Ontario Association of Residences Treating Youth

Centre of Excellence for Residential Care in Ontario

Residential Care and Treatment

A Longitudinal Study of Outcomes

Partners in Care V

May 2012

Partners in Care V is published by:

Residential Care and Treatment

A Longitudinal Study of Outcomes

Executive Summary

The Ontario Association of Residences Treating Youth (OARTY) is a provincial association representing 72 member agencies which since 1971 has been providing high quality treatment and residential care services to over 3,000 children, youth and young adults daily. OARTY strives to promote the provision of high quality residential care and treatment for vulnerable children, youth, young adults and their families.

OARTY funds and collects information for the Partners in Care (PIC) series of reports for the purpose of establishing the clinical profile of the clients in residential care; to establish referral patterns; to establish the types and costs of programs within our organization; and to track and monitor outcomes relating to the care of the children, youth, and young adults placed with our member agencies. In 1996, OARTY began a longitudinal study of outcomes based on nine measures. This report is the first analysis of the 1996-2011 OARTY longitudinal study.

In October 2011, OARTY initiated a second longitudinal study with additional instruments to measure outcomes for children who are medically fragile, autistic, dual diagnosis and FASD, in addition to children with serious emotional and behavioural problems. The second longitudinal study of outcomes is referred to as the Next Generation LS.

The findings of the 1996-2011 OARTY longitudinal study are as follows:

There are 888 unique clients in the longitudinal dataset, containing

a) 632 unique clients with follow-up data (i.e.) wave one and a matched client with wave two
b) 388 unique clients with three waves of test data
c) 252 unique clients with four waves of test data
d) 145 unique clients with five waves of test data
e) 60 unique clients with six waves of test data
f) 19 unique clients with seven waves of data

The difference in scores comparing wave 1 to wave 2, wave 2 to wave 3, wave 3 to wave 4 and so on was found for each measure. The difference was divided by the pooled standard deviation for the test pair. The resulting metric is a standardized change score, referred to either as *Cohen's d* or the *effect size*. International standards exist for interpreting the effect size. Depending on the value of the effect size, the treatment outcomes can be described as small, medium, large and very large.

The effect sizes for the Conners' Global Index, a measure of ADHD symptoms and psychological distress were between .51, and .68. This indicates a strong positive treatment effect according to international standards for social service agencies.

> OARTY agencies produced large positive improvements on the symptoms of ADHD at every occasion the child was tested. ADHD symptoms are indicative of neuro-behavioural disorders and of psychological distress in general.

The SA-45 Global Severity Index is a measure of broad psychiatric symptoms. The outcomes of children demonstrated a very strong treatment effect across four years of service. During each cycle of testing, the children made clinically significant improvement in a broadly based range of psychiatric symptoms. At the end of four years, 68% of all of the children moved from the clinical range to the normal range.

The FAB-C Problem Index is a measure of antisocial attitudes, negative feelings, poor peer relationships and misbehaviour. The outcomes of children demonstrated a very strong treatment effect across four years of service. The percentage of children whose scores crossed the threshold from at risk to normal scores varied from 57% to 83%; an average response rate of 69% across 3.3 years of service.

> The results of the FAB-C and the SA-45 indicate that OARTY agencies are able to eliminate a substantial portion of those at risk of serious emotional and behavioural problems within a three year period.

The Children's Global Assessment Scale is a measure of how well children are able to function at home and school. The outcomes of children demonstrated a strong treatment effect

across four years of service with the most substantial improvement occurring the first year of treatment.

After four years of service 67% of all children whose score was below the at risk threshold crossed over into the normal range. One third of the children were still functioning poorly. However, most of those who were functioning poorly had dropped out before the study was completed.

> The results on the CGAS indicate that the OARTY agencies are able to improve the child's ability to function in school, community and at home to a clinically significant degree. The only reservation is that complete success for the whole cohort of children at risk on this measure requires that the child spend sufficient time in treatment.

The Level of Assistance Scale (LOA) is a measure of how much adult support the child requires to get through the day in relation to basic tasks, such as eating, dressing and following rules. The outcomes for children on the LOA demonstrate a strong positive treatment effect, especially at the end of the first year of service. Fifty eight percent of the children in the at risk cohort ended up in the normal range

> The results on the LOA indicate that the OARTY agencies are able to improve the child's independence in daily living skills to a clinically significant degree.

The amount of adversity in the background of the children tested on the instruments above is significant. Compared to other longitudinal studies of normal populations, the children in the 1996-2011 OARTY longitudinal study had much higher levels of adversity, e.g., physical and sexual abuse, poverty, parental criminality, parental substance abuse and parental mental illness. The overall average number of different types of adversity shown by clients of OARTY member agencies is 4.0. Children with this much adversity have a 66% chance of failing to make the transition to young adulthood as functional, independent people.

The number of different types of adversity varies significantly by the type of client, with children who are completely normal on testing having the lowest number (2.1) and children who have intellectual deficits and serious behaviour problems having the highest number of different types of adversity (4.9).

The amount of adversity is reflected in a poor level of attachment for many children in relation to their closest caregivers.

Attachment was measured by the Parental Bonding Instrument, which has three scales; the first measuring the degree of affection the child perceives from his/her caregiver. 60% of children in the study began their placement with a very low affection score. The average for this at risk group was 30.7.

The average score measuring affection for the at risk children was below the 3rd percentile of teenagers all over the world. When these children were placed in residential care, they felt that their parents disliked them intensely.

