Author: A. McGrath and A. P. Thompson

Title: The relative predictive validity of the static and dynamic domain scores in risk-need assessment of juvenile offenders

Journal: Criminal Justice and Behavior **ISSN**: 0093-8548 1552-3594

Year: 2012 **Volume:** 39 **Issue:** 3 **Pages:** 250-263

Abstract: This study examined the predictive validity of the Australian Adaptation of the Youth Level of Service/Case Management Inventory (YLS/CMI-AA). The focus was on the subcomponents of the inventory, which represent one static and seven dynamic risk-need domains. Reoffending outcomes within 1 year of the inventory were obtained for a large sample (N = 3,568) of young people under juvenile justice supervision in the community. Logistic regression analyses investigated the relative contribution of YLS domain scores. The results showed that the static and four dynamic domain scores independently predicted recidivism and that the combination of those domain scores yielded a small improvement in prediction. A similar pattern of results was obtained from analyses of the simple additive scores for the YLS domains. The findings support the YLS/CMI-AA total score as a sufficiently useful predictor of risk, and they clarify the contribution of static and dynamic risk components.

URLs: http://researchoutput.csu.edu.au/R/-?func=dbin-jump-full&object_id=31822&local_base=GEN01-CSU01; http://dx.doi.org/10.1177/0093854811431917

Author Address: amcgrath@csu.edu.au

CRO Number: 31822

The relative predictive validity of the static and dynamic domain scores in risk-need assessment of juvenile offenders

Andrew McGrath

&

Anthony P. Thompson

Key words: Juvenile Offending, Risk Assessment, Recidivism, Youth Level of Service/Case Management Inventory

Author note

Andrew McGrath, School of Psychology, Charles Sturt University; Anthony P. Thompson, School of Psychology, Charles Sturt University.

The authors thank the NSW Department of Attorney General and Justice (Juvenile Justice) and the NSW Bureau of Crime Statistics and Research for assistance in undertaking the research. The opinions here do not necessarily reflect the views of these organizations, or any of their officers.

Correspondence concerning this article should be addressed to Andrew McGrath, School of Psychology, Charles Sturt University, Building C6, Panorama Avenue, Bathurst, NSW 2795, Australia. Email: amcgrath@csu.edu.au

Abstract

This study examined the predictive validity of the Australian Adaptation of the Youth Level of Service/Case Management Inventory (YLS/CMI-AA). The focus was on the sub-components of the inventory which represent one static and seven dynamic risk-need domains. Reoffending outcomes within one year of the inventory were obtained for a large sample (N = 3568) of young people under juvenile justice supervision in the community. Logistic regression analyses investigated the relative contribution of YLS domain scores. The results showed that the static and four dynamic domain scores independently predict recidivism, and that the combination of those domain scores yielded a small improvement in prediction. A similar pattern of results was obtained from analyses of the simple additive scores for the YLS domains. The findings support the YLS/CM-AA total score as a sufficiently useful predictor of risk and clarify the contribution of static and dynamic risk components.

The Role of Static and Dynamic Factors in Risk-Need Assessment of Juvenile Offenders

It is widely accepted that risk factors and treatment needs should be considered jointly when working with juvenile offenders (Day, Howells, & Rickwood, 2004; Hoge & Andrews, 1996). This dual approach applies broadly within correctional and forensic psychology, but the framework is especially relevant for young offenders. Juvenile justice legislation typically emphasizes protection of the public **and** guidance and assistance to young people as dual principles underpinning society's response to juvenile crime (Thompson, 2001). Within psychology, it is well-recognized that risk factors and treatment needs can be credibly and usefully assessed with structured psychometric inventories (Clements, 1996; Hollin, 2002; Lally, 2003). This approach has been broadly adopted in community juvenile justice systems and numerous inventories exist in Canada (Hoge & Andrews, 2006, 2011), the United States (Gavazzi, et al., 2003; Howell, 1995; Schwalbe, Fraser, Day, & Arnold, 2004), England and Wales (Baker, Jones, Roberts, & Merrington, 2003) and Australia (Thompson & Pope, 2005; Thompson & Putnins, 2003).

Conceptually, the distinction between static and dynamic risks turns on whether a factor that is associated with offending can be modified. Historical precursors (e.g., age at first offence, number of previous convictions) are useful for predicting risk of reoffending, but are not amenable to change in order to reduce this risk. By contrast, dynamic risk factors concern conditions that are currently related to offending and are potentially modifiable. The inclusion of static and dynamic risk factors in assessment inventories makes logical sense but it is also justified on other grounds. Foremost is the longstanding and substantive literature on factors that are associated with juvenile offending. This literature is both empirical and theoretical. In essence, the literature shows: 1) some unalterable aspects of demography and anti-social history are associated with propensity to crime and reoffending (Cottle, Lee &

Heilbrun, 2001; Nagin & Paternoster, 2000; Piquero, Farrington, & Blumstein, 2003), 2) a variety of modifiable psycho-social factors are associated with crime and risk is proportional to number and severity (Durlak, 1998; Farrington, 2002; Hubbard & Pratt; 2002), 3) the interplay of psycho-social risk factors can be theoretically constructed into coherent explanations of offending patterns (Moffitt, 1993; Paternoster & Brame, 1997; Sampson & Laub, 2003), and 4) psycho-social interventions that are appropriately targeted and delivered can reduce juvenile offending (Borduin, et al., 1995; Loeber & Farrington, 1998). A second pillar on which structured risk-need assessment rests is the compelling correctional strategy (*risk-need-responsivity)* promoted by Andrews and colleagues for the past 20 years (Andrews, Bonta, & Wormith, 2006; Andrews & Dowden, 2006; Andrews, et al., 1990). In summary, there is a solid rationale and sound evidence for attention to static and dynamic risk factors when working with juvenile offenders. However, more attention needs to be given to how dynamic risk factors are represented in risk-need inventories and their psychometric integrity. In the research reported here, we investigated static and dynamic factors in one particular risk-need inventory for youth called the Australian Adaptation of the Youth Level of Service /Case Management Inventory (YLS/CMI-AA) (Hoge & Andrews, 1995; Thompson & Pope, 2005).

