.

**Figure 1.1: Minimum criteria for completing a YORST**

A YORST is to be completed on every child and young offender that meet the following minimum criteria:

### Child offenders

- All children (aged 10-13 years) who have come to police attention for a *second* offence and/or incident.
- All children that are having a Youth Justice Family Group Conference (FGC) [s14(1)(e), s247(a)].

### Youth offenders

- Every young person referred by the police for a Youth Justice FGC (s247(b)).
- Every young person who is arrested and brought before the Youth Court and an FGC is required pursuant to;
  - Section 247(c) - the charge is denied and the young person has been remanded in custody
  - Section 247(d) - the Youth Justice Co-ordinator is directed to convene an FGC
  - Section 247(e) - the charge against the young person is proved and a FGC has not had the opportunity to consider ways in which the Court might deal with the young person for the offence that forms the basis of the charge.

### Police Youth Development Referrals

- Children or young people being referred to a Police Youth Development Programme (YDP).

## 1.2 Methods to test the predictive ability of the YORST

The aim of this report is to present findings on the analysis that:

- tested the accuracy of the existing total YORST score to predict recidivism
- investigated the relative contribution and utility of individual YORST items
- explored whether an alternative combination, formation or transformation of items or additional variables are better able to predict recidivism.

The predictive ability of the YORST was tested in a number of ways. Much of the methodology selected was drawn from similar work carried out in England to assess

the ASSET (Baker et al., 2003) and that in Australia to test the YLS/CMI-AA (Thompson and Pope, 2005; Upperton and Thompson, 2007).

In order to assess the predictive ability of the existing YORST it was necessary to test scores from a sample of YORSTs completed on young offenders against their actual recidivist behaviour in the 12 months following the YORST screen. This required the following four steps to be carried out:

1. identification of a sample of young offenders with completed YORST screens
2. retrieval of criminal history data for these young offenders including a 12 month period after the date the YORST was completed
3. identification of any time in custody for each young person in the 12 month follow-up in order to ensure that recidivist behaviour was based on equal time-at-large follow-up periods
4. application of the Lifetime Offender Seriousness Tool (LOST) to provide summarised criminal history data for each young person.<sup>2</sup> This LOST data enabled the testing of YORST predicted risk scores against actual recidivism behaviour. The LOST data also provided alternative predictor variables for consideration.

A description of each of these four steps appears below:

### **1. Identification of a sample of completed YORST screens**

The initial sample of YORSTs consisted of n=2245 screens carried out between 1 July 2009 and 31 October 2009. This sample was reduced to n=1999 after the following cases were removed from the sample:

- YORSTs with no total YORST score (n=21)
- YORSTs recorded as still being 'in progress' (n=163)
- where a young offender had more than one YORST completed in this period, just the first YORST for each young offender was retained (n=62)<sup>3</sup>

### **2. Retrieval of criminal history data**

The unique person ID associated with each of the n=1999 YORSTs was supplied to New Zealand Police in order for them to retrieve the criminal histories for each young person. Criminal histories were extracted for n=1981. IDs could not be identified for n=18; further investigation revealed these IDs were associated with youth with an alias, which had subsequently been matched and their records merged with their other ID. These youth were not included in the analysis as the automatically generated items on the original YORST would have been based on a partial criminal history.

---

2 The development of the LOST has been jointly funded by the Ministry of Social Development and New Zealand Police. The tool was initially created by Christine Jamieson when working for New Zealand Police, and further developed whilst working at the Ministry of Social Development

3 Subsequent offending that resulted in another YORST being completed became data to compare the initial YORST score against.

### 3. Identification of time in custody

Child Youth and Family (CYF) have the most useful records related to a young person being held in custody. It is possible to extract from their database custody related court orders and the length of time these court orders apply. The Children, Young Persons, and Their Families Act 1989 (the Act) provides for three situations and associated orders where a young person can be held in custody:

- detained in police custody following an arrest (section 236)
- remanded in custody pending a court hearing (section 238 1D or 1E)<sup>4</sup>
- sentenced to a supervision with residence order (section 311)

The unique person identifier used by police is different to that used by CYF, hence, names and dates of birth of each young offender had to be supplied to CYF which could then be matched to persons in their records. An initial run of the n=1981 found a positive match for n=1366. Those unmatched were either minor young offenders who had not had involvement with CYF or whose match had not been found because details held in the two systems differed (e.g. differences in names or date of births recorded).

Police records can identify if a young person has received a supervision with residence order, but do not record the duration. A police resolution report was run on the n=633 unmatched young offenders and a further 11 were identified as receiving supervision with residence in the 12 month follow-up period. These youth were manually searched for and a match subsequently found on the CYF database.

Of those youth where a match had been found (n=1377), CYF records indicated n=157 had received a custody order in the 12 month follow-up period, ranging from 1 day through to 394 days. These youth had their follow-up period extended by the number of days they were held in custody. This meant each young person had recidivist behaviour compared for a full 12 month period where they were at large. There were 16 young offenders whose extended 12 month follow-up period now ran beyond that of the criminal histories extracted and so were removed from the analysis. This left a final sample of n=1965 for analysis.<sup>5</sup>

The time in custody could be considered as an intervention that may have impacted on a young person's recidivism. This research did not attempt to quantify the impact on recidivism of this or any other interventions.

---

4 Under section 238 1E a young person can be remanded in police custody for up to 10 days. Under section 238.1D a young person is usually remanded to a CYF secure residence, however, there are a few occasions where they may be remanded to a CYF community-based family home. Community-based family homes are very restrictive (e.g. curfews enforced) but not secure residences. It was difficult to determine which type of residence had been used and so for the purposes of the research remand to either was considered to be in custody.

5 Analyses were run with and without time at large information included. The accuracy of the YORST improved with time at large information included but findings did not vary in regards to variables identified as significant predictors.



## 4. Application of the LOST

The LOST is a newly developed tool which simplifies the process of summarising a young person's criminal history data. This research provided the first opportunity to trial the tool to extract recidivist behaviour for a sample of young offenders over a specified period. In addition to actual recidivism, the tool was also able to provide a range of other summarised criminal history variables which could be examined as possible alternative predictor variables.

For each young person the date the YORST was completed was entered into the LOST and the tool was set to summarise criminal history for the following 12 months (where relevant the number of days a young person had been held in custody was added onto the 12 month period).<sup>6</sup> The LOST tool was then applied to the criminal histories of the final sample (n=1965).

### 1.2.1 Sub-group

A user ability assessment carried out in November and December 2010 revealed concerns that there were some raters who were not completing the YORST as intended (i.e. making insufficient enquires to ensure all items were completed accurately). In order that the existing YORST was tested against accurate scores a sub-set of more reliable raters were identified (n=30) who as a group had completed 330 YORSTs. These were individual Youth Aid Officers identified by their supervisors as using the tool as intended (i.e. carrying out thorough enquires to ensure valid rating of all YORST items).

### 1.2.2 Outcome variables

The primary outcome measure of recidivism was whether the young person had been re-apprehended for a new offence in the 12 month follow-up period. The diversionary emphasis of the New Zealand Youth Justice system means a relatively small proportion of child offenders (2%) or young offenders (29%) who come into contact with New Zealand Police are prosecuted. Therefore, while it is common in young offender research to use reconviction (following prosecution) data to assess recidivism, re-apprehension data was considered more appropriate in this case as it provided recidivism information that is relevant to a greater proportion of youth who come to police notice. The other advantage of using re-apprehension data was that it was in real time and was not subject to the procedural time delays of prosecution data where a Youth Court decision can take many months and sometimes years to be processed. The main limitation of this measure is that apprehensions represent the number of alleged offences and apprehensions. They do not tell us the number of proven offences<sup>7</sup>

6 The LOST has been designed to assess recidivism behaviour before and after an intervention. To get recidivism behaviour for a fixed 12 month period it was necessary to use the summarised data for an 'intervention' period, with the 'intervention' being the 12 months period from the date the YORST screen was carried out.

7 An apprehension can result in a number of outcomes and this varies for children and young person (e.g. caution/warnings, alternative actions, family group conferences). In 2008 less than a third of apprehensions of 14 to 16 year olds resulted in a prosecution, and of these around 20% the case against the young person was found to be not proven (Duncan, 2010).

**Primary outcome measure:** whether a young person had been re-apprehended for an offence in the 12 month follow-up period after the YORST was completed.

Offences were counted according to their occurrence date, to ensure only apprehensions for offences occurring after the YORST had been completed were included. Cases where there were new apprehensions for offences were coded as 1 and cases with no new apprehensions were coded as 0. A total of n=1303 (66.3%) of youth out of the full sample of n=1965 had been re-apprehended for one or more offences.

In assessing the predictive ability of a risk of re-offending tool, it is useful to assess its performance against a range of recidivism measures. Other outcome measures included:

- whether a young person had been re-apprehended for any violent offences
- if they had been re-apprehended for an offence of medium seriousness or higher.<sup>8</sup>

### 1.2.3 Predictor variables

The primary predictor variable of interest was the YORST total risk screening score. These scores range from 0 to 100 and are calculated by effectively summing the individual scores of all completed YORST items. The separate YORST items 1-14 were also assessed for individual predictive ability. This required items to be re-coded to a numerical score from 0 to 5 for analysis.

Other alternative predictor variables available for consideration included the following:

- demographic variables (age, gender and ethnicity)
- LOST variables relating to criminal history prior to the YORST being completed including:
  - a) age of first offence
  - b) length of offending career prior to YORST
  - c) number of offences prior
  - d) average frequency of offending prior
  - e) number of violent offences prior
  - f) single most serious offence prior
  - g) average seriousness of offending prior
  - h) accumulated seriousness of offending prior (average seriousness multiplied by length of offending career prior to YORST).

---

<sup>8</sup> The LOST tool used the Ministry of Justice (MOJ) Seriousness Scale as a measure of the seriousness of offending. This scale gives imprisonable offences a score according to how serious judges have deemed each offence based on the average number of days of imprisonment imposed over a specific time period. For this research, offences of medium seriousness were defined as offences with an MOJ seriousness score of 67.84 or higher. This is the simplified cut-off score for medium offences currently used in the YORST tool.

- LOST scores based on complete lifetime criminal histories including:
  - a) total number of offences ever
  - b) frequency counts of different categories of offences, including all offences classified as 'non-compliant'
  - c) frequency counts of truancy incidents
  - d) most serious offence ever
  - e) length of lifetime offending career
  - f) average frequency and seriousness of offences during offending career
  - g) total lifetime LOST (a sum of seriousness scores for all offences)
  - h) frequency counts of all victimisation occurrences (family violence and non-family violence related)
  - i) total estimated police cost associated with offending

#### 1.2.4 Analysis

Once data had been summarised by the LOST tool, relevant data for each young person was matched up with their original YORST screening data and imported into Statistical Package for the Social Sciences (SPSS) version 18 for analysis. Data was carefully screened for any errors, inconsistencies, omissions and to examine the frequency distribution of all variables. Once cleaned, variables were examined to see if re-coding or transformation was required to improve their distribution<sup>9</sup> Variables entered into logistic regression were assessed for linearity and multicollinearity.

The main stages of analysis are outlined below. Chapters of the report are organised around each stage.

