Situational and generalised conduct problems and later life outcomes: Evidence from a New Zealand birth cohort

David M. Fergusson, Ph.D.

Joseph M. Boden, Ph.D.

L. John Horwood, MSc.

Christchurch Health and Development Study

University of Otago, Christchurch School of Medicine and Health Sciences

Corresponding author: Prof. David M. Fergusson, Christchurch Health and Development Study, University of Otago, Christchurch School of Medicine and Health Sciences, PO Box 4345, Christchurch, New Zealand

Phone: +64 3 372 0406 Fax: +64 3 372 0407 Email: dm.fergusson@otago.ac.nz

The Authors declare no conflicts of interest.

Word Count: 6099

Number of Tables: 3

Abstract

Background: There is considerable evidence suggesting that many children show conduct problems that are specific to a given context (home; school). What is less well-understood is the extent to which children with situation-specific conduct problems show similar outcomes to those with generalised conduct problems.

Methods: Data were gathered as part of the Christchurch Health and Development Study, a 25-year longitudinal study of a birth cohort of New Zealand children. Information was obtained on: (a) conduct problems during the period 7-9 years (b) criminal offending during the period 16-25 years; (c) measures of DSM-IV mental disorders and suicidal behaviour over the interval 16-25 years; (d) measures of DSM-IV substance dependence over the interval 16-25 years; and (e) measures of relationship, pregnancy, and parenthood outcomes during the period 16-25 years.

Results: Latent-class modelling suggested three distinct groups of children with conduct problems: those with mother reports; those with teacher reports; and those with both mother and teacher reports. Both situation-specific and generalised conduct disorder were associated with increased risk of criminal offending, mental health disorders, substance dependence, and relationship and parenthood issues in late adolescence and early adulthood.

Conclusions: There is a need for recognition of the significance of situation-specific conduct problems in both developmental theory and in the treatment of childhood conduct disorders. A focus only on those children with generalised conduct problems is likely to overlook the features and needs of children whose conduct problems are confined to a specific context.

In recent decades there has been significant interest in the diagnosis and treatment of conduct disorders in children and young people (Scott, 2007, 2008). An important issue arising from research in this area is that of developing criteria to identify children and young people who are likely to experience adverse outcomes in later life because of conduct problems (Deligatti, Akin-Little, & Little, 2003; Kim-Cohen et al., 2005). At first sight this would seem to be a relatively straightforward task by applying standard diagnostic procedures. However, a complication that arises in the application of these diagnoses concerns the way in which situation-specific disorders should be managed. In particular, there is considerable evidence to suggest that many childhood conduct problems are context-specific, with the result that the child's behaviours are described as problematic in one context (e.g. home) but not another (e.g. school) (Arseneault et al., 2003; Fergusson & Horwood, 1989a; Lahey, Loeber, Quay, Frick, & Grimm, 1997; Matthys, Maassen, Cuperus, & van Engeland, 2001). The presence of these context specific behaviours appears to account for the well known findings that parent and teacher's reports of childhood conduct problems are only moderately correlated (Achenbach, McConaughy, & Howell, 1987; McCombs Thomas, Forehand, Armistead, Wierson, & Fauber, 1990; Spiker, Kraemer, Constantine, & Bryant, 1992; Touliatos & Lindholm, 1981). While this lack of across-source correlation has been attributed to a lack of reliability or validity of parent and teacher report data (McCombs Thomas et al., 1990; Spiker et al., 1992), structural equation modelling methods (Fergusson & Horwood, 1989a) suggest that this is unlikely to be the case. These analyses suggest the presence of reliable test variance in parent and teacher reports that appears to describe genuine between context differences in child behaviours (Fergusson & Horwood, 1989a).

These issues also emerge in the context of methods for identifying children with significant conduct problems. Specifically, a number of authors have advocated the use of multiple-gating methods in which only those children who show conduct problems in multiple contexts are eligible for treatment (LeBlanc, 1998; Loeber, Dishion, & Patterson, 1984; Severson, Walker, Hope-

Doolittle, Kratchowill, & Gresham, 2007; Walker & Severson, 1990). These methods may eliminate children whose conduct problems are situation-specific from access to treatment.