Like its parent version (YLS/CMI; Hoge & Andrews, 2006) and similar youth adaptations of the adult inventory (Andrews, Bonta, & Wormith, 2004), the YLS/CMI-AA incorporates a collection of checklist items that represent static and dynamic domains. A healthy research literature supports the Youth Level of Service family of inventories (YLS). For example, recent meta-analytic studies have found that the predictive validity for general recidivism is good. Schwalbe (2007) found a mean weighted effect size AUC of .641 based on 11 YLS studies. Olver, Stockdale and Wormith (2009) found a mean weighted correlation of .32 based on an overlapping but larger sample of 19 studies. These predictive validity indices were between YLS inventory total score and recidivism. Less attention has been paid to the relative contribution of static versus dynamic component scores to prediction. Olver et al. considered pursuing this in their meta-analysis but were dissuaded because such information was typically not provided in the studies reviewed. The relationship of static and dynamic risk factors to juvenile recidivism was examined in a meta-analysis conducted by Cottle et al. (2001). In the 22 studies that were examined, two static offence history variables were most strongly associated with recidivism and a number of dynamic family and social variables were also significant predictors. This meta-analysis only provides general support though for the inclusion of such components in risk-need assessment. The studies reviewed varied considerably on key variables such as age of juveniles and criterion for recidivism. Moreover, only six of the studies incorporated variables into a formal risk assessment approach. A similar pattern of findings was observed in a meta-analysis of the adult recidivism literature examining 131 studies (Gendreau, Little, & Goggin, 1996). Moderate effect sizes were observed for a number of static and dynamic predictors, with some composite evidence supporting the incremental validity of dynamic predictors.

Although few YLS studies address the predictive contribution of static and dynamic components, some relevant information is available. At the bivariate level, domain scores are typically found to correlate significantly with recidivism. Small to moderate coefficients are reported in a number of published (Thompson & Pope, 2005; Upperton & Thompson, 2007) and unpublished studies (Flores, Travis, & Latessa, 2003; Hoge & Andrews, 2006). In several of these (Flores, Travis, & Latessa; Thompson & Pope; Upperton & Thompson), the domain dealing with prior and current offences was the highest bivariate correlation, although not necessarily significantly stronger than some dynamic domains. Holsinger, Lowenkamp and Latessa (2006) found that the personality domain had the highest bivariate correlation with institutional misconducts.

The relative contribution of domain scores to predictive validity is approached in some studies but limited statistical analysis and methodological shortcomings qualify conclusions. For example, Marshall, Egan, English and Jones (2006) found that AUC indices for several file based measures of charges and convictions were virtually the same for the total YLS score as for the total score calculated from the dynamic domains. However, the criterion measures were retrospective in a relatively small sample of 94 male and female youth. Marczyk, Heilbrun, Lander, and Dematteo (2003, 2005) investigated the relationship between the YLS domain scores and decisions to certify youthful offenders to adult or juvenile court, as well as to subsequent arrests. The results were a mixed picture with scores from the YLS static and some dynamic domains showing predictive validity for certification decisions but not for recidivism. However, small sample size (N = 72 to 95), file based scoring, and acknowledged limitations in the measure of recidivism do substantially qualify the findings. In two related reports (Onifade, Davidson, Campbell, et al., 2008; Onifade, Davidson, Livsey, et al., 2008), the predictive validity of the instrument was assessed by regressing the eight YLS domain scores against a dichotomous recidivism variable. The logistic regression model accounted for a significant amount of variance in samples of intake youth (N = 287) and youth already on probation (N=238). However, the relative statistical contribution of each domain score was not elaborated further. In a large sample of approximately 1300 juvenile offenders, logistic regression analysis of domain scores showed only prior offenses, substance abuse and attitudes/orientation significantly predicted re-arrest, although the result related to attitude and orientation was not in the predicted direction(Flores, et al., 2003).

The contribution of YLS domain scores to predictive validity thus remains unclear. Apart from small sample sizes, variations in the criterion measure and different follow-up periods, the relative validity of static and dynamic domains has not been as thoroughly explored as it could be. For example, logistic regression has been used in a number of the YLS studies to show which static and dynamic domain scores contribute to prediction while controlling for the influence of other domains. However, the relative or incremental prediction of dynamic versus static domains has typically not been pursued through such analyses. One study that did examine predictive and incremental validity with the YLS only made comparisons between three adolescent risk assessment instruments, rather that within the YLS itself (Welsh, Schmidt, McKinnon, Chattha, & Meyers, 2008).