#### **Chapter 2 - Assessment of accuracy of the existing total YORST risk scores to predict recidivism.**

This included assessment of:

- the percentage of cases of high risk offenders correctly classified
- the accuracy of the prediction across the range of risk total risk scores (total risk scores generated by the YORST are plotted against the percentage of the sample re-apprehended). This included the proportion correctly classified according to the current low, medium and high risk cut-off points
- comparison of differences in mean YORST total risk scores of those who have been re-apprehended compared to those who have not
- the strength of the relationship between the YORST total risk score and those who have been re-apprehended (point bi-serial correlation)
- Receiver Operating Characteristic (ROC) analysis and calculating and plotting the area under the ROC curve (AUC) statistic to produce an index of a tool's overall

---

9 All variables that included the Ministry of Justice Seriousness Scores required a log transformation to create a more normal distribution. Other re-coding is described throughout the report.

accuracy.<sup>10</sup> To test the ability of the tool to predict recidivism across different groups of offenders, individual ROC statistics are calculated for young offenders of different ages, sex and ethnicity.

**Chapter3 - Item response analysis to examine the characteristics of individual YORST items and their likely utility.**

- the response frequency of individual items (endorsement proportions) to detect items which apply to greater than 90% of the sample (i.e. if 95% of the sample all scored as having peers who are 'all known repeat offenders' – this item has limited scope in differentiating between higher and lower levels of risk)
- association of individual items with measures of recidivism to identify those most useful in predicting risk of recidivism
- inter-correlation among items, to identify potential redundancy of highly correlated items
- subjecting existing items to logistical regression analysis to examine their relative contribution.

**Chapter 4 - Logistic regression model building explored whether an alternative combination, formation or transformation of items or additional variables was better able to predict recidivism.**

- analysis of alternative predictor variables to identify an alternative combination, formation or transformation of items and variables can better predict recidivism
- development of a revised model using a half the sample randomly selected (construction sample)
- revised model then tested for its ability to predict recidivism using the second half of the sample (validation sample).

**Endnote:** The analysis in this report has focussed on identifying which items are the best predictors of recidivism. However, the **final selection of items will also assess the role of different items in identifying criminogenic needs and informing intervention planning** whilst meeting the other objectives of the YORST.

---

10 See section 2.6 for explanation on ROC analysis

## 2 Accuracy of existing YORST

The first part of this chapter describes the recidivism behaviour of those young offenders who had a YORST screen carried out between July and October 2009. This provides important descriptive information on the sample of youth on whom subsequent analysis is based. Following this there are a series of analyses that test the accuracy of the existing YORST to predict the recidivism behaviour of this group.

### 2.1 Sample characteristics

The LOST tool provided a summary of the criminal histories of n=1965 youth who completed a YORST screen between July and October 2009. **In the 12 months following the YORST screen n=1303 (66.3%) had been re-apprehended for at least one new offence.** Further analysis of different sub-groups of youth found:

- re-apprehension rate was higher for males (70.0%, n=1049 of 1499) than females (54.6%, n=254 of 465)
- Māori offenders had slightly higher rates of re-apprehension (69.1%, n=717 of 1037) compared to New Zealand European (63.8%, n=458 of 718)
- older offenders (14-16 years) had higher rates of re-apprehension (67.0%, n=828 of 1236) compared to child offenders (10-13 years) (57.4%, n=116 of 202).

Frequency, type and seriousness of re-apprehensions for the complete sample in this 12 month follow-up period are presented in table 2.1. There appeared to be a large range of recidivist behaviour in the sample. The maximum number of re-apprehensions for any one person was 54 (the average was 3.6 offences per person). The MOJ seriousness score for the most serious offence for any one individual was 6430 (considered maximum seriousness in the simplified YORST scale), with the average most serious offence score for the group being 120 (medium or simplified seriousness scale).

**Table 2.1 Characteristics of re-apprehension offences**

	Min	Max	Mean (SD)	N
Count of offences	0	54	3.6 (5.5)	1965
Count of violent offences	0	13	0.39 (1.0)	1965
Single most serious offence score	0	6430	120.7 (337.1)	1965
Average offence seriousness	0	6887.5	238.1(620.5)	1965

This chapter now examines how accurate the existing YORST was in predicting the recidivism behaviour of this group of young offenders.

## 2.2 Percentage of cases of high-risk offenders correctly predicted

To calculate the percentage of cases correctly predicted by the YORST as high risk the same methodology described by Baker et al., 2003 in the testing of the UK's risk assessment tool the ASSET was utilised. This involved splitting YORST scores into high and low risk at a point corresponding to the proportions actually re-apprehended (66.3%). The highest 66.3% of YORST scores were considered high risk scores and assumed to predict recidivism (i.e. re-apprehension); the lowest 33.7% of YORST total risk scores were assumed to predict non-recidivism. High scores actually re-apprehended (50.5%) and low scores not re-apprehended (17.7%) were then counted as correct (the two shaded boxes in table 2.2).

**Table 2.2 Percentage of cases correctly predicted (n=1965)**

	YORST Low Risk (scores of 0 to 41) (n=657)	YORST High Risk (scores 42 to 100) (n=1308)	Total (n=1965)
Not re-apprehended	347 (17.7%)	315 (16.0%)	662 (33.7%)
Re-apprehended	310 (15.8%)	993 (50.5%)	1303 (66.3%)

This resulted in an overall total accuracy rate of the existing YORST of 68.2% (17.7% plus 50.5%). This compares well to the UK's ASSET which is far a more comprehensive 106 item full risk assessment tool. Over 12 months the ASSET correctly predicted reconviction for 67% of a group of n=1081 young offenders (Baker et al., 2003). After 24 months, the accuracy of the ASSET to predict reconviction increased to 69.4% (Baker et al., 2005).

## 2.3 Accuracy of the prediction across the range of total risk scores.

Total YORST risk scores can range from 0 to 100. These scores have been categorised into low, medium or high risk based on the following cut-off points:

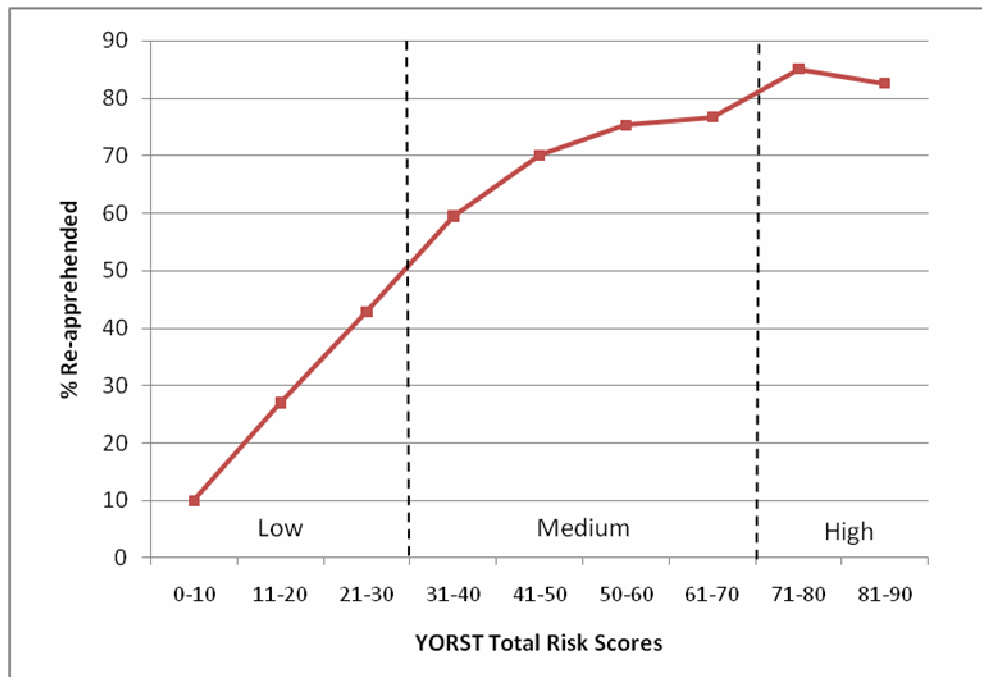
- Low risk = total YORST risk scores of 0 to 29
- Medium risk = total YORST risk scores of 30 to 69
- High risk = total YORST risk scores of 70 to 100.

Table 2.3 presents the percentage of cases re-apprehended for these three risk categories. The current cut-off points appear better able to differentiate between low and medium risk offenders, than medium and high risk, where both groups evidence a relatively high proportion who were re-apprehended.

**Table 2.3 Percentage of cases re-apprehended across low, medium and high YORST total risk scores**

	Re-apprehended		Not re-apprehended	
	n	%	n	%
Low risk (n=304)	100	32.7%	204	67.3%
Medium risk (n=1430)	1010	70.6%	420	29.4%
High risk (n=231)	193	83.5%	33	16.5%

The accuracy of the tool was further assessed by comparing the percentage of the sample re-apprehended for each of nine YORST risk score bands going up in ten point groupings (there are ten risk score bands possible, but in this sample there were no YORST scores over 90). Results appear in Figure 2.

**Figure 2.1 Existing YORST scores against percentage re-apprehended**

The graph shows a steady increase in the proportion of the sample re-apprehended as YORST total risk scores increased from 0 through to 50. Scores of over 50 show a flattening off, with a slight dip for scores of 81 to 90. This further supports the finding that the existing YORST total risk scores do a better job at distinguishing low to medium risk offenders than medium to high risk. In the future a review of the appropriateness of the cut-off points may be warranted. There were low numbers of young offenders with scores over 81 (n=23) as such the dip seen for this group may be the result of undue influence by single cases.

## 2.4 Mean YORST scores of those re-apprehended compared to those not

Another way to check the accuracy of the existing YORST was to compare the difference in mean total YORST risk scores of those who had been re-apprehended, to those that had not. Table 2.4 presents the results of an independent t-test showing there was a highly significant difference in total YORST risk scores between the two groups ( $p < 0.001$ ).

**Table 2.4 Difference in mean total YORST risk scores of those re-apprehended and those not re-apprehended**

	Mean (sd)	t	df	p
Re-apprehended	53.4 (15.4)	15.032 <sup>1</sup>	1127.2	.000
Not re-apprehended	40.7 (18.8)			

Table notes:

- 1 Levene's test for equality of variances was statistically significant so t score for 'equal variance not assumed' was used.

Mean total risk scores were also significantly different for youth who had been re-apprehended for violent offences ( $t=10.6$ ,  $p < 0.001$ ) and those who had been re-apprehended for medium seriousness offences or higher ( $t=17.0$ ,  $p < 0.001$ ).

## 2.5 Point bi-serial correlation

When comparing the ability of a tool to predict recidivism, the point bi-serial correlation is another common statistic.<sup>11</sup> The correlation of YORST total risk score against primary outcome was  $r=.34$ . This compares favourably to other well established more comprehensive full risk assessment tools. Olver et al., 2009 carried out a meta-analysis of three tools: Youth-adapted Psychopathy Checklist (PCL), the Structured Assessment of Violence Risk in Youth (SARVY) and the Youth Level of Service Case Management Inventory (YLS/CMI). The mean weighted  $r$  for predicting general recidivism for these tools was  $r=.33$  for the SARVY (across 7 studies),  $r=.32$  for the YLS/CMI (across 19 studies) and  $r=.33$  for the PCL (across 20 studies). The Australian version of the YLS/CMI-AA had an  $r=.28$  for predicting reconviction over 12 months (Thompson and Pope, 2005) and a  $r=.43$  for 24 months reconviction (Upperton and Thompson, 2007).