For all of these reasons, it is important to examine the similarities and differences in the longer-term prognosis of children with situation-specific and generalised conduct problems. One way of approaching this issue would be to study untreated groups of children with situational specific and generalised conduct problems into later life to determine the outcome risks for each group. It turns out that this experiment can be well approximated by studying existing cohorts of children born in the 1970s, who would have virtually no exposure to the array of treatments and interventions available to more recently born cohorts.

In this paper we use data from one such natural experiment to examine the long-term prognosis of children who showed situation specific and generalised conduct disorder using data from a 25-year longitudinal study of New Zealand born children. The aims of this study were two-fold:

1) To identify groups of children who showed stable patterns of situation-specific and generalised conduct problems in middle childhood (ages 7-9 years) using latent-class modelling techniques.

2) To relate the pattern of childhood conduct problems to long term outcomes observed up to the age of 25. These outcomes span criminal offending; mental health problems; substance dependence; and relationship and pregnancy/parenthood issues.

Method

The data reported here were gathered during the course of the Christchurch Health and Development Study (CHDS). The CHDS is a longitudinal study of an unselected birth cohort of 1,265 children born in the Christchurch (New Zealand) urban region during a 4-month period in mid-1977. This cohort has been studied at birth, 4 months, 1 year, annual intervals to age 16 years, and at ages 18, 21 and 25 years. Data have been gathered from a combination of sources including:

parental interviews; self-reports; psychometric tests; teacher reports; medical and other official records. All data were obtained with the informed consent of the participants.

Childhood conduct problems (7-9 years)

At ages 7, 8 and 9 years maternal and teacher reports of the child's tendencies to disruptive, oppositional and conduct disordered behaviours were obtained using items from the Rutter (Rutter, Tizard, & Whitmore, 1970) and Conners (1969; , 1970) parent and teacher questionnaires. In most cases, teacher reports at ages 7, 8, and 9 were obtained from different teachers. These items spanned a range of behaviours relating to disobedience and defiance of authority, fits of temper and irritability, aggression or cruelty towards others, destruction of property, lying, stealing and other similar behaviours. Confirmatory factor analysis of the selected items for each source (mothers, teachers) suggested that, in each case, the items could be scaled as unidimensional scales representing the extent of child conduct problems as reported by mothers and teachers at each age (Fergusson & Horwood, 1993; Fergusson, Horwood, & Lloyd, 1991). The reliabilities of these scales ranged from α = .83 to .85 for mother reports and α = .93 to .95 for teacher reports. The correlations between the mother reports and teacher reports at each age ranged from .27 to .36. In addition, the correlations between mother reports of conduct problems at ages 7, 8, and 9 ranged from .76 to .69, while for teacher reports ranged from .59 to .49.

Adverse outcomes in adulthood

Criminal offending

  Property/violent offending (16-25 years). This was assessed at ages 18, 21, and 25 using the Self Report Delinquency Inventory (SRDI, Elliott & Huizinga, 1989). These data were used to derive dichotomous measures of the percentage of individuals reporting: (a) ten or more property offences; and (b) ten or more violent offences; during the period 16-25 years.

Arrest/conviction (16-25 years). Individuals who reported being arrested or convicted of a criminal (non-traffic) offence during the interval 16-25 years were classified as having a having a history of official judicial contact.

Traffic offences (18-25 years). Sample members were classified as being repeated traffic offenders if they reported a total of 5 or more traffic offences of any type during the interval from 18-25 years.

Imprisonment (ever). Sample members were classified as having been imprisoned if they reported having ever been sentenced to prison at any time from age 17 onwards (the earliest any cohort member was sentenced to prison was age 17).

Overall number of offending problems. The five dichotomous measures described above were summed to create an overall count measure reflecting the number of criminal offending problems reported.

Mental health problems

Major depression/anxiety disorder. Assessed at ages 18, 21, and 25 via items from the Composite International Diagnostic Interview (CIDI, World Health Organization, 1993) for the assessment of DSM-IV (American Psychiatric Association, 1994) symptom criteria for major depression and anxiety disorders. Individuals who met DSM-IV diagnostic criteria for either a major depressive disorder or anxiety disorder during the period 16-25 years were classified as having a disorder.

Conduct/antisocial personality disorder (16-25 years). DSM-IV conduct disorder criteria were assessed at age 18 using the Self Report Delinquency Inventory (SRDI: Elliott & Huizinga, 1989). At ages 21 and 25 custom written survey items were used to assess DSM-IV diagnostic criteria for ASPD. Sample members were classified as having conduct disorder/ASPD if they met criteria for either disorder over the interval 16-25 years.