There are several reasons why it is important to advance understanding of the predictive validity of static and dynamic domains on risk-need inventories such as the YLS. First, although the theoretical, empirical and practical relevance of dynamic risk factors for juvenile crime is firmly established (Farrington, 1995; Gendreau, et al., 1996), translating these constructs into sound measures requires rigorous evaluation in keeping with psychological test standards (American Educational Research Association, 1999). Second, although third generation assessments are developed on the assumption that combing static and dynamic scores yields the best prediction of recidivism (Andrews, et al., 2006; Benda, Corwyn, & Toombs, 2001), empirical support for this needs to be more clearly demonstrated. The literature reviewed here suggests that some dynamic domains as measured on risk-need inventories contribute to prediction but that others are redundant. This gives rise to another reason for clarifying the relative contribution of the dynamic domains. Inventories such as the YLS already incorporate built-in weighting of dynamic domains by virtue of the number of items representing the domain. On the YLS/CMI (Hoge & Andrews, 2006) and YLS/CMI 2.0 (Hoge & Andrews, 2011), this varies from three to seven items. Disaggregating the relative predictive validity of dynamic domains may support such weighting, inform alternative weightings, or even support uncoupling risk from needs analysis as some have suggested (Howell, 1995).

It is important to note that these predictive validity considerations should not be construed as undermining the importance of assessing dynamic risk factors and intervening to address needs, improve lives and reduce criminal risk. However, if a key component of that well-supported correctional framework is its manifestation in contemporary risk-need inventories, then those inventories should serve that agenda psychometrically as strongly as possible. The research reported here contributes to this goal. The study was based on YLS/CMI-AA data for a large sample of juvenile offenders in New South Wales, Australia. Multivariate analyses were used to investigate the predictive and relative validity of static and dynamic domain scores.

Method

Data

Approval to access the data for this study was granted by the appropriate university and government ethics committees. Data were provided by the New South Wales (NSW) Department of Attorney General and Justice (Juvenile Justice) which in October 2002 adopted the YLS/CMI-AA as part of its assessment policy and procedures. The instrument includes risk/need items over eight domains: (1) *prior and current offences*; (2) *family and living circumstances*; (3) *education/employment*; (4) *peer relations*; (5) *substance abuse*; (6) *leisure/recreation*; (7) *personality/behavior*; (8) *attitudes/ beliefs*. The inventory consists of 47 items with all but one scored 0 or 1, depending on whether the risk factor is judged to apply to the young person or not. One item, related to age at first court order, is scored 0, 1, 2 with more weight given to younger offenders. Domain subscores vary between 3 and 9 while the total score can range from 0 to 48. The YLS/CMI-AA also includes three dichotomous items concerning strengths that may exist at the individual, family, or community level. These items though are to inform case planning and were not used in the current research.

In a computerized format, the inventory was completed for all young persons under various forms of supervision in the community and some young people in custody, by Juvenile Justice Officers who receive training on its use and are required to develop a case management plan based on the risk/need findings. Departmental policy requires repeated completion of the inventory at regular intervals. The data-set included scores for all items of all inventories completed up to early December 2005. This included 6890 sets of YLS scores. The data were screened to remove incomplete, duplicated and mistaken entries and restructured to link repeat assessments of the same individual. The result was 6632 sets of YLS scores relating to 4238 juvenile offenders from 40 juvenile justice locations throughout the state. We were able to obtain recidivism data for 4138 of these individuals from the NSW Bureau of Crime Statistics and Research reoffending database (ROD[1]). We eliminated four cases with data errors, 296 cases from 2002 when the inventory was being rolled out, and 270 cases where the inventory was used to assess youth serving custodial sentences. The final sample on which the results are based therefore incorporates all first YLS inventory results for 3568 young people that were completed between 1 January, 2003 and 2 December, 2005. The sample was 15.7% female and 84.3% male. Breakdown by ethnic origin was Australian not Aboriginal (44.7%), Aboriginal (29.7%), other ethnic origin (21.8%) and unspecified (3.8%). Age was distributed as 16.8% under 15 years, 41.3% aged 15 and 16 years, 39.9% aged 17 and 18 years, and 2.1% aged over 19 years.

Dependent variable

Choice of follow-up period is a matter of balancing practical and empirical considerations. From an empirical viewpoint, it is well established that longer follow-up periods will improve the predictive validity of risk instruments (Andrews at al., 2011; Yang, Wong, & Coid, 2010). However, for juvenile justice agencies, shorter periods may be more appropriate due to the age of the offender and sentencing time frames within which they operate. For instance, in 2010 the mean length of supervision orders made by Children's Court judges in the jurisdiction of the current study was 12 months (personal communication, C. Jones, Deputy Director, NSW Bureau of Crime Statistics and Research, September 2, 2011). Recidivism in the current study was thus defined as a re-offence occurring within one year of the administration of the YLS/CMI-AA that resulted in a conviction. Where the young person was charged and convicted of more than one offence, the offence attracting the most severe penalty was selected as the index offence. The analyses were based on the date of the actual index offence rather than the date of conviction, which varied depending on the police and court processing times. This calculation has the advantage of representing "real time" to offending. It reduced extraneous sources of time variability in our examination of the predictive reach of the inventory. The reoffending data retrieved from BOCSAR was for NSW court finalization dates up to the third quarter of 2007. Thus, there was ample time (approximately 11 months) for a conviction to be finalized and registered beyond the 1 year "real time" follow-up period of interest. Within 3 months of the first YLS, 20.9% of the sample had re-offended. Within 6 months, 33.7% had re-offended and within 12 months, 50.7% of the sample had re-offended.