## 2.6 Receiver Operating Characteristics (ROC) analysis

In addition to point bi-serial correlation perhaps the most commonly used technique to assess and compare the accuracy the accuracy of risk assessment/screening tools is to carry out Receiver Operating Characteristics (ROC) analysis. A ROC graph plots the proportion of young people in the sample correctly identified as

<sup>11</sup> Point bi-serial correlation is the same as an ordinary correlation except one of the variables is binary with only two possible outcomes.



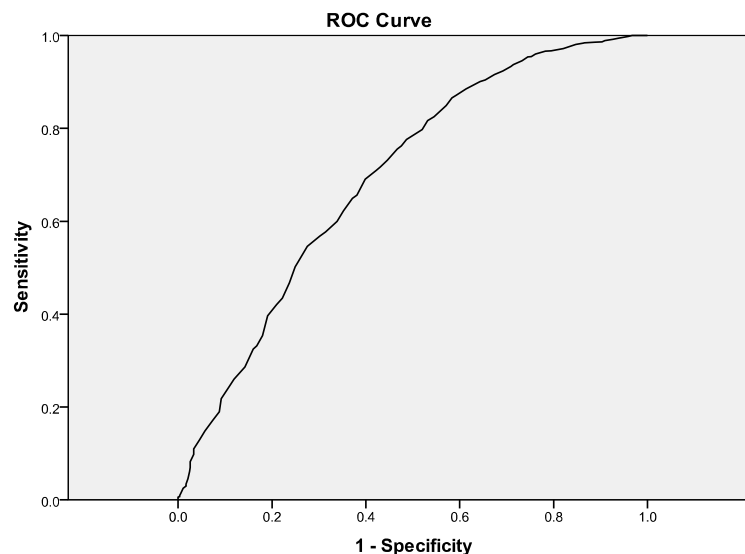
recidivists (sensitivity) against the proportion of offenders incorrectly identified as recidivist (1-specificity), across a range of different cut-off risk scores used to classify offenders as recidivists.

The area under the ROC curve (AUC) is an index of a tool's overall accuracy with scores produced ranging from 0 to 1 (0.5 indicates a chance-level accuracy; below 0.5 indicates a below-chance accuracy with tools incorrectly classifying the majority of offenders; above 0.5 indicates accuracy above-chance). A score of 1.0 indicates perfect discriminant accuracy or no false positive error. AUCs for an acceptable screening tool have been reported as being between 0.70 and 0.90 (Swets, 1988: cited in Vincent et al. (2009). Rice and Harris (2005) have produced tables to compare AUCs to other common measures of effect size (or significance of a statistical finding). They report that an AUC of .639 is equivalent to a moderate effect size, and one of .714 or higher is equivalent to a large effect size (highly significant).

ROC analysis was carried out on the total YORST risk scores to assess how accurately these were able to predict whether a young person was re-apprehended in the 12 month follow-up.

Figure 2.2 shows the ROC curve obtained. The AUC statistic was .695, a moderate effect size according to Rice and Harris (2005). This can be interpreted as a 69.5% probability that a randomly selected recidivist would score higher on the YORST total score than a randomly selected non-recidivist. When limited to the 'Special Group' of most reliable raters the AUC went up to .703, bringing it into the 'acceptable' range as defined by Swets, and just below the Rice and Harris's large effect size. The AUC score for the YORST again appears comparable to other more comprehensive tools. The AUC for the ASSET in predicting reconviction over 12 months was .712 (Baker et al., 2003) and .731 over 24 months. The YLS/CMI had a mean weighted AUC of .641 across 11 studies (Schwalbe, 2007) and the Australian version had an AUC of .67 over 12 months (Thompson and Pope, 2005) and .75 over 24 months (Upperton and Thompson, 2007). For a fuller comparison of YORST against other risk assessment tools see Mossman (2010a).

**Figure 2.2 ROC curve for the YORST total risk score (n=1965); AUC=.695**



Diagonal segments are produced by ties.

ROC analysis was also carried out to assess how well the tool predicted violent offending (AUC=.646) and offending of a medium seriousness (AUC=.702). For the Special Group respective scores were .647 and .738.

Finally ROC analyses were carried to see how well the tool performed for different sub-groups of offenders. Results appear in table 2.5, together with other ROC analysis results.

**Table 2.5 ROC analysis for existing YORST**

	N	AUC	SE	95% CI
<b>Complete sample</b>				
Any offending	1965	.695	.013	(.669, .720)
Violent offending	1965	.646	.014	(.618, .674)
Medium seriousness offending	1965	.702	.012	(.680, .725)
<b>Special Group sample</b>				
Any offending	330	.708	.030	(.649, .768)
Violent offending	330	.647	.034	(.581, .714)
Medium seriousness offending	330	.738	.027	(.685, .791)
<b>Sub-groups of complete sample</b>				
Males	1499	.694	.015	(.664, .724)
Females	465	.692	.025	(.643, .740)
New Zealand European	718	.724	.020	(.684, .764)
Māori	1037	.661	.019	(.624, .697)
10 to 13 years	202	.679	.040	(.601, .756)
14 to 16 years	1236	.691	.016	(.659, .723)

AUC statistics suggest the existing YORST appears to predict recidivism equally well for males and females, but is more accurate for New Zealand European than for Māori and also more accurate for older offenders. Closer examination of data suggested that Māori who were re-apprehended had higher YORST scores compared to New Zealand European young offenders who were re-apprehended. There are a number of possible reasons for this, either the raters may have inaccurately scored Māori on the YORST (inflated risk scores), or the factors included in the YORST are not effective in explaining Māori reconviction, or there is variability in the rate Māori young offenders are apprehended compared to New Zealand European. Unfortunately, it was not possible to determine which of these scenarios was the most likely explanation. Further research is required to assess the appropriateness of the YORST for young Māori offenders.

## 2.7 Summary

The accuracy of the existing YORST tool was comparable to other well established and more comprehensive risk assessment tools. YORST total risk scores correctly predicted re-apprehension status for 68.5% of the sample. Mean total risk scores were significantly higher for those who were re-apprehended for any offence, violent offences and those of medium seriousness or higher. Total risk scores had a point bi-serial correlation with re-apprehension status of  $r=.34$  and a moderate AUC of .695 for the entire sample and .703 for the Special Group of more reliable raters.

Analysis of the accuracy of the YORST across the score range found an increasing proportion of the sample was re-apprehended as total risk scores increased. However, the tool appeared better able to differentiate between low and medium risk offenders than between medium and high risk offenders. ROC analysis suggested the tool performed equally well for males and females, but was more accurate for New Zealand European compared to Māori and also more accurate for older offenders.



## 3 Item response analysis

Chapter two assessed the overall accuracy of the existing YORST in predicting re-apprehension. This chapter takes a close look at the individual items that make up the total YORST risk scores to examine their characteristics and how well they are able to perform in predicting recidivism. Their relative contribution to predicting recidivism is also explored using stepwise logistic regression analysis.

### 3.1 Item analysis – frequency of responses

There are 14 items scored between 0 and 5 that are added together to make up the total YORST risk score. The first consideration in examining the responses to these 14 items was whether the frequency of responses to individual items (endorsement proportions) has a distribution range that is going to be able to differentiate between the sample of young offenders. For instance, if over 90% of the sample are all receiving the same rating for an item (e.g. all have peers that are known repeat offenders – item 7) it will have limited scope in being able to differentiate between individuals of higher or lower levels of risk.

Endorsement proportions were all well below 90%; these ranged from 1.6% (item 13 – young person transient) through to 43.6 (item 9 – no previous care and protection history).

An earlier assessment of the quality of existing YORST data (Mossman, 2010b) found some items had more missing data than others. This held true for the current sample with item 10 (concern over alcohol and drugs) and item 14 (family history of offending) having the highest level of missing data (n=328 and n=236 respectively). In the current sample of 1965 young offenders, there were 1375 that had YORSTs completed with no missing data and just 5% of the sample had three or more items missing. However, missing data has implications for logistic model building which can only use cases with complete data. We know from Mossman (2010b) that unfortunately there was more missing data for older, high risk offenders. This means models developed may be biased towards younger lower risk offenders and potentially significant factors for the high-end offenders may be missed.

### 3.2 Inter-correlations between items

The next important consideration is the degree to which individual items are correlated with the recidivism and also with each other. Those with high correlations with re-apprehension are likely to be the most useful in predicting recidivism. Those with high levels of inter-correlation indicate potential redundancy among items. Table 3.1 provides the correlation of YORST items 1 to 14 with each other, and also with the two outcome measures (O1 is the primary outcome measure – re-apprehension for any offence in the 12 month follow-up period; O2Med represents re-apprehensions for any offence considered to be of medium seriousness or higher).

The first thing to notice is that all items were significantly correlated with both outcome variables, although, item 5 – seriousness of current offence, was only

significant to  $p < 0.05$ . These significant results are partly due to the large sample size the correlations are based on ranging from  $n=1637$  to  $1965$ . More useful is to consider the strength of the correlation coefficient ( $r$ ). Those shaded have an  $r$  of greater than  $0.25$  (a moderate correlation). The number of shaded boxes shows a high level of inter-correlation between items suggesting likely redundancy in terms of their ability to predict recidivism.

The strongest predictors in order of their strength were item 1 – time since last came to police notice for offending, item 3 – highest level of previous intervention, item 7- antisocial peers and item 2 – time since last came to notice for an incident. These items were also highly correlated with each other, along with many of the other items.

With highly inter-related items it is useful to identify their unique contribution once the relationship explained by other items has been accounted for (i.e. eliminate the redundancy among items). The multiple logistic regression analysis which follows assesses this.