Suicide attempt (16-25 years). Cohort members who indicated at least one suicide attempt during the period 16-25 years were classified as having attempted suicide during that period.

Overall number of mental health disorders. The three dichotomous measures described above were summed to create an overall count measure reflecting the number of mental health disorders reported.

Substance dependence

Nicotine dependence (18, 21, and 25 years). Assessed via custom written survey items designed to assess DSM-IV symptom criteria for dependence. Sample members who, at ages 18, 21, or 25, reported smoking 5 or more cigarettes per day and who also reported three or more of the DSM-IV symptom criteria were classified as nicotine dependent.

Alcohol dependence (16-25 years). Assessed over the intervals from 16-18 years, 18-21 years, and 21-25 years using items from the CIDI for DSM-IV symptom criteria for alcohol dependence. Individuals who met the relevant DSM-IV diagnostic criteria for alcohol dependence at any time during any of the intervals were classified as alcohol dependent.

Illicit drug dependence (16-25 years). Assessed via relevant items from the CIDI for DSM-IV symptom criteria for substance dependence. Sample members who met diagnostic criteria for cannabis or other illicit drug dependence at any time during the any of the intervals were classified as having illicit drug dependence.

Overall number of substance use problems. The three dichotomous measures described above were summed to create an overall count measure reflecting the number of substance use problems reported.

Relationship and pregnancy/parenthood issues

Interpartner violence (24-25 years). Assessed at age 25 using items from the Revised Conflict Tactics Scale (Straus, Hamby, Boney-McCoy, & Sugarman, 1996) concerning episodes of

psychological, physical or sexual violence within the most recent relationship lasting at least one month. Those reporting perpetrating or being victimised by violent behaviour with a partner were classified as being exposed to interpartner violence.

Ten or more sexual partnerships (18-25 years). Information concerning the number of sexual partnerships reported by participants during the periods 18-21 years and 21-25 were used to create a dichotomous measure of those individuals reporting 10 or more sexual partnerships during the interval 18-25 years.

Early pregnancy/parenthood. Sample members were classified as having an *early pregnancy* if they reported having become pregnant (females) or getting a partner pregnant (males) prior to age 21, and were classified as *early parents* if they reported either giving birth (females) or fathering a biological child (males) before age 21. Higher rates of pregnancy than parenthood were reported because: (a) some pregnancies were aborted; (b) some pregnancies were miscarried; (c) some pregnancies did not come to term until after the assessment at age 21.

Overall measure of relationship and pregnancy/parenthood issues. The four dichotomous measures described above were used to create an overall count measure of relationship and pregnancy/parenthood issues by summing the responses to the dichotomous outcome measures.

Statistical analysis

The analysis was conducted in three stages. In the first stage, mother and teacher ratings of conduct problems at 7, 8, and 9 years were dichotomised by classifying children as having conduct problems if the child's score placed them in the most disturbed 10% of the sample. The 10% cut-point was chosen on the basis of previous studies suggesting that in the region of 10% of the general child population meet criteria for conduct or oppositional defiant disorders (Anderson, Williams, McGee, & Silva, 1987; Fergusson, Horwood, & Lynskey, 1995; McGee et al., 1990). This process yielded $2^6 = 64$ response patterns describing child behaviour assessed by two sources (mother, teacher) at three times (ages 7, 8, 9). These response patterns formed the input data for fitting a series of latent

class models to the data, to determine whether it was possible to identify distinct groups of children displaying situation-specific and generalised conduct disorders. This analysis was conducted using PANMARK (van de Pol, Langeheine, & de Jong, 1999) and methods of maximum likelihood estimation. The investigation was conducted by fitting two, three, and four class models to the observed response patterns and using log likelihood ratio chi square values and Bayesian Information Criteria (BIC) to identify the best-fitting and most parsimonious model. All models assumed that the reporting accuracy of mother reports was consistent over time, and similarly for teacher reports. As explained in Results, the best-fitting model proved to be a four class model that classified children into classes of children with: (a) no conduct problems; (b) children with mother-reported conduct problems; (c) children with teacher-reported conduct problems; and (d) children with generalised conduct problems. From the parameters of the fitted model, estimates were derived of the posterior probabilities of latent class membership conditional on the observed pattern of responses of mothers and teachers. Children were classified to latent classes according to the highest estimated posterior membership probability (Fergusson & Horwood, 1989b).