Analyses

The data were analyzed using SPSS version 17. Analyses were conducted in three steps. Descriptive statistics were first calculated for the domain and total scores of the YLS/CMI-AA. The predictive validity of the total score was assessed using received operating characteristic (ROC) curve analysis (Rice & Harris, 1995). Second, logistic regression analyses were used to test the theoretical relevance of static and dynamic domains in predicting recidivism. Specifically, we investigated three models. Model 1 (static) tested the first domain of the YLS/CMI-AA as a predictor of recidivism. This domain represents age of first court order and other aspects of offence history. Hence, the first model is a test of a composite static predictor score. Model 2 (dynamic) tested the extent to which a weighted, linear combination of the seven dynamic YLS/CMI-AA domains can predict recidivism. Model 3 (static and dynamic) explored the linear combination of all eight domains of the inventory as a predictor of recidivism. To determine the relative predictive contribution of static and dynamic domains, individual increment tests were performed between Model 1 and Model 3, and between Model 2 and Model 3. This procedure involves comparing the likelihood ratio statistic for the full and reduced models. The difference between these two statistics has been shown to have an approximate chi-square distribution in larger samples. The degrees of freedom are equal to the difference in the number of predictor variables in the models (Kleinbaum & Klein, 2002). A significant chi-square statistic indicates that the full model accounts for significantly more variance than the reduced model. We also report the Nagelkerke R square statistic which can be interpreted as the proportion of variance explained by a logistic regression model (Nagelkerke, 1991). The third stage of our analyses focused on the predictive validity of additive scores from YLS/CMI-AA domains. The logistic regression analyses described examine optimal prediction when domain scores are weighted by coefficients in multinomial equations. In practice, static and dynamic domains on the inventory are weighted according to the number of items that comprise the domains. Thus, we analyzed the predictive validity of additive raw scores for the static domain, the combined seven dynamic domains, and all domains (total YLS/CMI-AA score). These results were compared to the findings of the logistic regression models.

Results

Descriptive statistics

The distributions of domain scores were all positively skewed except for *peer relations* and *leisure/recreation* which, having the smallest range of possible scores, were platykurtic. Means and standard deviations for the subscales can be seen in Table 1. The total scores on the YLS/CMI-AA were also positively skewed ranging from 0 (the minimum possible) to 48 (the maximum). The mean total score was 17.75 (SD = 9.65).

Table 1 about here

All subscales were significantly inter-correlated with most between .24 and .49 (see Table 2). The *family and living circumstances* domain had medium to large correlations (.40 to .53) with other dynamic domains. The highest correlation was between the *personality* and *attitudes* subscales (r = .65).

Table 2 about here

The total YLS/CMI-AA score was positively related to a re-offence within one year that resulted in a subsequent conviction (r = .26, p < 0.001). The AUC observed was .652, 95% CI [.634, .670] which indicates a 65% probability that a randomly selected recidivist would score higher on the total instrument than a randomly selected non-recidivist.

Logistic regression analyses

Linearity and collinearity assumptions for these analyses were tested using procedures recommended by Field (2010). For only one variable, *substance abuse,* was there evidence of a non-linear relationship with the logit of recidivism at a conservative probability level (p < .01). Collinearity diagnostics were acceptable. A summary of the logistic regression analyses are provided in Table 3. The results of Model 1 show that *prior and current offences* was a significant predictor of recidivism (OR = 1.30, p < 0.001). In Model 2, four of the dynamic domain scores were significantly related to recidivism: *education/employment*, *peer relations*, *substance abuse* and *attitudes/beliefs.* The highest OR was for peer relations OR= 1.19, p <.001. The remaining four dynamic domains were unrelated to recidivism in this multivariate model. For Model 3, using all the domain scores, five were significantly related to recidivism (*prior and current offences*, *education/employment*, *peer relations*, *substance abuse, attitudes/beliefs*). The subscale *prior and current offences* was the best predictor of recidivism when included in this multivariate model but its relationship with recidivism attenuated compared to Model 1 (*OR* = 1.18, $p < .001$). Odds ratios for the significant dynamic domains ranged from 1.14 ($p < .001$) for *peer relations* to 1.04 ($p = .043$) for *substance abuse*.

The increment tests showed that both static Model 1, $X^2(7) = 113.51$, p < .001, and dynamic Model 2, $X^2(1) = 57.87$, p < .001, accounted for significantly less variance than combined Model 3. The improvement using a combination of static and dynamic domains is reflected in the variance accounted for in Model 3 (Nagelkerke R square = .103) compared to Model 1 (Nagelkerke R square = .065) and Model 2 (Nagelkerke R square = .085). An alternative index of relative model utility is to consider the percentage accuracy in classification (PAC) for the binary criterion (recidivist versus non-recidivist). Using no model and only the base rates of recidivism, the PAC is 50.7%. For Model 1, the PAC improves to 59.8%. For Model 2, it is 60.2%, and for Model 3 it is 61.9%.

Insert Table 3 about here

Based on the results from preceding logistic analyses, a fourth model for predicting recidivism was run using only the five significant domain scores from Model 3. Model 4 was significant and showed that a degree of predictive utility equivalent to using all YLS/CMI-AA domain scores could be obtained from one static and four dynamic domains. Specifically, the non-significant chi-square statistic obtained from the increment test between Models 3 and 4 demonstrated the two models accounted for equivalent amounts of variance, $X^2(3) = 2.05$, $p = .56$. For Model 4, the R^2 estimate was.102 and PAC was 62.6%. All odds ratios and associated *p* values can be seen in Table 3[2].