**Figure 3.1 Correlations with YORST items 1 to 14 with each other and re-apprehension outcome variables (n=1637 to 1965)**

	O1	O2Med	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10	Item11	Item12	Item13	Item14
O1		<b>0.64</b>	<b>0.33</b>	<b>0.27</b>	<b>0.31</b>	<b>0.16</b>	0.06	<b>0.18</b>	<b>0.28</b>	<b>0.20</b>	<b>0.17</b>	<b>0.18</b>	<b>0.14</b>	<b>0.08</b>	<b>0.16</b>	<b>0.12</b>
O2Med	<b>0.64</b>		<b>0.30</b>	<b>0.27</b>	<b>0.31</b>	<b>0.18</b>	<b>0.10</b>	<b>0.21</b>	<b>0.28</b>	<b>0.21</b>	<b>0.21</b>	<b>0.17</b>	<b>0.15</b>	<b>0.09</b>	<b>0.19</b>	<b>0.15</b>
Item1	<b>0.33</b>	<b>0.30</b>		<b>0.45</b>	<b>0.55</b>	<b>0.27</b>	<b>0.10</b>	<b>0.35</b>	<b>0.43</b>	<b>0.33</b>	<b>0.31</b>	<b>0.34</b>	<b>0.21</b>	<b>0.10</b>	<b>0.34</b>	<b>0.22</b>
Item2	<b>0.27</b>	<b>0.27</b>	<b>0.45</b>		<b>0.42</b>	<b>0.22</b>	(0.03)	<b>0.23</b>	<b>0.31</b>	<b>0.30</b>	<b>0.33</b>	<b>0.28</b>	<b>0.20</b>	<b>0.09</b>	<b>0.30</b>	<b>0.23</b>
Item3	<b>0.31</b>	<b>0.31</b>	<b>0.55</b>	<b>0.42</b>		<b>0.31</b>	0.06	<b>0.37</b>	<b>0.43</b>	<b>0.39</b>	<b>0.38</b>	<b>0.41</b>	<b>0.20</b>	<b>0.11</b>	<b>0.32</b>	<b>0.26</b>
Item4	<b>0.16</b>	<b>0.18</b>	<b>0.27</b>	<b>0.22</b>	<b>0.31</b>		<b>0.09</b>	<b>0.17</b>	<b>0.29</b>	<b>0.13</b>	<b>0.33</b>	<b>0.10</b>	<b>0.21</b>	<b>0.20</b>	<b>0.27</b>	<b>0.30</b>
Item5	0.06	<b>0.10</b>	<b>0.10</b>	(0.03)	0.06	<b>0.09</b>		<b>0.39</b>	<b>0.22</b>	<b>0.13</b>	<b>0.08</b>	<b>0.09</b>	0.05	<b>0.07</b>	<b>0.15</b>	<b>0.08</b>
Item6	<b>0.18</b>	<b>0.21</b>	<b>0.35</b>	<b>0.23</b>	<b>0.37</b>	<b>0.17</b>	<b>0.39</b>		<b>0.41</b>	<b>0.34</b>	<b>0.31</b>	<b>0.43</b>	<b>0.23</b>	<b>0.13</b>	<b>0.43</b>	<b>0.28</b>
Item7	<b>0.28</b>	<b>0.28</b>	<b>0.43</b>	<b>0.31</b>	<b>0.43</b>	<b>0.29</b>	<b>0.22</b>	<b>0.41</b>		<b>0.35</b>	<b>0.33</b>	<b>0.41</b>	<b>0.22</b>	<b>0.21</b>	<b>0.41</b>	<b>0.37</b>
Item8	<b>0.20</b>	<b>0.21</b>	<b>0.33</b>	<b>0.30</b>	<b>0.39</b>	<b>0.13</b>	<b>0.13</b>	<b>0.34</b>	<b>0.35</b>		<b>0.30</b>	<b>0.38</b>	<b>0.19</b>	<b>0.11</b>	<b>0.41</b>	<b>0.23</b>
Item9	<b>0.17</b>	<b>0.21</b>	<b>0.31</b>	<b>0.33</b>	<b>0.38</b>	<b>0.33</b>	<b>0.08</b>	<b>0.31</b>	<b>0.33</b>	<b>0.30</b>		<b>0.31</b>	<b>0.40</b>	<b>0.12</b>	<b>0.48</b>	<b>0.38</b>
Item10	<b>0.18</b>	<b>0.17</b>	<b>0.34</b>	<b>0.28</b>	<b>0.41</b>	<b>0.10</b>	<b>0.09</b>	<b>0.43</b>	<b>0.41</b>	<b>0.38</b>	<b>0.31</b>		<b>0.20</b>	<b>0.06</b>	<b>0.43</b>	<b>0.26</b>
Item11	<b>0.14</b>	<b>0.15</b>	<b>0.21</b>	<b>0.20</b>	<b>0.20</b>	<b>0.21</b>	0.05	<b>0.23</b>	<b>0.22</b>	<b>0.19</b>	<b>0.40</b>	<b>0.20</b>		<b>0.17</b>	<b>0.41</b>	<b>0.42</b>
Item12	<b>0.08</b>	<b>0.09</b>	<b>0.10</b>	<b>0.09</b>	<b>0.11</b>	<b>0.20</b>	<b>0.07</b>	<b>0.13</b>	<b>0.21</b>	<b>0.11</b>	<b>0.12</b>	<b>0.06</b>	<b>0.17</b>		<b>0.18</b>	<b>0.27</b>
Item13	<b>0.16</b>	<b>0.19</b>	<b>0.34</b>	<b>0.30</b>	<b>0.32</b>	<b>0.27</b>	<b>0.15</b>	<b>0.43</b>	<b>0.41</b>	<b>0.41</b>	<b>0.48</b>	<b>0.43</b>	<b>0.41</b>	<b>0.18</b>		<b>0.44</b>
Item14	<b>0.12</b>	<b>0.15</b>	<b>0.22</b>	<b>0.23</b>	<b>0.26</b>	<b>0.30</b>	<b>0.08</b>	<b>0.28</b>	<b>0.37</b>	<b>0.23</b>	<b>0.38</b>	<b>0.26</b>	<b>0.42</b>	<b>0.27</b>	<b>0.44</b>	

Table notes: Figures in bold are statistically significant at  $p > .01$ ; un-bolded are statically significant at  $p < 0.05$ ; those in brackets are not statistically significant. Those shaded have the strongest correlation with  $r > .25$ .

### 3.3 Logistic Regression

The existing YORST items 1 to 14 were subjected to logistic regression analysis to reveal which were the best predictors of whether a young person was re-apprehended.

Items 1 to 14 were entered into a forward stepwise (wald) logistical regression. Results appearing in table 3.2 indicate just four items were significant predictors of re-apprehension (items 1, 2, 3 and 7). These items were combined into a statistically significant model that correctly predicted 72.3% group membership.

**Table 3.1 Logistic Regression analysis of existing YORST items**

Variables in the equation	B (SE)	Wald	df	Sig	Exp(B)
Constant	-1.018 (.142)	51.082	1	.000	.361
Item 1	.230 (.044)	27.381	1	.000	1.259
Item 2	.140 (.039)	13.044	1	.000	1.150
Item 3	.140 (.048)	8.618	1	.003	1.150
Item 7	.157 (.044)	12.744	1	.000	1.169

Table notes: % correctly predicted = 72.3;  $R^2 = .201$  (Nagelkerke) .145 (Cox & Snell); Model Chi-square=209.338,  $p = .000$ .<sup>12</sup>

The odds ratios (Exp[B]) provide the simplest way to assess the relative contribution of each item. An odds ratio of greater than 1 indicates that as the predictor increases, the odds of the outcome occurring increases. Conversely, a value of less than 1 indicates that as the predictor increases, the odds of the outcome occurring decreases. Item 1 had the greatest contribution in predicting re-apprehension with the highest odds ratio of 1.259, this was followed by item 7 (odds ratio = 1.169) and then equally by items 2 and 3 (odds ratio = 1.150).

This suggests for this sample of young offenders, the automatically generated items related to a young person's offending history are most useful in predicting recidivism: Item 1 and 2 time since last came to police notice for offending or incidents and item 3 highest level of previous intervention. The only subjectively rated dynamic factor that was significant was the number of influential peers known to police.

### 3.4 Modifications to existing YORST

Considering the item response analysis above it appears there are a number of possible ways the existing YORST could be revised to improve its predictive ability. These include:

- removing redundant items

<sup>12</sup> Logistic regression does not have an equivalent to the R-squared that is found in multiple regression; instead pseudo-R-square statistics are calculated Nagelkerke $R^2$  and Cox & Snell  $R^2$ . These statistics do not mean what R-squared means in multiple regression (the proportion of variance explained by the predictors), but still provide a gauge of the substantive significance of the model, with larger values indicating a better fitting model.



- applying weights to items based on logistic regression analysis
- re-coding of items to better reflect their ability to predict re-apprehension.

### 3.4.1 Removing redundant items

A new YORST total risk score was calculated based only on items 1, 2, 3 and 7. The removal of redundant items appeared to improve the prediction with ROC analysis producing an AUC of .723 (.787 for the Special Group). This suggests some of the existing items are erroneously associated with re-apprehension.

It could be argued that if all that was required from the YORST was a prediction of recidivism, it would be better to use just these four items. However, this is not the case with the YORST which also aims to capture information on criminogenic need. Hence, although not required for predicting risk of recidivism, retention of other items enables the collection of useful case management information. However, further work to revise items to improve their validity may be warranted.

### 3.4.2 Weighting of items based on logistic regression model

SPSS can use the logistic regression model reported in table 3.1 to calculate the probability of an event occurring (i.e. re-apprehension). These estimated probabilities for each case are equivalent to risk scores. A ROC analysis of these predicted probability scores allows assessment of whether a more sophisticated calculation of risk scores using different weightings of items based on the logistic regression model is superior to the current more simple cumulative risk assessment were responses to items are added together. This was a key question raised in Phase I of the YORST research (Mossman, 2010b). When a ROC analysis was carried out on these predicted probabilities, similar but not greatly improved levels of prediction were achieved (AUC = .729; and .788 for the Special Group).

### 3.4.3 Re-coding of items

Another question raised in Phase I of the YORST research was whether a smaller number of response items might better predict recidivism (Mossman, 2010b). Availability of recidivism data now allows this to be objectively assessed. Responses to each item were plotted against the mean of the re-apprehension outcome variable (see graphs in figure 3.1). This was done for the whole sample and also the Special Group (SG) to assess how the responses from the two groups compared.

These graphs show, for items 1 to 14, the proportion of the group re-apprehended for each response 0 to 5.<sup>13</sup> A steady linear incremental increase in the proportion re-apprehended with each increase in response option 0 to 5 would suggest the existing responses are usefully predicting re-apprehension. However, departures from this pattern show areas for potential improvement or reveal items that are not useful predictors (e.g. the graphed response options for item 5 and 13 which currently