In the second stage, the child's conduct problem status at ages 7-9 was related to a series of outcomes (criminal offending; mental health; substance dependence; relationship and pregnancy/parenthood issues) observed from ages 16-25 to determine the extent to which adjustment in middle childhood was related to later outcomes. The associations between conduct problems latent classes and outcomes were modelled using Poisson regression models with design variates representing the conduct problems latent classes. The parameters of the fitted models were then used to estimate the Incidence Rate Ratio (IRR) and 95% confidence interval for each conduct problems latent class for each outcome.

Finally, all analyses were replicated by using a 5% cutting point on mother- and teacher-reported conduct problems.

Sample sizes and missing data

The present analyses were based on those cohort members for whom information was available on childhood conduct problems at ages 7, 8, and 9, and for whom information was available on outcomes during the period 16-25 years. Overall 1046 participants (83% of the original cohort) had data on child conduct problems at age 7-9 years and were included in the latent class analysis. This sample was reduced to 983 (78% of the cohort) for the analysis of outcome data to age 25.

To test for selection bias arising from the processes of sample attrition, the sample included in the analyses was compared with remaining cohort members on a series of measures collected at the time of birth. These comparisons suggested evidence of small but statistically significant ($p <$ .05) tendencies for individuals from socio-economically disadvantaged backgrounds (low parental education, low SES family, single parent family) to be under-represented in the analysis sample. To examine the extent to which the study findings may have been influenced by these small biases the analyses were repeated using the data weighting methods described by Carlin et al (Carlin, Wolfe, Coffey, & Patton, 1999). These analyses produced almost identical conclusions to the results reported here, suggesting that the findings were unlikely to have been influenced by selection bias.

Results

The classification of conduct problems

As explained in Methods, measures of conduct problems were obtained from mothers and teachers at ages 7, 8 and 9 years. For each age and source the conduct problems scale scores were dichotomised to classify children in the most disturbed 10% of the sample as having conduct problems, and the other cohort members as not having these problems.

Table 1a summarises the goodness of fit indices of 2, 3 and 4 class models fitted to the data. Fit indices include the log likelihood ratio chi square statistic and the Bayesian Information Criterion (BIC). The Table shows that, on the basis of both fit criteria, a 4 class model provided the best fit to the data. The estimated model parameters and standard errors for the fitted 4 class model

are shown in Table 1b. These parameters represent: (a) *latent class probabilities* – the estimated unconditional probabilities of membership of each latent class; and (b) *response probabilities* - the probabilities that a member of a given latent class will be described as having conduct problems in a given year by their mothers or teachers. The Table shows that the latent class model classified children into the following groups:

1. *No Problems.* This group comprised 83.2% of the cohort. Children in this group had low probabilities of being described as having conduct problems by their mother (2.0%) or teacher (3.5%).

2. *Mother Reported Problems.* This group comprised 5.8% of the cohort. Children in this group had a high probability (69%) of being described by their mother as having a conduct problem at any age but a low probability (3.3%) of being described by their teacher as having conduct problems.

3. *Teacher Reported Problems.* This group comprised 7.6% of the cohort. Children in this group had a high probability (55%) of being described by their teacher as having conduct problems at a given age but a low probability (11%) of being described as having conduct problems by their mother.

4. *Generalised Conduct Problems.* This group comprised 3.4% of the cohort. Children in this group had high probabilities of being described as having conduct problems at any given age on the basis of both mother reports (81%) and teacher reports (65%).

INSERT TABLES 1A, 1B HERE

Using the parameters of the fitted four-class model, children were classified into latent classes on the basis of the estimated posterior probabilities of latent class membership (see Methods). This method of assignment was estimated to result in a 97.4% rate of correct classification of participants to latent classes.

To illustrate the association between the latent class assignment and the original behaviour scores Table 2 shows the observed mean scores on the measures mother and teacher conduct problem at ages 7-9 years. For simplicity of presentation all measures have been scaled to a mean of 100 and standard deviation of 10. Examination of the table provides strong validation for the latent class model. Specifically, those assigned to the mother reported problems group had consistently high mean scores on mother reports, whereas the teacher scores were similar to those in the no problems group. Those assigned to the teacher reported problems group had consistently high means on teacher reports but mean scores for mother reports were only slightly elevated compared to the no problems group. Those in the generalised conduct problems group had consistently elevated mean scores on both mother and teacher reports.