Predictive validity of additive scores

The point-biserial correlations of the additive scores from the YLS with recidivism at one year showed the static score (*prior and current offences*) had a point-biserial correlation with recidivism of $r = .22$, $p < .01$. For the additive sum of the seven dynamic domains, the correlation with recidivism was $r = .24$, $p < .01$. Finally, the additive sum of all domains (YLS total score) was correlated with recidivism $r = .26$, $p < .01$. The correlation coefficients for the static and dynamic scores were statistically equivalent ($t = 1.38$, $p = .083$). However, the static score and recidivism correlation was significantly lower than the total score and recidivism correlation ($t = 2.72$, $p = .003$). Likewise, the dynamic score and recidivism correlation was significantly lower than the total score and recidivism correlation ($t = 6.15$, $p < .001$). These findings are consistent with the results of the logistic regressions, where Models 1 (static) and Model 2 (dynamic) were both found to account for significantly less variance than Model 3 (static and dynamic). As the additive score correlation coefficients are not directly comparable to the logistic regression analyses, for comparative purposes, the YLS additive total score was regressed against one-year recidivism. This resulted in a Nagelkerke R-square of .090 ($OR = 1.03$, $p < .001$, X^2 (1) = 249.13, $p < .001$). In comparison, the amount of variance accounted for by Model 3 was a Nagelkerke R square of .103 (see Table 1). The weighting used in the multivariate logistic regression therefore appears to result in marginally better predictive validity.

Discussion

The aim of this study was to examine the relative predictive validity of the static and dynamic domain scores of the YLS/CMI-AA. Both static and dynamic risks for juvenile offending are well supported in the theoretical and intervention literatures. Empirical support is also well established for such variables, although their relative contribution to both the emergence and continuation of offending is less clear. In structured risk-need inventories, scores for static and dynamic risks are known to be related to future offending. However, previous YLS studies showed that the static and only some dynamic domain scores predict recidivism. Moreover, the relative predictive contribution of domain scores has not been incisively pursued. In essence, the current study showed that individually, the static domain and a linear combination of the 7 dynamic domains both predicted recidivism. A linear combination of all static and dynamic domains resulted in a significant improvement in prediction. The same degree of prediction, though, could be obtained from a linear combination of the static and four of the dynamic domains. The simple additive scores from the YLS showed a similar pattern of association with recidivism. Specifically, the static domain score and the sum of the dynamic domain scores were both correlated with recidivism. However, the overall total YLS score yielded a higher degree of association with recidivism. These findings are discussed in more detail next.

Analysis showed that the first YLS domain (*prior and current offences)* was, on its own, a significant predictor of recidivism and accounted for an estimated 6.5% of the variance. Logistic regression analyses of the seven YLS dynamic domain scores accounted for an estimated 8.5 % of the variance and showed that only four contributed significantly to predicting recidivism. These domains were *education/employment*, *peer relations*, *substance abuse* and *attitudes/beliefs. Substance abuse* was the only dynamic score that significantly predicted increased odds of re-arrest in the logistic regression analysis of YLS domains by Flores, Travis, and Latessa (2003). In the multivariate models of current study, three dynamic domains (*family and living circumstances, leisure/recreation, personality/behavior)* were redundant for improving prediction. To a degree, this was reflected in the bivariate domain intercorrelations. *Family and living circumstances* had medium to large correlations with all other dynamic domains, and *personality/behavior* was highly correlated with *attitudes/beliefs*. The refined model based on the best 5 predictor domains was equivalent in variance accounted for as the model incorporating all YLS domains (Model 4 versus Model 3). However, risk prediction is only one benefit of a structured risk assessment inventory. Dynamic domains are also essential for needs assessment, case planning, and intervention. The logistic regression results show that even though some dynamic domains are redundant for prediction, they can serve other purposes without necessarily degrading prediction. This argument would be less tenable if the redundant dynamic domains were not related to recidivism at the bivariate level. However, all three redundant domains were correlated with reoffending in this study and hence satisfy one of the essential criteria for dynamic risk factors (Hanson, 2009).

It is important to be clear that the logistic regression results show how the domains of the YLS/CMI-AA could be weighted to predict recidivism based on the data from our large sample. That calculation is different from using simple additive scores from the domains which is how they are weighted in practice. Using the actual domain scores, we were able to show that the static and aggregated dynamic components predict recidivism, and that the total score shows a degree of improvement over each. Although a significant improvement in prediction could be gained from a more complex weighting of domains scores, the actual improvement (1% of variance) is small. Also, the point has been made previously that complex, computer generated risk estimates are not transparent to practitioners (Baker, et al., 2003). Transparency promotes the faithful application and systemic integrity of risk-need assessment which is the defining feature of fourth generation offender assessment (Andrews, et al., 2006). Within the risk-need-responsivity framework, the total YLS score is used to classify youth into several categories of risk (Hoge & Andrews, 2006, 2011). Our research findings support the value of the YLS total score for doing that although appropriate cut scores need to be determined and justified (McGrath & Thompson, 2011).

The amount of variance in recidivism accounted for by the predictive models in the current study was modest. It was at best 10% with a related AUC for the YLS/CMI-AA total score of .652. These indices are consistent with results from recent meta-analyses. For example, Schwalbe (2007) found a mean weighted AUC of .641 from 11 YLS predictive validity studies and Olver, Stockdale, and Wormith (2009) found a mean weighted correlation of .32 from 19 YLS studies. A variety of factors can limit predictive accuracy including length of the follow-up period, base rate of reoffending, sample characteristics especially related to gender and ethnic composition, integrity of YLS scoring, impact of remedial interventions, official versus non-official recidivism, and jurisdictional changes of youth residence. A recent meta-analysis in fact found that the majority of variance in the predictive validity of 9 risk assessment tools could be explained by methodological differences of this type (Yang, Wong, & Coid, 2010). In addition, Andrews and colleagues found the predictive validity of the YLS family of instruments to be superior for studies conducted in Canada (Andrews, et al., 2011). It is also worth pointing out that predictive accuracy is undermined by outcomes that statistically would be considered outliers. For example, the fit statistics in Model 4 were poor with two outliers flagged. This was for two youth who had YLS total scores of 45 and 47 but who did not re-offend in the follow-up period. Removal of these cases improved the fit statistics. However, these cases were bona-fide and not removed from the data. It is not uncommon in the field of juvenile justice to observe such outcomes. Indeed, it is the aim of juvenile justice interventions to help high risk youth to desist from offending. At the same time, it is true that some low risk youth will unexpectedly reoffend. These cases will attenuate predictive validity. Nevertheless, the current findings demonstrate that the YLS/CMI-AA is a useful tool for predicting risk and informing case management.