The outcomes for the at risk children demonstrated a strong treatment effect every time they were tested. The results improved year by year. Indeed, the outcomes at the end of four years demonstrated a very strong treatment effect.

The second scale of the PBI is able to identify children who feel they are neglected and unsafe when reflecting on his/her closest caregiver. 28% of the children felt they were neglected and unsafe when first admitted. The average score for this at risk cohort was 40.4. In comparison with the normative reference group, the OARTY at risk group was below the 16th percentile.

The outcomes for the children who were in the at risk group demonstrated a very strong treatment effect. The largest change occurred at the end of 3.25 years of treatment. During the wave 4 to 5 test period, 94% of the at risk cohort had crossed the threshold indicating they feel safe and were not neglected. Eighty percent of those at risk on this measure, returned to the normal range during the study.

The third scale of the PBI measures the degree of unfairness, i.e., the child feels he/she was "over-controlled" resulting in the child feeling his/her parents did not understand him/her.

52% of the children felt they were treated unfairly at wave #1. The average score for this at risk cohort was 72.9. Less than 2% of teenagers in the normative reference group produced a score equal to or higher than the OARTY sample.

The outcomes for children in this at risk group demonstrated a very strong treatment effect. The most substantial change occurred at the end of 3.25 years of treatment. During the wave 4-5 test period, 47% of the at risk cohort had crossed the threshold indicating they feel their needs were understood and they were fairly treated.

Children who think that they are intensely disliked or treated unfairly by the very people they feel closest to are children with an insecure attachment. The change that OARTY members create is a change in perception. The children came to believe that they are loved, safe and treated fairly. They began to feel securely attached.

The measures used in this longitudinal study are directly related to the most potent factors predicting life span outcomes for children with serious emotional and behavioural problems. The instruments identify children who are at risk of developing enduring mental illness or an inability to function as independent adults. The criteria for determining who is at risk were based on other longitudinal studies referenced in this report.

The risk factors measured were:

- hyperactivity, impulsiveness, inattention and emotionality

- broad psychiatric symptoms mapped to the DSM-IV R
- attitudes, feelings and behaviour
- ability to perform the social roles of student, family member and community member successfully
- degree of adult support required to perform 22 daily living tasks, such as eating and crossing the street, ranging from totally independent to fully dependent on the adult to do everything
- degree of adversity in the background before admission
- degree of affection felt by the child in relation to his/her closest caregivers
- degree of neglect felt by the child in relation to his/her closest caregivers
- degree of unfairness and over-control felt by the child in relation to his/her closest caregivers

The degree of adversity does not change over time, but these risk factors affect the life span outcomes.

For all of the measures subject to change, the data shows a strong to very strong positive treatment effect. The majority of children who started the study within the at risk zone crossed over to the range of scores indicative of a normal child.

Table of Contents

Context for the Report

The Ontario Association of Residences Treating Youth (OARTY) has 72 member agencies which provided 804,906 days of care in 2011. The clients of the residential service included 273 adults and 3,388 children, for a total of 3,661. In 1996, OARTY began a longitudinal study of outcomes based on nine measures. The members were provided with software, referred to as the OARTY Information System (OIS), to collect the data and export it to OARTY.

This report is the first analysis of the 1996-2011 OARTY longitudinal study. In October 2011, OARTY initiated a second longitudinal study with additional instruments to measure outcomes for children who are medically fragile, autistic, dual diagnosis and FASD, in addition to children with serious emotional and behavioural problems. New software was provided to the member agencies to collect the data and export it to OARTY.

The types of homes providing care and treatment is displayed below:
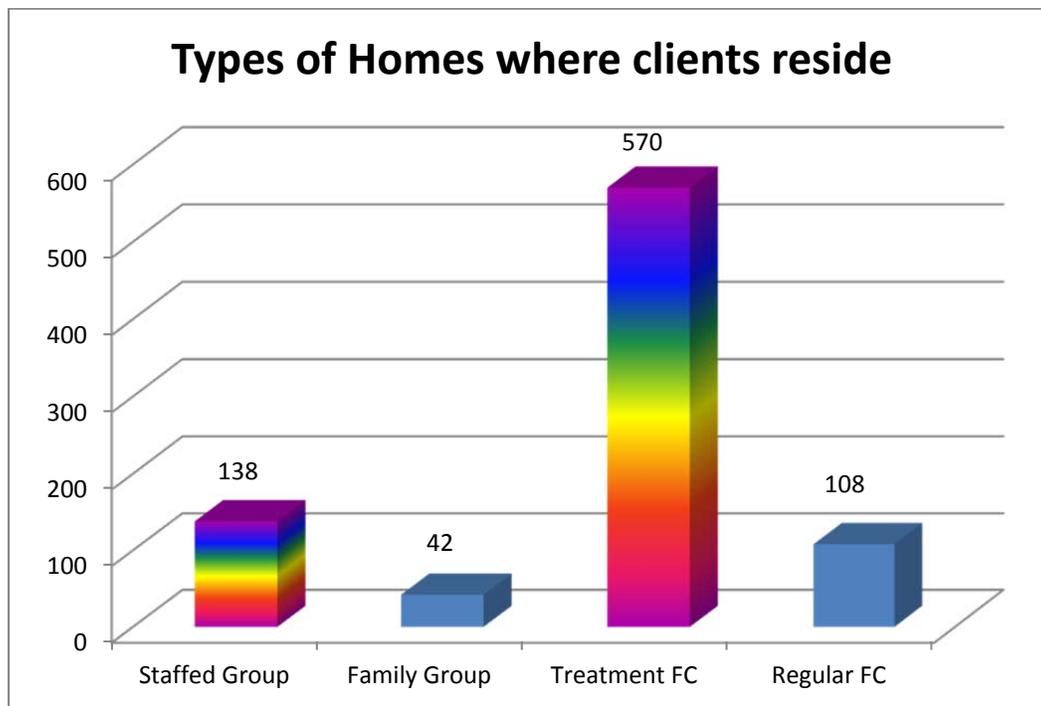
The retrospective longitudinal study was based on five agencies that provided service in staffed group and in treatment foster care, indicated by the rainbow coloured bars.