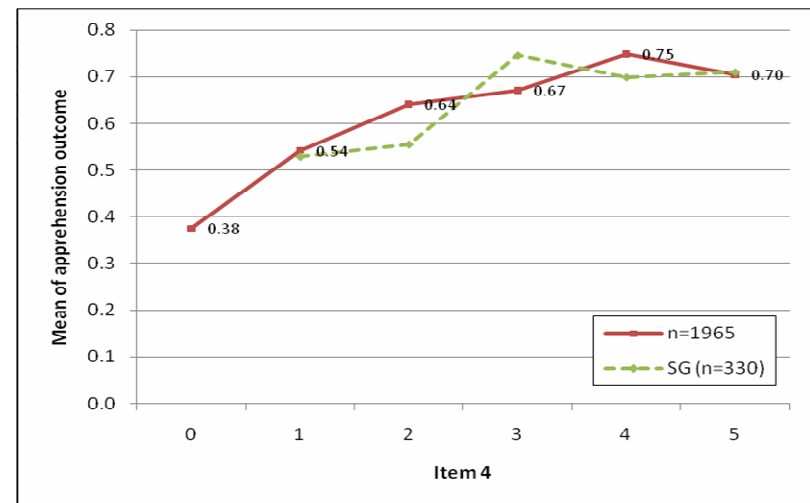
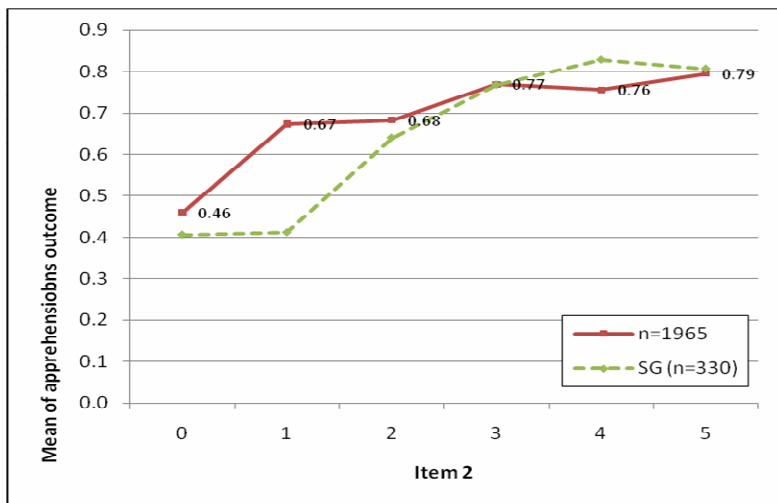
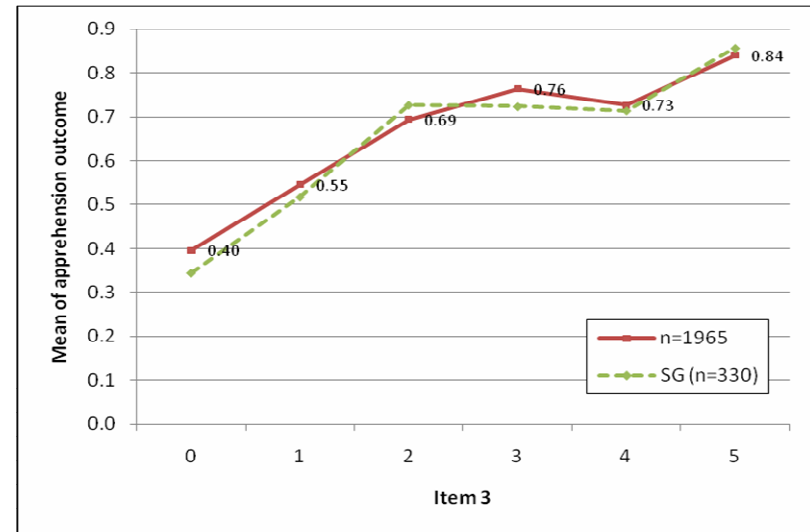
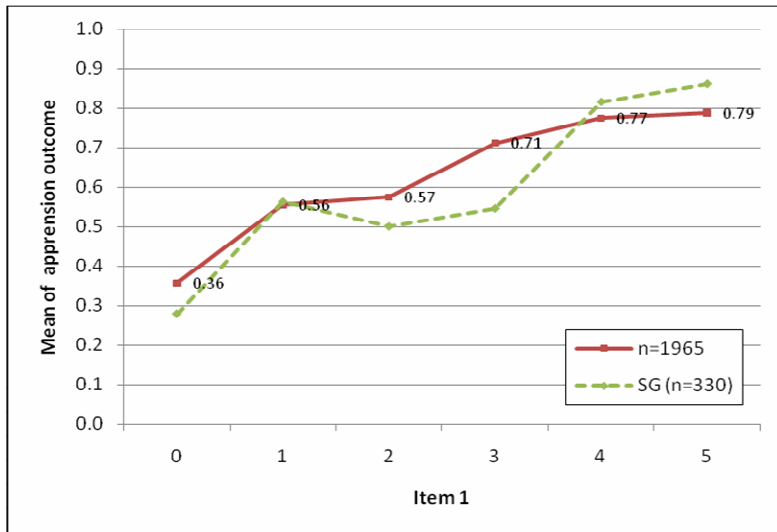
---

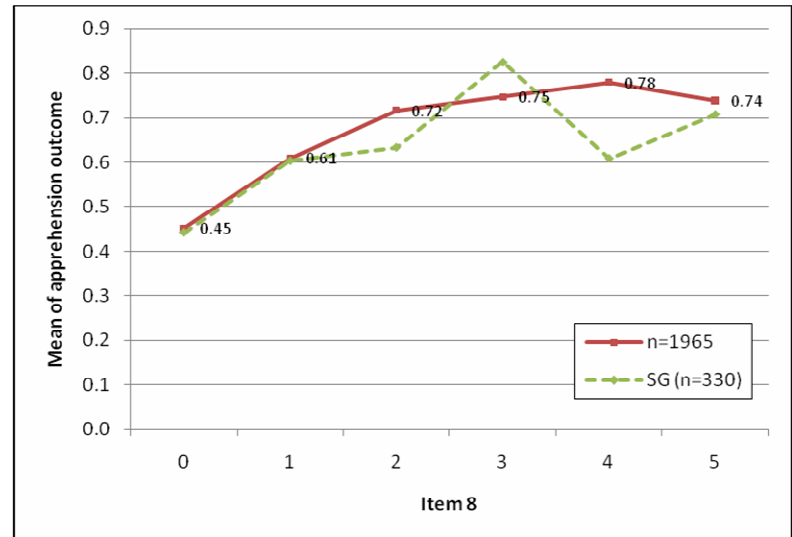
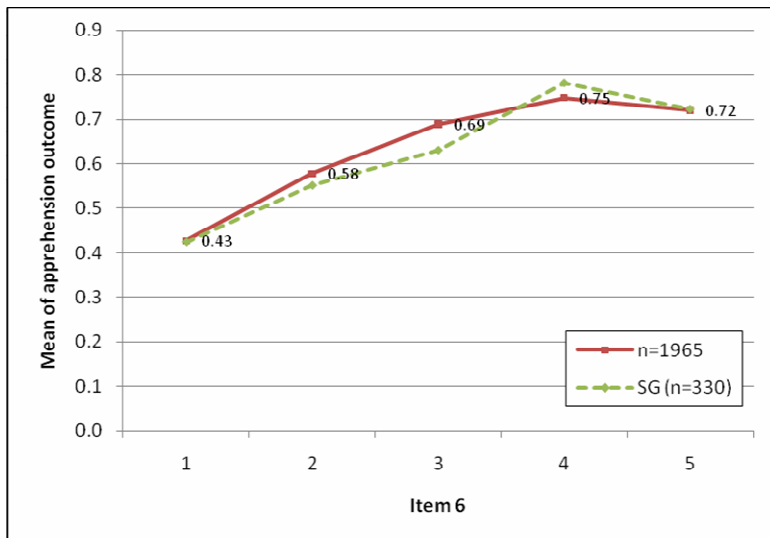
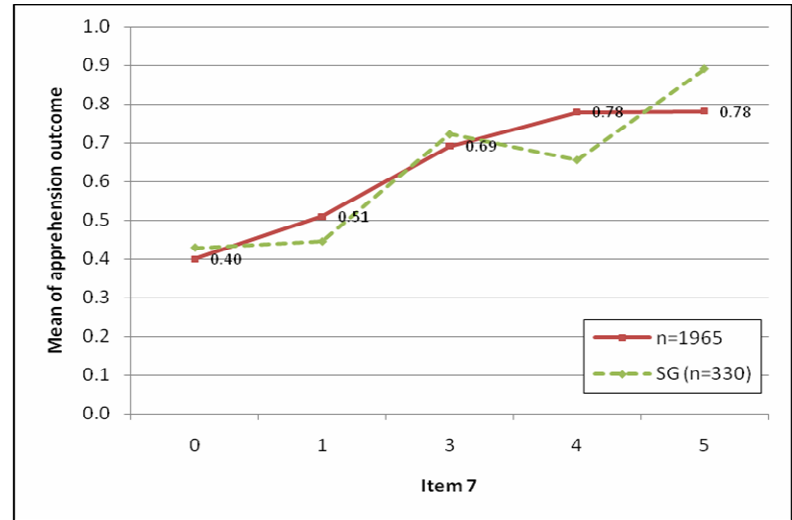
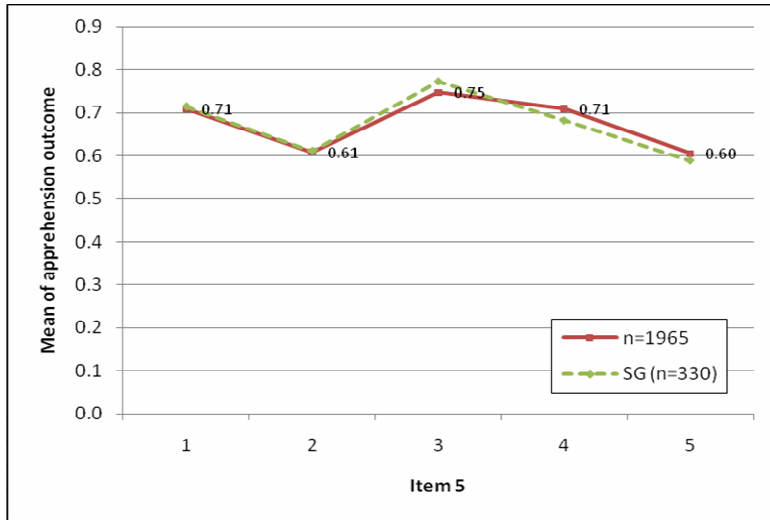
13 When the YORST was first developed response options for items varied slightly to provide some weighting to certain responses. Items 5 and 6 start at 1 not 0; item 7 has no response option of 2; items 9, 11, 12 and 14 have no response option of 1; item 10 had no response option of 3; item 11 has no response option of 1.