The results of this study support the predictive validity of the YLS/CMI-AA and in particular clarify the relative contribution of the static and dynamic domains and their related scores. It is important that YLS inventories be validated with youth samples in the jurisdictions that they serve (Hoge & Andrews, 2006, 2011). At the same time, the results reported here contribute to a substantial empirical literature on YLS inventories that details their predictive credibility (Olver, Stockdale, & Wormith, 2009; Schwalbe, 2007). The methodological strengths of this study, such as the large sample size and "real time" measure of reoffending, reinforce this conclusion. Our focus on domain scores has been in keeping with the need for attention to scale components to provide what Olver et al. referred to as "a more nuanced examination" (p. 349) of risk assessment instruments. Future directions for our research involve replicating this study with more recent data. The data presented here was from the first three full years of using the YLS/CMI-AA. Use of this particular version continues to date and has become a prominent aspect of agency culture. There are many benefits that accrue from a jurisdictional program of ongoing risk assessment research. Moreover, there is evidence that agency level adherence to risk-need-responsively principles can positively influence a variety of juvenile justice outcomes (Andrews, Bonta, & Wormith, 2006).

Table 1: Means and standard deviations for YLS subscales and total score

Subscale	N	Mean	SD	Min.	Max.
Prior and current offences	3568	3.57	1.77	0	9
Family and living	3568	2.35	1.91	0	7
Education/employment	3568	2.58	2.15	0	7
Peer relations	3568	2.13	1.34	0	4
Substance abuse	3568	2.32	1.98	0	6
Leisure/recreation	3568	1.52	1.18	0	3
Personality/behavior	3568	2.13	2.08	0	7
Attitudes/beliefs	3568	1.14	1.45	0	5
Total score	3568	17.75	9.65	0	48

Table 2: Correlations between YLS domain and total scores and with recidivism

Subscale	1	2	3	4	5	6	7	8	9
Prior and current offences (1)	(.22)								
Family and living (2)	.39	(.19)							
Education/employment (3)	.34	.48	(.19)						
Peer relations (4)	.41	.50	.44	(.22)					
Substance abuse (5)	.24	.40	.24	.41	(.14)				
Leisure/recreation (6)	.27	.46	.44	.47	.31	(.16)			
Personality/behavior (7)	.36	.53	.53	.41	.29	.36	(.16)		
Attitudes/beliefs (8)	.35	.53	.49	.46	.30	.44	.65	(.18)	
Total score (9)	.60	.78	.73	.71	.59	.63	.76	.75	(.26)

***Note*: correlation with recidivism at one year on diagonal**
All correlations significant (*p* < .01)

Table 3: Logistic regression models for predicting recidivism at 12 months

	Model 1 (static)					Model 2 (dynamic)					Model 3 (static +dynamic)					Model 4 (refined)				
	B	SE	Wald	OR (95% CI)	p	B	SE	Wald	OR (95% CI)	p	B	SE	Wald	OR (95% CI)	p	B	SE	Wald	OR (95% CI)	p
Prior and current offences	0.26	.02	166.13	1.30 [1.25, 1.37]	.001						0.16	0.02	50.47	1.18 [1.13, 1.23]	.001	0.17	0.02	53.42	1.18 [1.13, 1.23]	.001
Family and living						0.05	0.02	3.67	1.05 [0.99, 1.10]	.055	0.03	0.03	1.02	1.03 [0.98, 1.08]	.312					
Education/ employment						0.07	0.02	11.16	1.07 [1.03, 1.12]	.001	0.06	0.02	7.94	1.06 [1.02, 1.11]	.005	0.07	0.02	11.13	1.07 [1.03, 1.11]	.001
Peer relations						0.18	0.03	28.50	1.19 [1.12, 1.27]	.001	0.13	0.03	15.05	1.14 [1.07, 1.22]	.001	0.14	0.03	19.47	1.15 [1.08, 1.23]	.001
Substance abuse						0.04	0.02	4.70	1.04 [1.00, 1.09]	.030	0.04	0.02	4.11	1.04 [1.01, 1.08]	.043	0.05	0.02	5.63	1.05 [1.01, 1.09]	.018
Leisure/ recreation						0.03	0.04	0.58	1.03 [0.96, 1.10]	.448	0.03	0.04	0.71	1.03 [0.96, 1.11]	.399					
Personality/ behavior						0.01	0.02	0.06	1.00 [0.96, 1.05]	.801	0.01	0.02	0.10	0.99 [0.95, 1.04]	.756					
Attitudes/beliefs						0.08	0.06	5.45	1.08 [1.01, 1.16]	.020	0.07	0.04	4.01	1.07 [1.00, 1.15]	.045	0.08	0.03	6.87	1.08 [1.02, 1.15]	.009
Model X^2				179.60					235.25					286.87					284.82	
Nagelkerke R-square				.065					.085					.103					.102	

***Note*: All model X^2 were significant ($p < .001$), Model 1 *df* = 1, Model 2 *df* = 7, Model 3 *df* = 8, Model 4 *df* = 7**

References

American Educational Research Association (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2004). *Level of Supervision/Case Management Inventory (LSI/CMI): An offender assessment system. User's guide.* Toronto: Multi-Health Systems.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency, 52*, 7-27.