Status of the Next Generation Longitudinal Study

The next generation longitudinal study, referred to as the Next Generation LS, has had a strong beginning. OARTY has received the baseline scores for all clients currently in treatment in 21 agencies out of the 72 member agencies. The next generation data repository includes 1,146 clients.

There are 798 clients in the Partners In Care 5 dataset that were served in 2011; this represents 22% of the grand total of 3,661 clients served in 2011. 337 cases in the database were admitted and discharged before Jan 1, 2011 and 11 clients admitted after Jan 1, 2012

During the initial phase of the Next Generation LS, OARTY members were trained on the instruments and the software. The instruments are dual purpose providing assessment and outcomes on a case by case basis.

Cases in the Retrospective Longitudinal Study

There are 888 unique clients in the 1996-2011 OARTY Longitudinal Study, containing

- 632 unique clients with follow-up data, i.e., wave one and a matched client with wave two
- 388 unique clients with three waves of test data
- 252 unique clients with four waves of test data
- 145 unique clients with five waves of test data
- 60 unique clients with six waves of test data
- 19 unique clients with seven waves of data

1996-2011 OARTY Longitudinal Study

OARTY implemented an outcome testing protocol in fiscal year 1996. The protocol was based on a multi-gate risk screening system, comprised of a battery of tests, measuring pre-morbid symptoms of disorder at gate one, followed by another battery of tests measuring catalytic and protective factors at gate two.

The screening system and the actual instruments are evidence based and in wide use throughout the world. The system identifies children at risk of developing enduring mental health problems or at risk of functional impairment as a young adult. The term functional impairment means that the young person is unable to make the transition to adulthood as an independent capable adult.

The gate #1 instruments measure the premorbid risk factors for adverse outcomes that seriously compromise the health and safety of either the child or the public. These concepts were measured by:

Instrument	Concept Measured
Conners' Global Index	ADHD and psychological distress
SA-45's Global Severity Index	emotional disturbance
FAB-C Problem Index	
Children's Global Assessment Scale	adaptive failure
Level of Assistance	level of adult support in daily living

The prognosis for a child with serious problems, as measured in the gate #1 instruments, in the areas of mental health, fitting into society or the child's behaviour is determined by a combination of the accumulation of very stressful events and adverse conditions from early life and the quality of the child's attachment to his or her primary care givers. These concepts were measured by:

Instrument	Concept Measured
Parental Bonding Instrument	quality of attachment
Sociodemographic Checklist	adverse family history

Five OARTY member agencies participated in a 15 year longitudinal study of all their clients in residence. The children were tested within 30 days after admission and every nine months thereafter until they were discharged.

Intercorrelation for all Measures

A correlation analysis was performed for all measures on the 1st wave of testing. The correlation matrix shows a low level of correlation with each other. Specifically, the highest correlation was between the CGAS and Level of Assistance (r = -.459, sig = .000, n =388). The Level of Assistance scale was also moderately correlated with the CGI (r = .400, sig = .000, n = 387). This fact means each instrument produces new information about the outcomes not included in the other tests.

Nature of a Longitudinal Design

Each individual client of OARTY who was enrolled in the longitudinal study had a unique Client ID# so that the corresponding data could be tracked throughout their time at a member agency; as OARTY does not collect or store identifying information on the children in care we cannot track the individual client once they transfer out. As a result of the unique Client ID#, we are able to observe how the scores change from the first test to the last for each

client. Each child has a series of test pairs, i.e., change from time 1 to time 2, time 2 to time 3, time 3 to time 4, etc. By convention, these pairs are referred to wave 1-2, wave 2-3, wave 3-4 and so on.

The change in scores was assessed with three statistical measures:

- Effect Size, (i.e.. Cohen's d[i])
- T-score of the difference in means
- Analysis of Variance or Regression Analysis as appropriate

The effect size is the ratio between the raw difference in scores between one test and the next divided by the pooled standard deviation. The guidelines for interpreting effect sizes in the social sciences are:

- small effect size, $r = 0.1 - 0.23$
- medium effect size, $r = 0.24 - 0.36$
- large effect size, $r = 0.37$ or larger[1].

Cohen's d displays very little "small sample bias" and is a valid indicator of treatment outcomes with samples between 10 and 25.[2]

The T-score of the difference in means is a measure that determines if the difference could be explained by random chance; otherwise, the difference is considered statistically significant.