INSERT TABLE 2 HERE

The long term outcomes of childhood conduct problems

Table 3 shows the associations between the latent class assignment and a series of outcomes measured up to the age of 25. These outcomes span the domains of: offending; mental health; substance use; and parenting and partnerships. The associations between the conduct disorder latent class groups and the overall number of problems measure in each outcome domain were modelled using Poisson regression, with latent classes represented by design variates. From the parameters of these models, estimates of the Incidence Rate Ratio (IRR) and 95% confidence interval (CI) were obtained for each conduct problems group relative to those with no problems, as well as pair-wise tests of significance between groups. The Table shows that:

1. On all outcome measures (offending; mental health; substance dependence; relationship and pregnancy/parenthood issues) children with conduct problems were at significantly ($p < .05$) increased risk of later adverse outcomes, with risk ratios ranging from 1.39 to 3.74.

2. The risk ratios for the three conduct problems groups were similar. This impression was supported by the results of pairwise comparisons, which showed in all case that the outcomes of the groups were not significantly different (p > .05).

INSERT TABLE 3 HERE

Analyses using 5% cut off score

The analyses reported above were repeated using conduct problem scale scores that were dichotomised to classify children in the most disturbed 5% of the sample as having conduct problems, and the other cohort members as not having these problems. The latent class analyses and analyses of the associations between the conduct problems latent classes and later adverse outcomes were conducted as described above, with the following results:

1. Latent class modelling showed that, on the basis of both fit criteria, a 4 class model provided an adequate fit to the data. Those with no conduct problems comprised 90.5% of the cohort; those with mother-reported conduct problems comprised 3.5% of the cohort; those with teacher-reported conduct problems comprised 4.5% of the cohort, and those with generalised conduct problems comprised 1.5% of the cohort.

2. Analyses of the associations between the conduct problems latent classes and later adverse outcomes showed that those with either situation-specific or generalised conduct problems had significantly (p < .05) higher numbers of adverse outcomes than individuals with no conduct problems. However, there were no significant (all p values > .05) differences found between any of the three conduct problems groups.

Discussion

This paper has used data gathered over the course of a 25-year longitudinal study to examine the linkages between conduct problems in middle childhood and later outcomes in young adulthood

including criminal offending, mental health, substance dependence and parenting/partnerships. The focus of this analysis was on identifying the frequency with which situational conduct problems were present and the extent to which situational problems were related to longer term outcomes. The key findings from this analysis are given below.

The prevalence of situational and generalised conduct problems

In the first stage of the analysis, a constrained latent class model was fitted to data on dichotomised mother and teacher report obtained at ages 7, 8, and 9 years. This model provides a categorical analogue of the method/trait model fitted to these data in an earlier analysis (Fergusson & Horwood, 1993). In general the latent class model was well-fitting and suggested four distinct groups of children: (a) those described as not having conduct problems (83.2%); (b) those described as having conduct problems by their mother but not their teachers (5.8%); (c) those described as having conduct problems by their teachers and not their mother (7.6%); and (d) those described as having conduct problems by both their mother and teachers (3.4%). This finding is consistent with the conclusion that the majority of childhood conduct problems are situation-specific and that only a minority of children have conduct problems that occur across contexts. Using a 10% cutting point suggested that while 17% of children had significant conduct problems reported by mother or teacher, only 3.4% had generalised conduct problems.

These findings reflect the well-known fact that mother and teacher reports of child behaviours are only modestly correlated, with the result that problems are frequently reported by one source but not another (Achenbach et al., 1987; McCombs Thomas et al., 1990; Spiker et al., 1992; Touliatos & Lindholm, 1981). In the past the poor correlation between parent and teacher reports has been interpreted as being due to a lack of reliability or validity in such reports (McCombs Thomas et al., 1990; Spiker et al., 1992). However, the alternative explanation is that parent and teacher reports of child behaviour are generally accurate, and that poor agreement

between reports reflects the fact that for many children conduct problems are situational and occur in one context but not others (Fergusson & Horwood, 1989a).

The long-term consequences of situational and generalised conduct problems

The second stage of the analysis examined the extent to which situational and generalised conduct problems were predictive of long term outcomes. This analysis led to three major conclusions.