Andrews, D. A., & Dowden, C. (2006). Risk principle of case classification in correctional treatment - A meta-analytic investigation. *International Journal of Offender Therapy and Comparative Criminology, 50*, 88-100.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 28*, 369-404.

Andrews, D. A., Bonta, J., Wormith, J. S., Guzzo, L., Brews, A., Rettinger, J., et al. (2011). Sources of variability in estimates of predictive validity. *Criminal Justice and Behavior, 38*, 413-432. doi: 10.1177/0093854811401990

Baker, K., Jones, S., Roberts, C., & Merrington, S. (2003). *ASSET-The evaluation of the validity and reliability of the Youth Justice Board's assessment for young offenders. Findings from the first two years of the use of ASSET*. Oxford: Centre for Criminological Research.

Benda, B. B., Corwyn, R. F., & Toombs, N. J. (2001). Recidivism among adolescent serious offenders. *Criminal Justice and Behavior, 28*, 588-613.

Borduin, C. M., Mann, B., Cone, L., Henggeler, S., Fucci, B., Blaske, D., et al. (1995). Multisystemic treatment of serious juvenile offenders: Long-term prevention of criminality and violence. *Journal of Consulting & Clinical Psychology, 63*, 569-578.

Clements, C. B. (1996). Offender classification: Two decades of progress. *Criminal Justice & Behavior, 23*, 121-143.

Cottle, C., C., Lee, R., J., & Heilbrun, K. (2001). The prediction of criminal recidivism in juveniles: A meta-analysis. *Criminal Justice and Behavior, 28*, 367-394.

Day, A., Howells, K., & Rickwood, D. (2004). *Current trends in the rehabilitation of juvenile* offenders. Canberra: Australian Institute of Criminology, 2004.

Durlak, J. A. (1998). Common risk and protective factors in successful prevention programs. *American Journal of Orthopsychiatry, 68*, 512-520.

Farrington, D. P. (2002). Key results from the first forty years of the Cambridge study in delinquent development. In T. P. Thornberry & M. D. Krohn (Eds.), *Taking stock of delinquency: An overview of findings from contemprary longitudinal studies* (pp. 137-183). New York: Kluwer/Plenum.

Flores, A. W., Travis, L. F., & Latessa, E. J. (2003). *Case classification for juvenile corrections: An assessment of the Youth Level of Service Case Management Inventory (YLS/CMI)*. Cincinnati, OH: Centre for Criminal Justice Research, University of Cincinnati.

Gavazzi, S. M., Slade, D., Buettner, C. K., Patridge, C., Yarcheck, C. M., & Andrews, D. W. (2003). Toward conceptual development and empirical measurement of global risk indicators in the lives of court-involved youth. *Psychological Reports, 92*, 599-615.

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology, 34*, 575-607.

Hanson, R. K. (2009). The psychological assessment of risk for crime and violence. *Canadian Psychology, 50*, 172-182.

Hoge, R. D., & Andrews, D. A. (1995). *Australian adaptation of the Youth Level of Service/Case Management Inventory*. North Tonawanda: NY: Multi-Health Systems.

Hoge, R. D., & Andrews, D. A. (1996). *Assessing the youthful offender: Issues and techniques*. New York: Plenum Press.

Hoge, R. D., & Andrews, D. A. (2006). *Youth Level of Service/Case Management Inventory. User's Manual*. North Tonawanda, NY: Multi-Health Systems.

Hoge, R. D., & Andrews, D. A. (2011). *YLS/CMI 2.0. Youth Level of Service/Case Management Inventory 2.0*. North Tonawanda, NY: Multi-Health Systems.

Hollin, C. R. (2002). Risk-needs assessment and allocation to offender programmes. In J. McGuire (Ed.), *Offender rehabilitation and treatment: Effective programmes and policies to reduce reoffending* (pp. 309-332). Etobicoke, Canada: John Wiley.

Holsinger, A. M., Lowenkamp, C. T., & Latessa, E. J. (2006). Predicting institutional misconduct using the youth level of service/case management inventory. *American Journal of Criminal Justice, 30*, 267-284.

Howell, J. C. (1995). *Guide for implementing the comprehensive strategy for serious, violent, and chronic juvenile offenders*. Washington, DC: Office of Juvenile Justice and Deliquency Prevention.

Hua, J., & Fitzgerald, J. (2006). *Matching court records to measure reoffending*. Sydney: NSW Bureau of Crime Statistics and Research.

Hubbard, D. J., & Pratt, T. C. (2002). A meta-analysis of the predictors of delinquency among girls. *Journal of Offender Rehabilitation, 34*(3), 1-13.

Kleinbaum, D. G., & Klein, M. (2002). *Logistic regression: A self-learning text*. New York: Springer.

Lally, S. J. (2003). What tests are acceptable for use in forensic evaluations? A survey of experts. *Professional Psychology: Research & Practice, 34*, 491-498.

Loeber, R., & Farrington, D. P. (Eds.). (1998). *Serious & violent juvenile offenders: Risk factors and successful interventions*. Thousand Oaks: Sage Publications.