The Analysis of Variance or the Regression Analysis is used to assess whether a specific aspect of the case may be correlated with the change in scores. For each test pair, appropriate statistics were used to examine possible relationships between the difference in score and the following variables:

a. Member agency where the child resides
b. Gender
c. Age at time of test
d. Length of time in treatment
e. Degree of adversity before admission
f. History of various adverse conditions
   - Life long poverty
   - Sexual abuse
   - Physical abuse
   - Family member in jail
   - Family member with mental illness
   - Family member with intellectual deficit
   - Family member abusing drugs or alcohol

---

[1] Jacob Cohen (1988). *Statistical Power Analysis for the Behavioral Sciences* (second ed.). Lawrence Erlbaum Associates

[2] Lipsey, Mark W. & Wilson, David B. (1993), "The Efficacy of Psychological, Educational and Behavioural Treatment: Confirmation from Meta-analysis", *American Psychologist*, 48(12): 1181-1209

- Child abusing drugs or alcohol
- Child has a history of failure in school from primary grades
- Child's mother began parenting in her teens

CRITERIA FOR DETERMINING A POSITIVE OUTCOME

The outcomes for children served were assessed by nine distinct measures that focus on distinct aspects of the child's mental health. Each measure is widely used for this purpose around the world. The introduction to each measure includes a review of the literature proving that the measure is valid and reliable.

All of the measures have a range of scores that are described as normal, borderline or clinical according to the test publishers. The normal range is quite wide, for example, a t-score between 40 and 60 on the Conners' CGI is considered normal. If a normal child is tested across several years, his or her scores would move up or down within the normal range; this is not indicative of a presenting issue.

All of the tests have a range of scores that would be considered of borderline clinical significance, meaning the child may be developing a serious problem in the area measured by the test.

All of the tests have a range of scores in the clinical range, with prevalence rates of less than 2% in the normal population. Scores in the clinical range indicate a child is at risk of mental illness or functional impairment. The primary service goal for OARTY member agencies is to reduce the number of children who are at risk; this can be measured by the number of children that cross the at risk threshold into the normal or borderline range. Each test has been assigned a threshold score that indicates the boundary line between at risk and normal or borderline. Crossing the boundary line equals a positive outcome.

Some children do appear to get worse before they get better. This is a common occurrence in the first year of treatment when the child begins to think about his situation in life and becomes distressed over how unfair life has been. When children cross the at risk threshold in a negative direction, this is referred to as deterioration.

The tests included in this study can only be administered with the child's consent. As such it is to be expected that several children will miss one or more tests. This will be referred to in the discussion below as "missing tests".

Each child included in the longitudinal study has a different number of test pairs depending on when they entered and how long they were in treatment. The longitudinal study covers a period of four years of treatment with five test pairs. If a child does not complete all five test pairs, we referred to this situation as "research drop-outs".

Conners' Global Index

The Conners' Global Index or CGI is a 10-question behavior checklist completed by either or both the parent and teacher. The term "parent" in this context refers to his current caregiver. This could be a foster parent, adoptive parent or child care worker. The CGI

measures the frequency of four types of behaviour: hyperactivity, inattention, impulsiveness and emotionality. The CGI is used to measure the degree of ADHD, a neuro-behavioural disorder that responds well to stimulant medication. The CGI also identifies children who display high levels of ADHD symptoms caused by other factors such as psychological distress and co-morbid psychiatric diagnoses. As a result of the many factors producing ADHD symptoms, the CGI is a leading indicator of a number of mental health issues. Longitudinal studies have found that ADHD along with poor peer relations in early childhood are the best predictors of children who are at risk for developing enduring serious mental health problems[3].

The CGI is a valid indicator of mental health needs. The T-score from the Conners' is highly correlated (r = 0.82) with the Total Problems T-score from the Achenbach (Achenbach, 1991, page 85). Secondly, the Conners' Global Index was highly correlated with the total score of the Kovacs Children's Depression Inventory (CDI) at .71 (Conners' Manual, page 131). These findings demonstrate convergent validity. The average scores of the CGI for children with no disorder (52.7), those diagnosed with ADHD (69.8) and those diagnosed with another disorder (66.7) demonstrate discriminant validity.[4]

The normal range on the Conners' CGI is between 40 and 60. In order to observe the number of positive outcomes, we started the data analysis by selecting the children whose t-scores on the CGI were above 70; the high end of the range considered borderline clinical significance. There are 609 children who were tested using the CGI; 61% had a score above the threshold of 70. The statistics for this at risk cohort are as follows: mean = 81.55, SD = 6.098, n = 262. A t-score of 81.55 means this cohort had more ADHD symptoms than 1 in 1,000 children in society. The graph below illustrates this cohort over time.

---

[3] Rutter, Michael and Seija Sandberg (1985), "Epidemiology of Child psychiatric Disorder: methodological Issues and Some Substantive Findings", *Child Psychiatry and Human Development*, 15(4), 209-233

[4] Sullivan, Jeremy R. & Riccio, Cynthia A. (2007), "Diagnostic Group Differences in Parent and Teacher Ratings on the BRIEF and Conners' Scales", *Journal of Attention Disorders*, 11 (3), 398-406