In all cases children with situational or generalised conduct problems were at increased risks of later crime, mental health problems, substance dependence and relationship/parenthood issues than children with no significant problems. Further, in most comparisons children with generalised conduct problems had the highest rates of adverse outcomes but in all cases these rates were not significantly different from children having situational conduct problems. These conclusions are consistent with the view that children with situational conduct problems are an at risk group for a wide range of future adverse outcomes, although the risks for this group were somewhat lower than for those with generalised conduct problems.

Implications

The findings above have important implications for the identification of children who may benefit from intervention for childhood conduct problems. Currently the literature on this topic has emphasised the importance of treating children who show generalised conduct problems that are present at home and at school. This emphasis has been evident in so-called multiple gating systems that classify children as eligible for treatment if and only if there is an agreement about the child's status from multiple sources (LeBlanc, 1998; Loeber et al., 1984; Severson et al., 2007; Walker & Severson, 1990). In cases where gating involves both parent and teacher report these procedures will result in children who show situation-specific conduct disorders being disqualified from treatment. The present study leads to two major conclusions about the methods of classification which exclude children with situational disorders from treatment. First, since situational conduct

problems are the most common, these methods will eliminate the majority of children with significant conduct problems from being eligible for treatment. Second, those with situational conduct problems appear to at increased risks of long-term adverse outcome if these problems are not treated.

The results of the present study also have implications for single gating procedures, in which classification decisions are made using a single piece of information. Although there were wide discrepancies between maternal and teacher reports of conduct problems in the present study, it was clear that those who were rated as having conduct disorder by either mothers or teachers demonstrated similar adverse outcomes later in life. This suggests that the most appropriate single gating approach for the classification of conduct problems in children would involve obtaining information from multiple informants, but classifying children as conduct disordered on the basis of a single report of disorder, rather than multiple reports as in multiple gating, thereby increasing the sensitivity of the screening and classification process.

Limitations

The limitations of present the study include that the classification of conduct problems has been made on the basis of distributional criteria rather than standardised diagnostic criteria, and that observational data to cross validate mother and teacher reports was not available. Further the data were gathered within a specific cohort studied in a specific social context at a specific historical period. The extent to which the findings from this cohort generalise to other setting and context remains to be established.

However, within these limitations the present study leads to the conclusion that: (a) the majority of childhood conduct problems appear to be situation-specific rather than generalised; (b) children with situational conduct problems are at increased risks of later adverse outcomes, with levels of risk that are similar to those faced by children who exhibit generalised conduct problems; and (c)

there is a need for greater recognition of the significance of situation-specific conduct problems in

both developmental theory and in the treatment of childhood conduct disorders.

Key points

- A number of children show evidence of conduct problems that are specific to a single context such as home or school.
- It is unclear whether children with conduct problems in one setting share the same risks as those who show conduct problems in multiple settings.
- The present study showed that the majority of children displayed conduct problems in one context rather than multiple contexts.
- However, the risk of adverse outcomes for children with conduct problems was similar, irrespective of the context in which conduct problems were observed.
- A focus on those children who display conduct problems in multiple contexts is likely to overlook the needs of children whose conduct problems are confined to a specific context.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders (4th ed.)*. Washington, DC: American Psychiatric Association.

Anderson, J., Williams, S., McGee, R., & Silva, P. A. (1987). DSM-III disorders in preadolescent children. *Archives of General Psychiatry, 44*, 69-76.

Arseneault, L., Moffitt, T. E., Caspi, A., Taylor, A., Rijsdijk, F. V., Jaffee, S. R., Ablow, J. C., & Measelle, J. R. (2003). Strong genetic effects on cross-situational antisocial behaviour among 5-year-old children according to mothers, teachers, examiner-observers, and twins' self-reports. *Journal of Child Psychology and Psychiatry, 44*(6), 832-848.

Carlin, J. B., Wolfe, R., Coffey, C., & Patton, G. C. (1999). Tutorial in Biostatistics. Analysis of binary outcomes in longitudinal studies using weighted estimating equations and discrete-time survival methods: Prevalence and incidence of smoking in an adolescent cohort. *Statistics in Medicine, 18*, 2655-2679.

Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126*(6), 884-888.

Conners, C. K. (1970). Symptom patterns in hyperkinetic, neurotic and normal children. *Child Development, 41*(3), 667-682.

Deligatti, N., Akin-Little, A., & Little, S. G. (2003). Conduct disorder in girls: Diagnostic and intervention issues. *Psychology in the Schools, 40*(2), 183-192.