Marczyk, G. R., Heilbrun, K., Lander, T., & DeMatteo, D. (2003). Predicting juvenile recidivism with the PCL:YV, MAYSI, and YLS/CMI. *International Journal of Forensic Mental Health, 2*, 7-18.

Marczyk, G. R., Heilbrun, K., Lander, T., & Dematteo, D. (2005). Juvenile decertification: Developing a model for classification and prediction. *Criminal Justice and Behavior, 32*, 278-301.

Marshall, J., Egan, V., English, M., & Jones, R. M. (2006). The relative validity of psychopathy versus risk/needs-based assessments in the prediction of adolescent offending behavior. *Legal and Criminological Psychology, 11*, 197-210.

McGrath, A. (2009). Offenders' perceptions of the sentencing process—A study of deterrence and stigmatisation in the NSW Children's Court. *Australian & New Zealand Journal of Criminology, 42*, 24-46.

McGrath, A., & Thompson, A. P. (2011). *Actuarial risk assessment for juvenile offenders : Lessons from an Australian normative study*. Paper presented at the 32nd International Congress of Law and Mental Health, Berlin.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*, 674-701.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78*, 691-692.

Nagin, D., & Paternoster, R. (2000). Population heterogeneity and state dependence: State of the evidence and directions for future research. *Journal of Quantitative Criminology, 16*, 117-144.

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior, 36*, 329-353.

Onifade, E., Davidson, W., Campbell, C., Turke, G., Malinowski, J., & Turner, K. (2008). Predicting recidivism in probationers with the Youth Level of Service Case Management Inventory (YLS/CMI). *Criminal Justice and Behavior, 35*, 474-483.

Onifade, E., Davidson, W., Livsey, S., Turke, G., Horton, C., Malinowski, J., et al. (2008). Risk assessment: Identifying patterns of risk in young offenders with the Youth Level of Service/Case Management Inventory. *Journal of Criminal Justice, 36*, 165-173.

Paternoster, R., & Brame, R. (1997). Multiple routes to delinquency? A test of developmental and general theories of crime. *Criminology, 35*, 49-84.

Piquero, A. R., Farrington, D. P., & Blumstein, A. (2003). The Criminal Career Paradigm. In M. Tonry (Ed.), *Crime and justice: An annual review of research* (Vol. 30, pp. 359-506). Chicago: University of Chicago.

Rice, M. E., & Harris, G. T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology, 63*, 737-748.

Sampson, R. J., & Laub, J. H. (2003). Life-course desisters? Trajectories of crime among delinquent boys followed to age 70. *Criminology, 41*, 555-592.

Schwalbe, C. S. (2007). Risk assessment for juvenile justice: A meta-analysis. *Law & Human Behavior, 31*, 449-462.

Schwalbe, C. S., Fraser, M. W., Day, S. H., & Arnold, E. M. (2004). North Carolina Assessment of Risk (NCAR): Reliability and predictive validity with juvenile offenders. *Journal of Offender Rehabilitation, 40*, 1-22.

Smith, D. A., & Paternoster, R. (1990). Formal processing and future delinquency: Deviance amplification as a selection artefact. *Law and Society Review, 24*, 1109-1131.

Thompson, A. P. (2001). Juvenile offending : An overview in support of assessing risk factors, needs, and strengths. *Australian Educational and Developmental Psychologist, 18*, 39-56.

Thompson, A. P., & 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*, 207-214.

Thompson, A. P., & Putnins, A. L. (2003). Risk-need assessment inventories for juvenile offenders in Australia. *Psychiatry, Psychology and Law, 10*, 324-333.

Upperton, R. A., & Thompson, A. P. (2007). Predicting juvenile offenders recidivism: Risk-need assessment and juvenile justice officers. *Psychiatry, Psychology and Law, 14*, 138-146.

Weatherburn, D. J., Lind, B., & Hua, J. (2003). *Contact with the New South Wales court and prison systems : The influence of age, Indigenous status and gender*. Sydney: NSW Bureau of Crime Statistics and Research.

Welsh, J. L., Schmidt, F., McKinnon, L., Chattha, H. K., & Meyers, J. R. (2008). A comparative study of adolescent risk assessment instruments: Predictive and incremental validity. *Assessment, 15*, 104-115.

Yang, M., Wong, S. C., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin, 136*, 740-767.

Author bios

Andrew McGrath, PhD

Andrew McGrath lectures in forensic and developmental psychology at Charles Sturt University in Bathurst, Australia. He received his PhD from the University of Sydney in 2007. His research interests are in juvenile justice, criminological theory, risk assessment and bail decisions.

Anthony P. Thompson, PhD
Tony Thompson has worked with both juvenile and adult offenders. As an Associate Professor, his teaching and research interests have included a range of related topics.

Endnotes

[1] ROD contains all finalised juvenile and adult criminal appearances in NSW higher courts (District and Supreme), NSW Local Courts and the NSW Children's Court since 1994. Further information regarding the matching procedures used and their accuracy can be found in Hua and Fitzgerald (2006) and in Weatherburn, Lind, and Hua (2003).

[2] These data were also analysed using Cox's regression, which, because it uses time to reoffend as the dependent variable, maximises follow-up times. See McGrath (2009) for a discussion of this analytic technique. The results of these analyses were essentially the same as the logistic regression analyses with recidivism at 12 months as the dependent variable. In addition, the same logistic regression models were run using re-offence within 6 months of YLS administration as the dependent variable. A similar final model was obtained, with the addition of subscale 6 (leisure and recreation), which was significantly associated with recidivism (OR = 1.09, p = 0.026).