GRAPH #1   The Impact of Treatment on ADHD Symptoms

Changes in ADHD symptoms
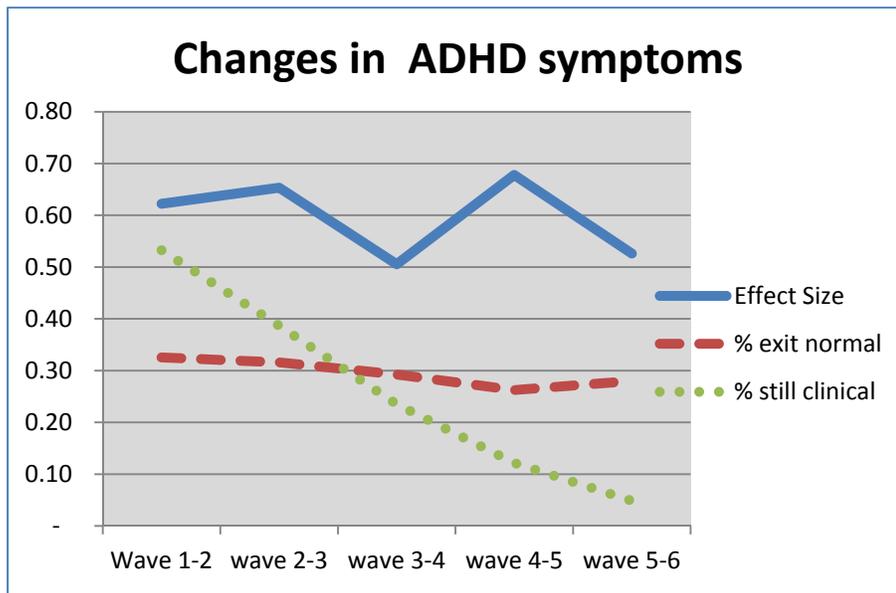
Graph #1 displays the average effect sizes for the CGI across 4.0 years of treatment, indicating by the solid line. The graph shows that OARTY members consistently produced a large effect at each and every wave of testing.

As the testing continued, the at risk group expanded by 152 children, who deteriorated after the 1st test. The children who deteriorate represents 33% of the total at risk group on the CGI.

During each cycle of testing, a substantial percentage of children crossed the at risk threshold on the CGI showing clinically significant improvement in ADHD symptoms. This percentage is indicated by the heavy dashed line, labelled "% exit normal". At the end of four years, 48% of all the children at risk had a positive outcome.

The opposite story is the percentage of non-responders who remained within or deteriorated into the at risk range over time. This group is indicated on the graph by the dotted line. The percentage of non-responders falls from 53% after three months to 5% after four years. Some of these non-responders became drop-outs before the study was completed. The final position of all cases in the expanded at risk group is displayed in graph #1.1:

GRAPH #1.1   4-YEAR ACCUMULATED OUTCOMES (CGI)

**4- year accumulated outcomes
Conners' Global Index**

24%

48%

52%

■ % became normal

■ % still clinical

■ % of drop outs
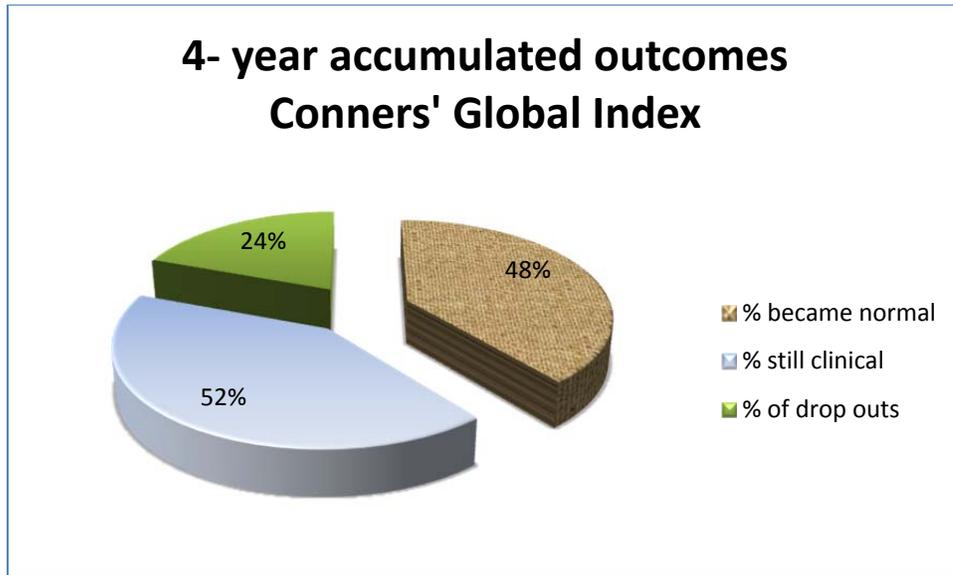
52% of children in the at risk group who were still in the clinical range at the end of four years is high compared to the other tests. This suggests that ADHD symptoms are very persistent. This has been confirmed by Michael Rutter (ibid).

About one-half or 24% of the group who were still in the clinical range dropped out of the research. The drop out group may have shown improvement had they continued.

CORRELATION OF FACTORS WITH THE PATTERN OF CHANGE OBSERVED

We asked the question, "what factors are correlated with the pattern of change observed?" Using Analysis of Variance and Multifactor Regression Analysis, we found:

The average CGI score for the at risk cohort varies significantly by agency shortly after admission (F = 3.5, n = 369, sig = .008). The average scores for all agencies were higher than 80, meaning these children have more ADHD symptoms than one in 1,000 children in society. The clients for some agencies are marginally higher than in other agencies. The clients from different agencies produced the same average score after one year of treatment; i.e., there was no statistically significant difference at wave 2.

This means that the five participating agencies are treating the same population of children. If each agency had large differences between the children at intake on this measure, then we could apply the findings across agencies. Indeed, we would have to declare that there were five longitudinal studies rather than one.

The fact that there were no differences in the CGI measure between agencies at later stages of the research means that the agencies are equally effective.