Elliott, D. S., & Huizinga, D. (1989). Improving self-reported measures of delinquency. In M. W. Klein (Ed.), *Cross-national research in self-reported crime and delinquency* (pp. 155-186). Boston, MA: Kluwer.

Fergusson, D. M., & Horwood, L. J. (1989a). Estimation of method and trait variance in ratings of conduct disorder. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 30*(3), 365-378.

Fergusson, D. M., & Horwood, L. J. (1989b). A latent class model of smoking experimentation in children. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 30*(5), 761-773.

Fergusson, D. M., & Horwood, L. J. (1993). The structure, stability and correlations of the trait components of conduct disorder, attention deficit and anxiety/withdrawal reports. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 34*(5), 749-766.

Fergusson, D. M., Horwood, L. J., & Lloyd, M. (1991). Confirmatory factor models of attention deficit and conduct disorder. *Journal of Child Psychology & Psychiatry, 32*(2), 257-274.

Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1995). The stability of disruptive childhood behaviors. *Journal of Abnormal Child Psychology, 23*(3), 379-396.

Kim-Cohen, J., Arseneault, L., Caspi, A., Tomas, M. P., Taylor, A., & Moffitt, T. E. (2005). Validity of DSM-IV conduct disorder in 41/2-5-year-old children: A longitudinal epidemiological study. *Am J Psychiatry, 162*(6), 1108-1117.

Lahey, B. B., Loeber, R., Quay, H. C., Frick, P. J., & Grimm, J. (1997). Oppositional defiant disorder and conduct disorder. In T. A. Widiger, A. J. Frances, H. A. Pincus, R. Ross, M. B. First & W. Davis (Eds.), *DSM-IV sourcebook* (Vol. 3, pp. 189–209). Washington, DC: American Psychiatric Association.

LeBlanc, M. (1998). Screening of serious and violent juvenile offenders: Identification, classification, and prediction. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions*. London: Sage.

Loeber, R., Dishion, T. J., & Patterson, G. R. (1984). Multiple Gating: A Multistage Assessment Procedure for Identifying Youths at Risk for Delinquency. *Journal of Research in Crime and Delinquency, 21*(1), 7-32.

Matthys, W., Maassen, G. H., Cuperus, J. M., & van Engeland, H. (2001). The assessment of the situational specificity of children's problem behaviour in peer-peer context. *Journal of Child Psychology and Psychiatry, 42*(3), 413-420.

McCombs Thomas, A., Forehand, R., Armistead, L., Wierson, M., & Fauber, R. (1990). Cross-informant consistency in externalizing and internalizing problems in early adolescence. *Journal of Psychopathology and Behavioral Assessment, 12*(3), 255-262.

McGee, R., Feehan, M., Williams, S., Partridge, F., Silva, P. A., & Kelly, J. (1990). DSM-III disorders in a large sample of adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 29*(4), 611-619.

Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, Health and Behaviour*. London: Longmans.

Scott, S. (2007). Conduct disorders in children. *BMJ, 334*(7595), 646.

Scott, S. (2008). An update on interventions for conduct disorder. *Adv Psychiatr Treat, 14*, 61-70.

Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratchowill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology, 45*(2), 193-223.

Spiker, D., Kraemer, H. C., Constantine, N. A., & Bryant, D. (1992). Reliability and validity of behavior problem checklists as measures of stable traits in low birth weight, premature preschoolers. *Child Development, 63*(6), 1481-1496.

Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The Revised Conflict Tactics Scales (CTS2). Development and preliminary psychometric data. *Journal of Family Issues, 17*(3), 283-316.

Touliatos, J., & Lindholm, B. W. (1981). Congruence of parents' and teachers' ratings of children's behavior problems. *Journal of Abnormal Child Psychology, 9*(3), 347-354.

van de Pol, F., Langeheine, R., & de Jong, W. (1999). *PANMARK 3 User's Manual. Panel Analysis Using Markov Chains A Latent Class Analysis Program* (*Version 2*). Voorburg, The Netherlands: Netherlands Central Bureau of Statistics.

Walker, H. M., & Severson, H. H. (1990). *Systematic screening for behavior disorders (SSBD).* Longmont, CO: Sopris West.

World Health Organization. (1993). *Composite International Diagnostic Interview (CIDI).* Geneva, Switzerland: World Health Organization.