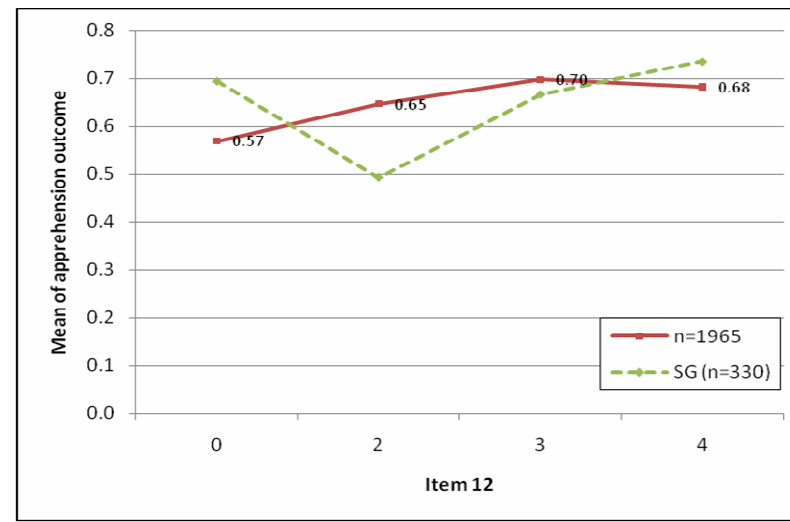
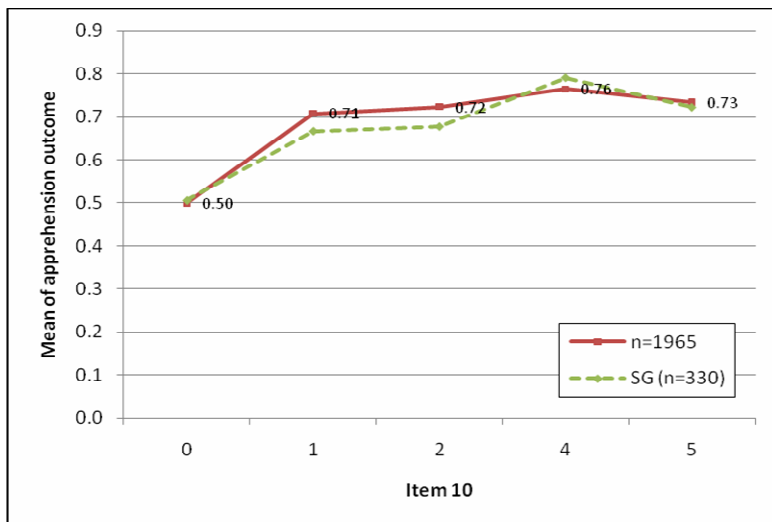
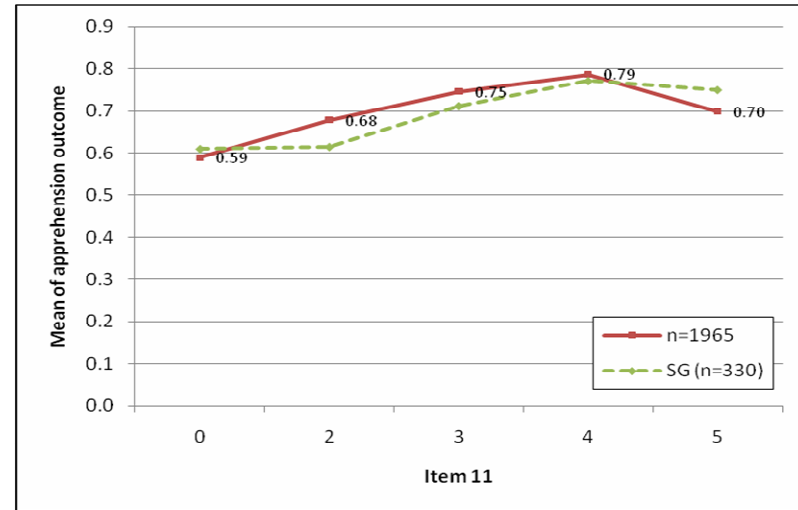
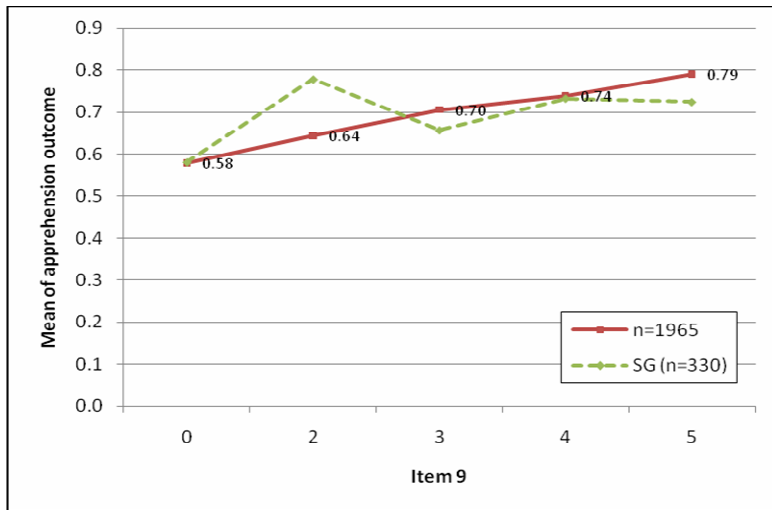
suggest no useful prediction of recidivism). Other graphs indicate re-coding could improve their utility. For example with item 1, response 1 (last came to police attention for offending - over 2 years) and response 2 (last came to police attention for offending - 1 to 2 years) have almost identical proportions of re-apprehensions (56 and 57%); hence, there is a strong case to combine these. Responses 3 to 5 also have very similar proportions of re-apprehensions (71 to 79%) and may benefit from being combined.

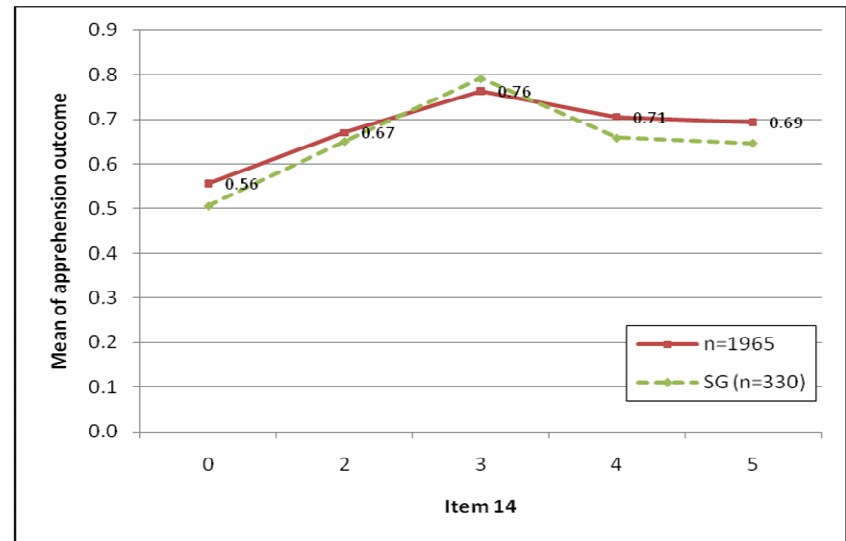
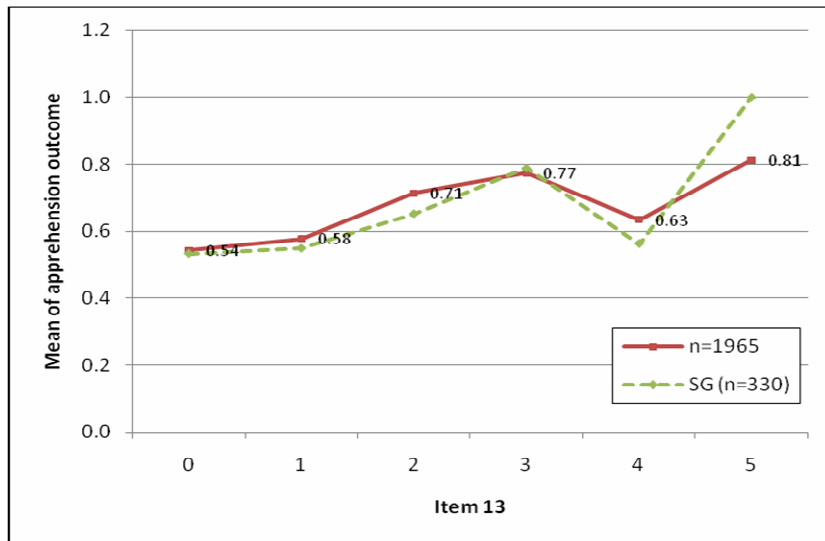
Each item was closely examined in this way and recoded with responses ranging from 0 through to 2 (new codes appears in Appendix 2).

**Figure 3.1 Individual item responses against proportion of group re-apprehended (mean of apprehension outcome)**









These re-coded items were then subjected to logistical regression to see which items might become more useful in predicting re-apprehension. Results appear in table 3.2.

The re-coded item 10 (concern over alcohol and drug use) now became a significant predictor along with the same re-coded items 1, 2, 3 and 7. This new model predicted a 72.9% correct group classification which was similar to the original items (see table 3.1). The predicted probabilities using this model achieved a slightly higher AUC of .735 (and .774 for the Special Group). However, overall the re-coding appeared to provide no great improvement on existing item responses.

**Table 3.2 Logistic Regression analysis of existing YORST items**

Variables in the equation	B (SE)	Wald	df	Sig	Exp(B)
Constant	-1.310(.170)	23.620	1	.000	.270
Item 1	.476(.098)	15.458	1	.000	1.610
Item 2	.312(.079)	5.76	1	.000	1.367
Item 3	.299 (.127)	11.387	1	.018	1.349
Item 7	.323 (.096)	4.645	1	.001	1.381
Item 10	.158 (.073)	59.231	1	.031	1.171

Table notes: % correctly predicted = 72.6;  $R^2 = .208$  (Nagelkerke); .149 (Cox & Snell); Model Chi-square=216.631;  $p = .000$ .

### 3.5 Summary

Item response analysis found a good distribution of responses across the sample suggesting the potential to be able to differentiate between high and low risk offenders. All items had significant correlations with re-apprehension status, but the strongest correlations with re-apprehension were item 1 – time since last came to police notice for offending, item 3 – highest level of previous intervention, item 7- antisocial peers and item 2 – time since last came to notice for an incident. There was a high level of inter-correlation among items suggesting some redundancy amongst these. This was confirmed with a logistical regression analysis which found only four items significantly contributed to the prediction of re-apprehension (Items 1, 2, 3 and 7).

A number of modifications to the existing YORST were examined to see if they were able to improve its predictive ability. This included the removal of redundant items (AUC of .723; .787 for the Special Group), applying weightings to items based on the logistic regression model (AUC = .729; and .788 for the Special Group), and re-coding of item-based responses to each item and their relationship with re-apprehension (AUC = .735; .774 for the Special Group). All modifications showed some improvement compared with the accuracy of the existing YORST. However, the most important finding from the item analysis was the identification of those which were or were not useful in predicting recidivism. This provides important information when considering the role of different items and how they may be improved to further enhance their primary objective, be it predicting recidivism or identifying areas of criminogenic needs or another of the YORSTs potential uses.





## 4 Logistic model building

This chapter presents findings from the next stage of the analysis which was to explore if an alternative combination, formation or transformation of predictor variables could better predict recidivism. An alternative model was developed using a construction sample and then tested on a validation sample.

### 4.1 Identifying alternative predictor variables

The LOST tool provided a range of alternative data on recidivist behaviour (see section 1.1.3). Some of these were similar to variables already included in the YORST (e.g., item 4 provided a grouping of age of first offence and the LOST provided a continuous variable of actual age of first offence). These alternative but similar variables were grouped with comparable YORST items and then examined to see if the alternatives offered any likely improvement in the prediction of re-apprehension. These groupings of variables and their correlations with re-apprehension are described below.

#### Frequency of offending

- *Item 1* (time since last came to police notice for offending) was compared with;
- count of the total number of offences prior; and also
- prior count recode (a re-coded version of this variable where 0 = no prior offences, 1=1 to 3 prior offences, and 2=4 or more prior offences); and with
- average yearly frequency of offences prior.

*Prior count recode* had the highest correlation with re-apprehension ( $r=.356$ ), followed by *item 1* ( $r=.332$ ) and then the *total count of offences prior* ( $r=.270$ ). *Average yearly frequency* was weakly but significantly negatively correlated with re-apprehension ( $r=-.088$ ).<sup>14</sup>

#### Seriousness of offending

- *Item 3* (highest level of previous intervention); was compared with
- sum of the seriousness scores of all prior offending;<sup>15</sup>
- single most seriousness score prior;
- the log of both these scores; and
- average yearly seriousness of offending prior.

*Item 3* had the highest correlation with re-apprehension ( $r=.311$ ). The only non-significant variable was average yearly seriousness prior. The log version of these

---

14 Average frequency can produce skewed data as a result of first time offenders with very short offending careers inflating their average frequency compared to those with longer offending careers. The same effect was anticipated with average seriousness of offending.

15 This variable was created by multiplying average seriousness by number of offences occurring in the same 12 months.

variables had improved correlations over the original seriousness variables ( $r=.292$  for log of sum of seriousness prior, and  $r=.257$  for log of single most serious offence prior).

### **Duration of offending career**

- *Item 4* (groupings of age of first offence); was compared with
- actual age of first offence; and
- length of offending career prior in years.

*Length of offending career* in years prior had the highest correlation ( $r=.209$ ), followed by *item 4* ( $r=.155$ ) and then *actual age* ( $r=.144$ ).