Female clients have significantly more ADHD symptoms than males at the beginning of treatment (females = 84, males = 81, F = 16.9, n = 368, sig = .000). Examining the cohort that was still at risk three years later, females had significantly higher ADHD scores than males (females = 82, males = 74, F = 4.3, sig = .04)During the study, the majority of at risk girls crossed the threshold into the healthier range, but the remaining group was still significantly worse compared to the boys.

There was no difference in the level of ADHD symptoms for different groups of clients by different forms of adversity in their background. Clients who were sexually abused, physically abused or grew up in poverty did not have a worse profile than children who were not abused. Clients whose parents were disturbed or dangerous did not have a worse profile than children whose families were not.

The Symptom Assessment – 45 (SA-45)

The SA-45 is a one page form with 45 questions completed by the teenager. This type of assessment instrument uses a statement of self to represent the DSM diagnostic criteria. The teenager responds by rating "how much the problem has bothered or distressed you during the last 7 days including today" on a five point scale: not at all (1), a little bit (2) ... extremely (5).

The questions are evidently linked to psychiatric symptoms. For example, "feeling afraid to travel on buses, subways or trains", "feeling tense or keyed up", "shouting and throwing things" and "the idea that you should be punished for your sins". The content is clearly about thoughts and feelings.

The SA-45 has very low false positive and false negative rates of 10% each for identifying inpatient versus outpatient samples.

Borduin et al (1995) found that the Global Severity Index of the SA-45 "represents the best single indicator of the respondent's psychiatric functioning" (page 572)[5] . The GSI t-score of 66 or higher was used to identify the clients who were at risk of mental illness. Less than 5% of young people in society have a t-score in this range. The graph below illustrates what happened to this cohort over time.

---

[5] Borduin, C., Mann, B., Cone, L., Hengeller, S., Fucci. B., Blaske, D., & Williams, R. (1995), "Multisystemic Treatment of Serious Juvenile Offenders: Long-Term Prevention of Criminality and Violence", *Journal of Clinical and Consulting Psychology*, 63, 4, 569-578

Graph #2   Changes in Broad Psychiatric Symptoms
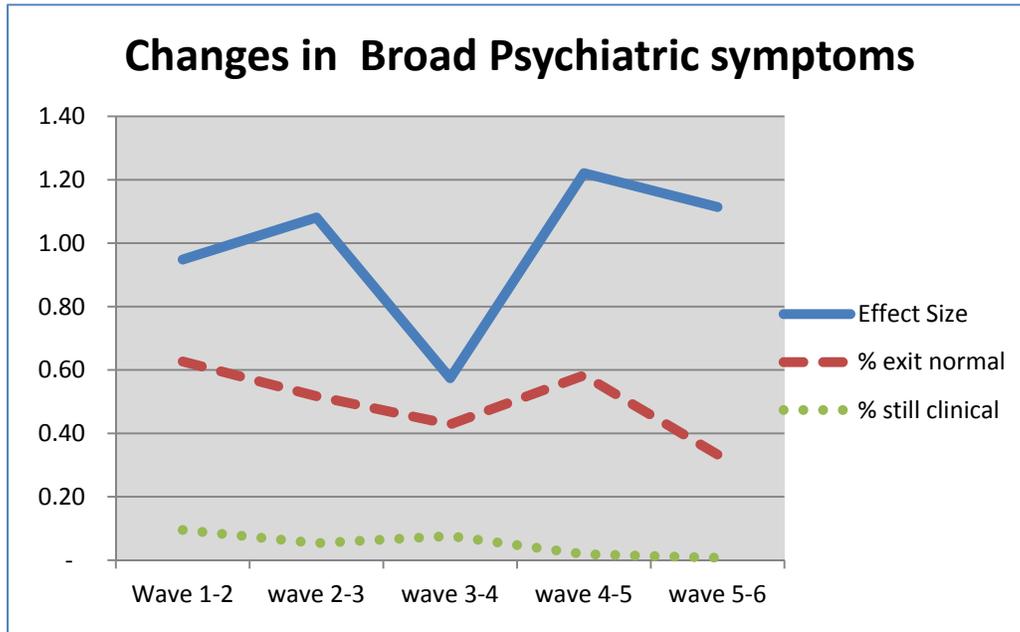
Three hundred and eighty-two (382) youth were tested with the SA-45. Only 32%, or 122 children, tested above the threshold for being at risk of mental illness. The initial scores for this at risk cohort are as follows: mean = 71.52, SD = 6.108, n = 122.
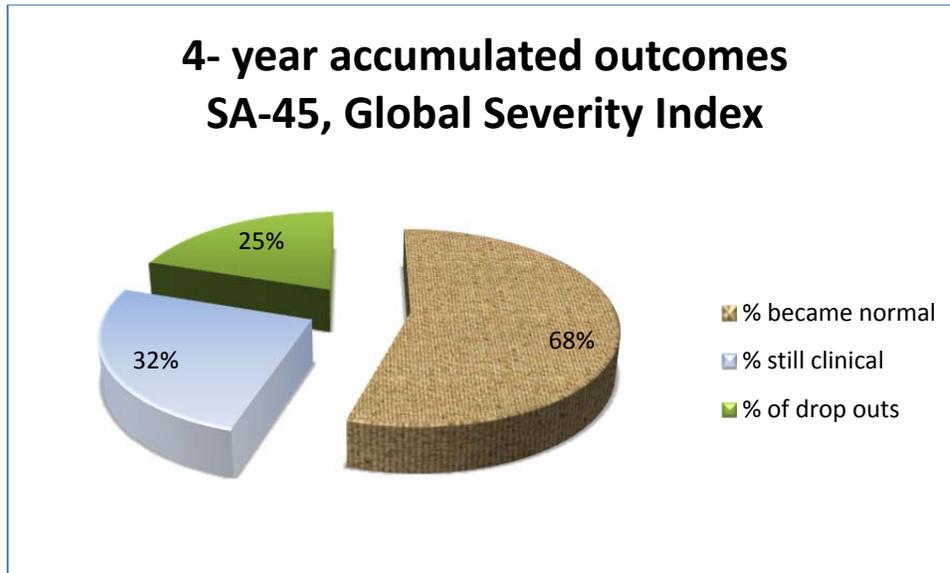
The solid line at the top indicates that there was a very strong treatment effect across four years of service in relation to broadly based psychiatric symptoms.