Table 1 Summary of latent class models

a) Goodness of fit statistics for 2, 3 and 4 class models

Number of latent classes	Log likelihood ratio chi square		Bayesian Information Criteria (BIC)
	χ^2 (df)	p	
2 classes	180.1 (58)	<.0001	3423
3 classes	97.0 (55)	<.0001	3360
4 classes	59.0 (52)	.23	3343

b) Fitted parameters (standard errors) for the 4 class model

	Latent Class			
	No Problems	Mother Reported Problems	Teacher Reported Problems	Generalised Problems
Latent class probabilities	.832 (.023)	.058 (.014)	.076 (.020)	.034 (.011)
Response probabilities				
Mother report	.020 (.004)	.690 (.070)	.115 (.043)	.807 (.076)
Teacher report	.035 (.006)	.033 (.043)	.549 (.084)	.652 (.095)

Table 2 Mean mother and teacher reported conduct scores (7-9 years) by latent class assignment

	Latent Class			
Measure	No Problems (n=899)	Mother Reported Problems (n=45)	Teacher Reported Problems (n=64)	Generalised Problems (n=38)
Mother report				
7 years	97.6	120.9	104.7	123.1
8 years	97.6	119.2	104.9	125.8
9 years	97.6	119.8	105.2	124.2
Teacher report				
7 years	97.9	98.1	119.6	118.1
8 years	97.9	97.9	118.0	121.3
9 years	97.8	99.5	119.4	120.8

Table 3 Associations between conduct problems latent classes (7-9 years) and outcomes, ages 16-25

Outcome	No Problems (n=853)	Mother Reported Problems (n=41)	Teacher Reported Problems (n=55)	Generalised Problems (n=34)
		Conduct Problems Latent Class		
Offending				
% ten or more self-reported property offences ages 17-25	6.3	17.1	11.3	14.7
% ten or more self-reported violent offences ages 17-25	4.8	14.6	13.2	23.5
% arrested/convicted ages 16-25	9.2	24.4	24.5	41.2
% sent to prison ages 16-25	0.0	2.4	3.8	2.9
% reporting 5+ traffic offences ages 18-25	10.6	23.7	30.2	32.4
Mean (SD) number of offences	0.30[a] (0.7)	0.80[b] (1.23)	0.83[b] (1.10)	1.15[b] (1.35)
IRR (95% CI)	1.00 (-)	2.63 (1.83-3.17)	2.71 (1.97-3.73)	3.74 (2.67-5.24)
Mental Health				
% major depression/anxiety disorder ages 16-25	51.3	65.9	52.8	61.8
% conduct/anti-social personality disorder ages 16-25	6.9	22.0	27.3	26.5
% attempting suicide ages 16-25	5.7	19.5	13.2	14.7
Mean (SD) number of mental health problems	0.64[a] (0.68)	1.07[b] (0.82)	0.91[b] (0.91)	1.03[b] (0.58)
IRR (95% CI)	1.00 (-)	1.69 (1.24-2.30)	1.43 (1.07-1.91)	1.62 (1.15-2.28)

Substance dependence

% nicotine dependent ages 18-25	28.4	46.3	45.3	58.8
% alcohol dependent ages 16-25	11.6	17.1	20.8	17.7
% illicit drug dependent ages 16-25	12.3	22.0	22.6	26.5
Mean (SD) number of substance dependence problems	0.52[a] (0.79)	0.85[b] (0.96)	0.89[b] (0.97)	1.03[b] (0.94)
IRR (95% CI)	1.00 (-)	1.63 (1.16-2.31)	1.70 (1.26-2.29)	1.97 (1.40-2.78)

Relationship and pregnancy/parenthood issues

% reporting interpartner violence age 25	9.6	21.2	9.5	25.9
% reporting 10 or more sex partners ages 18-25	29.2	44.7	47.2	38.2
% becoming a parent by age 21	12.4	34.2	17.5	35.3
% pregnant/got partner pregnant by age 21	16.7	41.5	21.8	41.2
Mean (SD) number of relationship and pregnancy/parenthood issues	0.65[a] (0.85)	1.34[b] (1.26)	0.90[b] (0.99)	1.35[b] (0.98)
IRR (95% CI)	1.00 (-)	2.08 (1.58-2.74)	1.39 (1.04-1.85)	2.10 (1.55-2.83)

Note: Differing superscripts [a, b] indicate statistically significant ($p<.05$) differences between groups