### **Engagement in education**

- *Item 8* (education /employment status); was compared with
- total truancy count.<sup>16</sup>

*Item 8* had the highest correlation ( $r=.198$ ). Truancy count was not significant and was not correlated with item 8 either.

### **Care and protection issues**

- *Item 9* (CYF Care & Protection history); was compared with
- *item 11* (Family Violence records);
- *Item 13* (concern in living situation); and
- *total count of family violence victim occurrences*.<sup>17</sup>

Item 13 had the highest correlation ( $r=.162$ ). The total count of family violence victim occurrences was weakly but significantly correlated ( $r=.064$ ).

## **4.2 Additional LOST predictors**

Other possible criminal history predictor variables provided by the LOST were also explored. Those found to be strong predictors included:

- Total counts of different types of offences (11 categories):
  - a) anti-social behaviour
  - b) burglary/trespass
  - c) violence
  - d) dishonesty.
- Other LOST specific measures:
  - a) log of lifetime LOST (cumulative seriousness scores - ever)

---

16 The 'truancy count' is calculated by LOST and relates to the number of truancy incidents (1Ts) against a young person recorded by police officers in NIA.

17 The number of 'family violence occurrences' is calculated by LOST and relates to the number of occurrences where the young person was recorded as a victim of family violence.

- b) log of most serious offence score (ever)
- c) number of violent offences prior
- d) total duration of offending in days (ever)
- e) number of offences classed as non-compliant (ever)<sup>18</sup>

All of these additional LOST variables are lifetime measures of offending (with the exception of 'number of violent offences prior'). Hence, despite strong correlations with re-apprehension, these could not be added into any model as they also included the behaviour that was being predicted. For example anti-social behaviour offences would be those occurring in the 12 month follow-up period, plus all others previous. In the future it might be worth exploring if similar variables could be calculated in LOST for the period prior to the 'intervention' in addition to the lifetime counts.

## 4.2 Logistic model building

Logistic regression was used to explore if alternative predictors identified above, when combined with the existing YORST items found to significantly predict recidivism, could more accurately predict re-apprehension.

To control for over-fitting of the data, the cases in the complete dataset were randomly assigned into one of two samples, the construction sample and the validation sample. Models were developed using the 'construction sample', and then re-tested on the second half of the sample (the validation sample).

Prior to model building the two samples were then compared to make sure there were no significant differences in their make-up on key outcome variables. No significant differences were found between the two samples on demographic make-up, YORST item responses and recidivism variables in the 12 month follow-up period (i.e. number of offences, number of violent offences, single most serious offence score, and cumulative seriousness).

Different combinations of predictor variables were examined using forward stepwise logistical regression to select the predictor variables that performed the best. Those listed below were the final selection of variables to be considered for the new model.

### Demographics

- age
- gender (categorical variable)
- ethnicity (categorical variable – NZE, Māori , Pacific, other)

### YORST items

- Item 1 (re-coded)
- Item 2

---

<sup>18</sup> Offences classed as non-compliant are those that relate to the offender being non-compliant with the enforcement of the law e.g. assaulting a police officer, breach of probation.

- Item 3
- Item 7
- Item 10 (re-coded)

#### LOST variables

- prior count of offences recode
- number of violent offences prior
- log of sum of all seriousness scores prior
- log of single most seriousness score prior
- length of offending career prior

These variables were entered into forward stepwise logistic regression and the findings appear in table 4.1.

**Table 4.1 Logistic Regression analysis of alternative variables and existing YORST items**

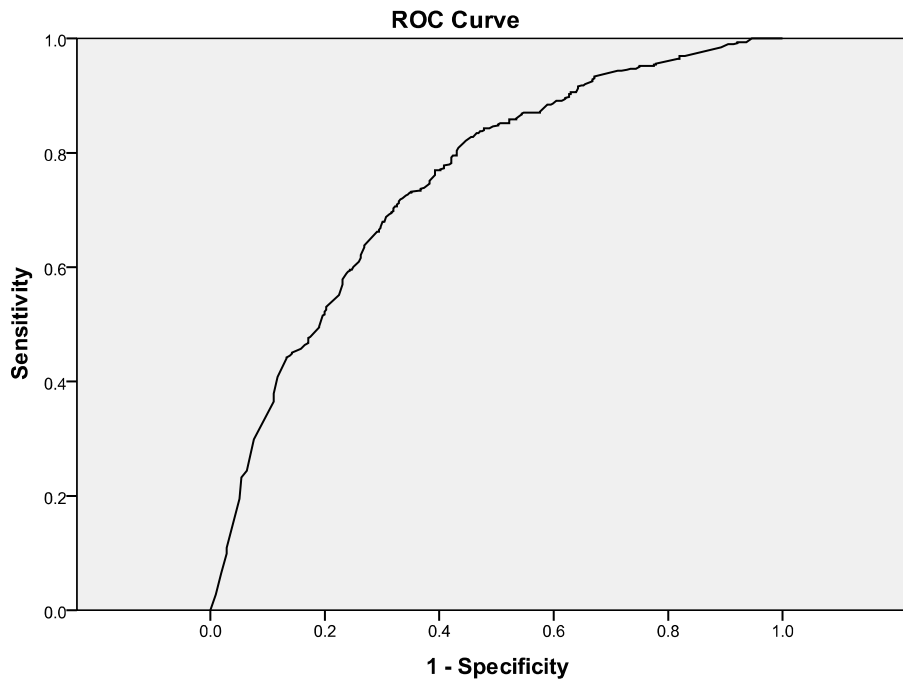
Variables in the equation	B (SE)	Wald	df	Sig	Exp(B)
Constant	--2.094 (312)	45.069	1	.000	.123
Gender (male)	.586(.193)	9.237	1	.002	1.797
Item 1 (re-code)	.441(.134)	10.843	1	.001	1.554
Item 2	.148 (.052)	8.214	1	.004	1.159
Item 7	.178(.060)	8.702	1	.003	1.195
Prior count of offences	.549 (.206)	7.120	1	.008	1.731

Table notes: % correctly predicted = 71.9;  $R^2 = .215$  (Nagelkerke); .154 (Cox & Snell); Model Chi-square=125.0;  $p=.000$ .

Just five variables were found to significantly predict re-apprehension. These were re-coded YORST item 1, and existing YORST items 2 and 7, together with being male and the total number of prior offences (re-coded). Being male had the greatest contribution to predicting re-apprehension with an odds ratio of 1.797, this was followed by total number of previous offences (odds ratio=1.731), then item 1 re-coded (odds ratio=1.554), then item 7 (odds ratio=1.195) and then item 2 (odds ratio=1.159). The overall performance of this model was very similar to the original model presented in table 3.1. There was a slight improvement in fit of the model with Nagelkerke  $R^2$  increasing to .215 compared to .201 in the original model.

The predicted probabilities produced by this model were subjected to ROC analysis, which resulted in an improved AUC=.748 (and .826 for the Special Group). Figure 4.1 shows the ROC curve obtained.

**Figure 4.1 ROC curve for revised model (construction sample n=982);  
AUC=.738**



Further ROC analysis suggested the model performed well across different sub-groups of offenders (see table 2.3). Overall AUC scores are respectable compared to other risk assessment tools, particularly for the Special Group of YORST raters.

**Table 4.2 ROC analysis for existing YORST**

	N	AUC	SE	95% CI
Construction sample	983	.748	.017	(.714, .782)
Construction sample (Special Group)	159	.826	.036	(.756, .896)
<b>Sub-groups of complete sample</b>				
Males	738	.742	.021	(.701, .783)
Females	224	.720	.035	(.652, .787)
New Zealand European	353	.786	.026	(.734, .838)
Māori	524	.726	.025	(.677, .774)
10 to 13 years	94	.762	.054	(.656, .868)
14 to 16 years	633	.726	.022	(.683, .769)

**4.2.1 Comparison of construction and validation sample**

The model presented in 4.1 was then applied to the validation sample. Predicted probabilities were calculated and then subjected to ROC analysis to assess how well

the model could be generalised to explain other data sets. Results presented in table 4.3 suggest the model maintained its predictive ability with this new dataset.

**Table 4.3 ROC analysis for existing YORST (validation sample)**

	N	AUC	SE	95% CI
Validation sample	982	.747	.018	(.712, .783)
Validation sample (Special Group)	171	.783	.038	(.707, .858)
<b>Sub-groups of complete sample</b>				
Males	761	.723	.022	(.679, .766)
Females	221	.793	.032	(.732, .855)
New Zealand European	365	.793	.026	(.742, .843)
Māori	513	.712	.027	(.658, .766)
10 to 13 years	108	.774	.048	(.680, .868)
14 to 16 years	603	.722	.024	(.679, .769)

This new model compares well to other actuarial models developed to predict recidivism in young offenders. In New Zealand a model was developed to predict recidivism of a group of young offenders who participated in a multi-systemic type programme in Christchurch (Reducing Young Offenders Programme). The model developed achieved an AUC of .74 on a validated sample and was based on three factors (number of prior intelligence notings, number of police recorded occurrences and a composite variable that included gender, ethnicity, age of first offending, number of prior care and protection notifications, number of prior court dates). In Australia a model based on four factors (age, educational status, frequency of contacts with criminal justice system) predicted reconviction of a sample of young offenders with an AUC of .763 (Weatherburn et al., 2007). For comparisons against other models see Mossman (2010a).

#### 4.2.2 Utility of the new model

International best practice recommends the use of both static and dynamic factors in assessing risk of recidivism. This is so that assessments can inform case management as well identifying level of risk. A limitation of the new model is the over-reliance on static historical factors (factors that cannot be targeted by an intervention to reduce the likelihood of recidivism). There is just one factor that is truly dynamic (item 7 – influential peers known to police), although, items 1 and 2 (time since last came to police notice for offending or incidents), while approximating a static factor (frequency of prior offending) have been designed in such a way as to be able to reflect change over time. However, overall the use of these items alone does little to inform case management and with the over-reliance on static factors the model will not be sensitive to changes in risk level over time.

#### 4.2.3 Summary

A range of alternative predictor variables were examined to explore if an alternative combination, formation or transformation of predictor variables could better predict

recidivism. Most of the alternatives concerned historical / static risk factors as they were summarised variables extracted from a young person's criminal history.

Logistic modelling was only able to identify two alternative variables that significantly predicted re-apprehension in a construction sample. These were consideration of a young person's gender and the number of prior offences in addition to YORST items 1, 2 and 7 already identified. The new model based on these variables showed a slight improvement on the initial model presented in table 3.1. This new model had comparable ROC results to other modifications presented in chapter 3 (AUC=.748; and .826 for the Special Group) and appeared more accurate than the initial model in its ability to predict re-apprehension status across different sub-groups. The model remained a good predictor of re-apprehension when re-tested on a new validation sample.

A limitation of the new model is the over-reliance on static historical factors. While providing a slightly more accurate prediction of risk, overall the model does little to inform case management and with the over-reliance on static factors means it will not be sensitive to changes in risk level over time.





## 5 Conclusion

The analysis presented in this report has shown that accuracy of the existing YORST and its prediction of recidivism is comparable to other well established and more comprehensive risk assessment tools (e.g. ASSET, YCL/CMI). It should be noted, however, that predicting recidivism in young offenders is not an easy task and none of the tools do exceptionally well at this.