During each cycle of testing, a substantial percentage of children crossed the at risk threshold on the Global Severity Index demonstrating clinically significant improvement in a broadly based range of psychiatric symptoms. This percentage is indicated by the heavy dashed line, labelled "% exit normal". At the end of four years, 68% of all of the children at risk had a positive outcome.

The opposite story is the percentage of non-responders who remain within the at risk range over time. The highest percentage of non-responders is 7% dating from the first set of test pairs. This demonstrates that the OARTY member agencies were able to substantially reduce the psychiatric symptoms of anxiety and depression within one year of treatment. New cases of at risk youth appeared in later waves as the young people developed self reflective capacity and began to self identify their fears and sadness.

OARTY member agencies demonstrate a strong treatment effect with each new wave of testing. The percentage of children that remained in the clinical range over time represents 32% of the expanded at risk group.   Graph #2.1 displays the status at the end of four years.

GRAPH #2.1   4-YEAR ACCUMULATED OUTCOMES (SA-45)

**4- year accumulated outcomes
SA-45, Global Severity Index**

25%

32%

68%

- % became normal
- % still clinical
- % of drop outs

Most of the group who remained in the clinical range did not remain in the research for the full four years. The drop out group (25%) may have shown improvement had they continued.

A series of Analysis of Variance did not identify any significant relationships between the changes in the SA45 t-scores and other variables, such as the agency, gender of the youth, history of abuse and history of parental criminality or mental illness.

Feelings, Attitudes and Behaviour (FAB-C)

The FAB-C consists of 48 yes/no questions that ask the child to rate whether the statement describes him or her. The FAB-C is appropriate for children ages 6 to 15 years of age. Children with a first grade reading level can complete the test.

The FAB-C is not designed to map directly to DSM disorders. It is designed to identify the attitudes and feelings of school age children who seek out mental health services. In addition, it examines the child's behaviour with peers, a critical area for assessment. The FAB-C also has a lie scale. Higher scores indicate increasing severity of emotional and behavioral problems.

The FAB-C was validated with a population of children who had already been screened by clinicians and assigned to a clinic or assessed as not requiring intervention. The ability of the FAB-C to correctly classify children in a clinical sample using the problem index score was tested. The false positive rate was 27%; the test implies these children should be in the clinic sample when the clinicians had determined they were not in need of treatment. The false negative rate was 37% meaning that these children were classified as normal when, in fact, they were already in the clinic sample receiving treatment. In comparison, the SA-45 performed much better with false positive false and negative rates of 10% each. High scores on the FAB-C

problem index have been proven to identify children with significant mental health needs 63% of the time.

The FAB-C shows strong convergent and discriminant validity with the Conners' Rating System, the CBCL parent rated instruments, the Child Depression Inventory and clinician ratings of the same child. Associating with negative peers on the FAB-C is inversely correlated with social activities on the CBCL (-.27) and positively correlated with the clinician's rating of negative peer relations. The FAB-C conduct scale was correlated with aggression (.40) and antisocial behaviour (.49).

The internal validation studies for the FAB-C demonstrate excellent psychometric properties. The FAB-C Problem t-score of 66 or higher was used to identify the children who were at risk of serious emotional and behavioural problems. Less than 5% of young people in society have a t-score in this range. The graph below illustrates what happened to this cohort over time.