Modifications to the existing YORST through the re-coding of items, applying weights to predictive items or including additional predictor variables were able to improve slightly its predictive ability. However, the most useful finding was the identification of which items are most important in predicting recidivism.

Analysis revealed that while the total risk scores significantly predict recidivism, in fact just four items that are contributing to this prediction (items 1, 2, 3, and 7). Three of these are static criminal history variables that are automatically generated (Items 1 – time since last came to police notice for offending, 2 – time last came to police notice for an incident, and 3 – highest level of previous intervention) with just one involving a subjective rating of a criminogenic need that could be targeted through intervention (item 7 – antisocial peers). Other items whilst each significantly correlated with re-apprehension, offered no additional predictive value once the variance explained by items 1, 2, 3 and 7 had been accounted for.

Model building also found a new variable, total frequency of prior offending, is also useful in predicting recidivism. However, while providing a slightly more accurate picture of risk, the predominantly static factors that made up this new model, mean it has little value in identifying areas of need to inform case management and are less sensitive to change in a young person's level of risk over time. In the future it would be good to devote more time into developing and re-designing dynamic risk factors that are more useful in predicting level or risk.

Identification of which items that are most useful in predicting recidivism has important implications for the practical application of the YORST. As noted at the start of this report, while the primary objective of the YORST is to identify high risk offenders, this is not its only objective. Identifying areas of need that can be targeted through interventions such as alternative action plans or youth development programmes is also of high importance. Other objectives are the collection of standardised information, assisting with information sharing among agencies and being able to measure change in risk level following an intervention. Having now established which items are most useful in predicting recidivism, items can now be reviewed and their role and potential utility in respect of the various objectives reconsidered. For example, items that more useful for identification of criminogenic need, could be redesigned to better capture this information.

The current screening process should also be reviewed in the light of these findings. It would be good to reconsider who should be given a YORST screen and whether they should be screened using all 14 items, or whether there may be advantages of screening a greater proportion of child and young offenders with a smaller number of

items. Those identified as high risk could then receive a fuller screen of risk and need.

# References

- Andrews, D. A., Bonta, J. and Wormith, J. S. (2006). 'The Recent Past and Near Future of Risk and/or Need Assessment'. *Crime Delinquency*, 52 (1): 7-27.
- Baker, K., Jones, S., Merrington, S. and Roberts, C. (2005). *Further Development of Asset*. London: Youth Justice Board for England and Wales.
- Baker, K., Jones, S., Roberts, C. and Merrington, S. (2003). *ASSET: The Evaluation of the Validity and Reliability of Youth Justice Board's Assessment for Young Offenders*. London: Youth Justice Board for England and Wales.
- Duncan, A. (2009). *Child and Youth Offending Statistics in New Zealand: 1992 to 2007*. Wellington: Ministry of Justice.
- Galletly, A. (2006). *Frequency and severity of offending young people in New Zealand: Descriptive analysis and development*. Unpublished Masters, Canterbury University, Christchurch.
- Ministry of Justice. (2002). *Youth Justice and Social Sector Service Delivery to the Children and Young People Convicted in relation to the death of Michael Choy*. November 2002. Wellington: Ministry of Justice.
- Ministry of Justice and Development, M. o. S. (2002a). *Report of the Ministerial Taskforce on Youth Offending*. April 2002. Wellington: Ministry of Justice and Ministry of Social Development.
- Ministry of Justice and Development, M. o. S. (2002b). *Youth Offending Strategy: Preventing and reducing offending and re-offending by children and young people*. April 2002. Wellington: Ministry of Justice and Ministry of Social Development.
- Mossman, S.E. (2010a). *Screening and assessment of young offenders risk of recidivism: Literature review*. August 2010. Final report prepared for New Zealand Police, Wellington.
- Mossman, S.E. (2010b). *Research to validate the YORST Phase I: Review of existing YORST data*. August 2010. Final report prepared for New Zealand Police, Wellington.
- New Zealand Police. (2009a). *YORST: Youth Offending Risk Screening Tool*. Internal document outlining policy guidelines for completing the YORST. Police and & CYF Police, 25/05/09.
- New Zealand Police. (2009b). *Request for Proposal: Research to Validate the New Zealand Police Youth Offending Risk Screening Tool (YORST) - First year*. Wellington: New Zealand Police.
- Olver, M. E., Stockdale, K. C. and Wormith, S. (2009). 'Risk assessment with young offenders: A meta-analysis of three assessment measures'. *Criminal Justice and Behaviour*, 36: 329-353.
- Rice, M., and Harris, G. (2005). Comparing Effect Sizes in Follow-up Studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, 29 (5), 615-620.
- Schwalbe, C. S. (2007). 'Risk assessment for juvenile justice: A meta-analysis'. *Law and Human Behavior*, 31 (5): 449-462.
- The Werry Centre. (2009). *A Literature Review: Mental health and alcohol and other drug screening, assessment and treatment for youth justice populations*.

Auckland: The Werry Centre for Child and Adolescent Mental Health Workforce Development.

- Thompson, A. P. and Pope, Z. (2005). 'Assessing juvenile offenders: Preliminary data for the Australian Adaptation of the Youth Level of Service/Case Management Inventory (Hoge & Andrews, 1995)'. *Australian Psychologist*, 40 (3): 207-214.
- Upperton, R. A. and Thompson, A. P. (2007). 'Predicting juvenile offender recidivism: Risk-need assessment and juvenile justice officers'. *Psychiatry Psychology and Law*, 14 (1): 138-146.
- Vincent, G., Terry, A. and Many, S. (2009). 'Chapter 12: Risk/Needs Tools of Antisocial Behavior and Violence Among Youthful Populations'. In, J. T. Andrade (Ed.), *Handbook of Violence Risk Assessment and Treatment: New Approaches for Mental Health Professionals*. New York: Springer Publishing Company, LLC.
- Weatherburn, D., Cush, R. and Saunders, P. (2007). 'Screening juvenile offenders for further assessment and intervention'. *Crime and Justice Bulletin*, (109): 1:11.
- Wilson, N. and Rolleston, R. (2004). *A Risk-Need Profile Using Four Measures for Youth Offenders Incarcerated in Young Offender Units*. Retrieved 7 December, 2009, from <http://www.corrections.govt.nz/research/a-risk-need-profile-youth-offenders.html>

# Appendices

## Appendix 1: The YORST



### YOUTH OFFENDING RISK SCREENING TOOL

NAME						NIA Person ID No:	
(Child/YP):		Surname		First name(s)		File no:	
DOB		Age	Gender		Male	Female	Date RST Completed by (QID)
ETHNICITY	European	Pacific	Asian	Other			
	Maori		Iwi		Hapu		
Incident / Offence Code		Incident / Offence Description					

#### Part (A) Offending Factors

Time since last came to Police notice for their <b>offending</b> ?						
1	No previous	Over 2 yrs	1 to 2 yrs	Less than 1 yr	1 to 6 mths	Under 1 mth
	0	1	2	3	4	5
Time since last came to Police notice for <b>incidents</b> (e.g. 1J, 2M, 1T) relating to them and/or serious behaviour incident at school?						
2	No previous	Over 2 yrs	1 to 2 yrs	Less than 1 yr	1 to 6 mths	Under 1 mth
	0	1	2	3	4	5
Highest level of previous intervention? (final outcome)						
3	No previous	Noting	Warning	Alt. Action	FGC	Youth Court
	0	1	2	3	4	5
At what age was offending first reported to Police (if first offence use current age)?						
4	No offences	15+	14	13	10 to 12	Under 10
	0	1	2	3	4	5
Rate the seriousness of the current primary offence using the youth offence rating tool (see A4 list).						
5	Minimum	Minimum / Medium	Medium	Medium / Maximum	Maximum	
	1	2	3	4	5	
Is the nature (MO) of current or previous offending of a concerning nature?						
6	Very Low	Low	Medium	High	Extreme	
	1	2	3	4	5	
Comments re Question 6:						

#### Part (B) Peer Group Factors

Influential peers known to Police?						
7	None	Very few known	Some known	Many known	All known repeat offenders	Unknown
	0	1	3	4	5	0

#### Part (C) Education / Employment Factors (contact the school, but not the employer)

Current school / education / course or employment status							
8	Full time well engaged	Full time some issues	Mostly attends	Irregular attendance	Stood down / suspended	Not attending (school / job)	Unknown
	0	1	2	3	4	5	0

#### Part (D) Care & Protection History

Has a notification been made to CYF for this family or child / young person?					
9	No	Notification concerning another sibling	Notification concerning this child / young person	Some form of intervention provided by Child, Youth & Family	Currently / previously in the custody of CYF (101 status)
	0	2	3	4	5

#### Part (E) Alcohol and/or Drug Use

Is their use of alcohol or drugs causing concern? (consider the long term effects of the type of drugs used).						
10	No concern	Slight	Moderate	Serious	Very Serious	Unknown
	0	1	2	4	5	0

Part (F) Family Factors						
If there are FAMILY VIOLENCE records in NIA for this family / address, what is the highest FV score?						
11	Zero Records	Records, but no score	Score from 1 - 8	Score from 9 - 16	Score 17 or over	
	0	2	3	4	5	
Where do they live? (socio economic area decile rating of local state primary school)						
12	8 - 10	4 - 7	2 - 3	1	Transient / Motor Camp	
	0	2	3	4	5	
Are there concerns in the living situation? e.g. parent / caregiver support and supervision of child / young person, parental mental health problems, drug and alcohol use, suspected child abuse and / or unrecorded family violence						
13	None	Very minor concerns	Some concerns	Major concerns	Some major concerns	Young Person Transient
	0	1	2	3	4	5
Detail Concerns:						
Family members have offending history?						
14	None	Parent(s) with minor history	Parent/s with major history (imprisonment)	Parent(s) have offended within past 12 months	Sibling(s) have offended within last 12 months	Unknown
	0	2	3	4	5	0
Any General Comments:						
Information Sources						
Spoken To		Child / young person	Parent / caregiver	School / course provider / MOE	Child Youth & Family	Other agency
	This time					
	Previously					
	Not At All					
Scoring Instructions				Risk Screening		
Questions			Answers		YORST Score	
No. of Questions	Max	Sum of the Scores (Above)		=		%
Not Answered:	Answered:	x 5	Max. Total for Answered Questions		=	x 100 =
Dynamic Risk Factors						
Static Factor Results					Dynamic YORST Score	
Sum of Dynamic Factors						
Maximum Possible Total for Dynamic Factors					45	x 100 =
Youth Aid Response						Your Station
Warning	AA	FGC	Youth Court	Police Youth Development	Other	

## Appendix 2: Re-coding of existing YORST

YORST Item	Revised coding for responses		
	0	1	2
Item 1	0	1 or 2	3, 4 or 5
Item 2	0	1 or 2	3, 4 or 5
Item 3	0 or 1	2, 3, or 4	5
Item 4	0	-	1, 2, 3, 4, 5
Item 5	0	-	1, 2, 3, 4, 5
Item 6	1	2	3, 4 or 5
Item 7	0	1	3, 4 or 5
Item 8	0	1	2, 3, 4, or 5
Item 9	0	-	2, 3, 4, or 5
Item 10	0	-	1, 2, 4, or 5
Item 11	0	-	2, 3, 4, or 5
Item 12	0	-	2, 3, or 4
Item 13	0 or 1	-	2, 3, 4, or 5
Item 14	0	-	2, 3, 4, or 5