Graph #3   Changes in the FAB-C
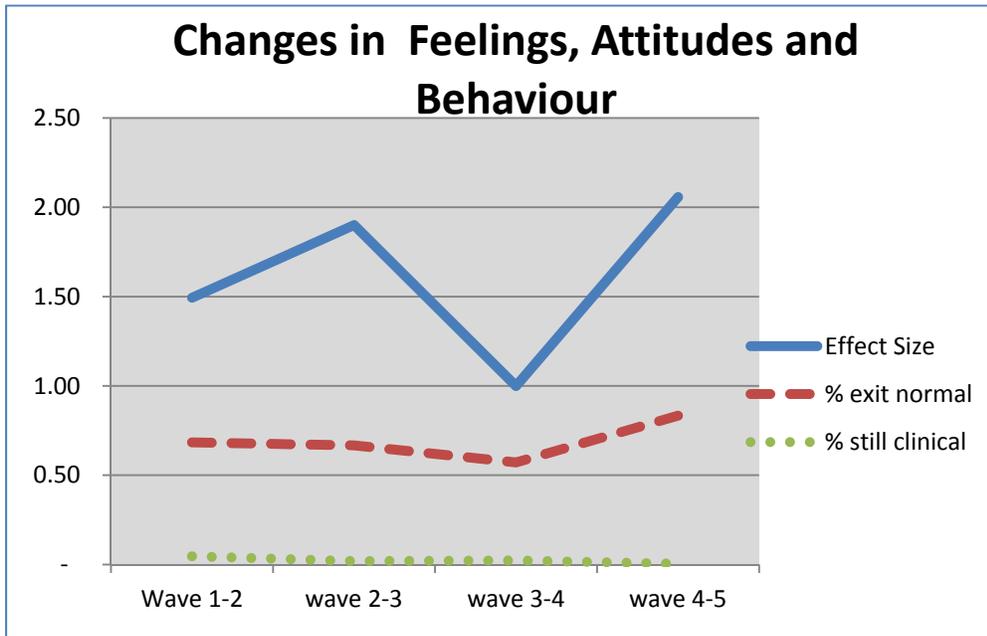
Thirty percent of the children who responded to the FAB-C had a score on the 1st test above t-score 65; the threshold for being at risk of serious emotional and behavioural problems. The initial scores for this at risk cohort are as follows: mean = 73.89, SD = 7.959, n = 64.

The solid line at the top indicates there is a very strong treatment effect across four years of service in relation to serious emotional and behavioural problems. The average treatment effect across four pairs of tests was 1.61.

The percentage of children whose scores crossed the threshold from at risk to normal scores varied from 57% to 83%, an average response rate of 69% across 3.3 years of service.

The results indicate OARTY member agencies are able to eliminate a substantial portion of those at risk of serious emotional and behavioural problems within a three year period for school age children and young adolescents.

Graph #3.1 displays the status of the at risk cases after four years of service. The graph shows a large percentage of the at risk children became normal in relation to this area of mental health.

GRAPH #3.1   4-YEAR ACCUMULATED OUTCOMES (FAB-C)

4- year accumulated outcomes
FAB-C, problem index

10%
21%
79%

- % became normal
- % still clinical
- % of drop outs
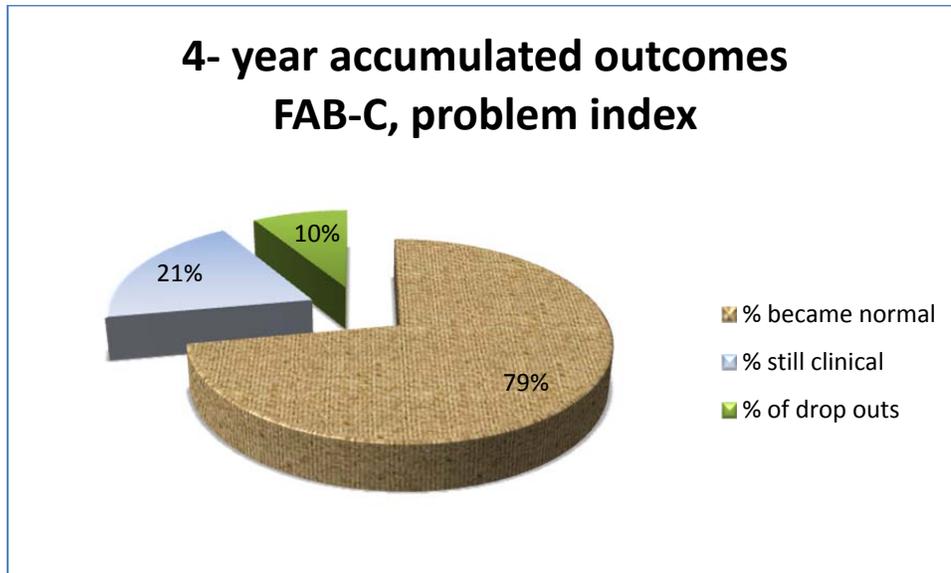
Half of the children in the clinical range did not complete the five pairs of the longitudinal study. The drop out group (10%) may have shown improvement had they continued.

An analysis of variance determined that the agency where the child was placed was not correlated with the scores at 1st and 2nd test periods.

Children's Global Assessment Scale (CGAS)

The CGAS[6] is a clinician rated criterion referenced scale for assessing the functioning of children and adolescents. Youth are assessed relative to how they perform their roles at home, school and community in the month prior to the test. The clinician must gather the information necessary to score the test from open ended assessment interviews with the child, parent and school including information in prior assessments.

The CGAS rating scale was constructed from ten exemplars or generic clinical profiles. Each exemplar is progressive from superior functioning to severe impairment. The worker evaluates the child's level of functioning to the exemplar best matched to the child's role

[6] Shaffer, D., Gould, M., Brasic, J., Ambrosinin, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983), "Children's Global Assessment Scale", *Archives of General Psychiatry*, 40, 1228-1231

performance. The manual instructs the worker to rate the child according to his/her worst level of functioning in the past month.

Haugen Schorre & Vandvik (2004) found 243 references to the CGAS on PubMed[7]. The CGAS has been used across the world in a triage function as gating criteria for admission to psychiatric hospitals[8] . The CGAS score can also be used as part of an assessment battery to determine the severity of the child's mental health needs[9]. The CGAS supports formulation and treatment planning. The test has been used to profile a population of children receiving service. It has also been used as a pre-post measure of outcomes.

The scientific basis for the CGAS was established in the research from developmental psychopathology. Stroufe & Rutter (1984) found that the strongest predictors[10] in childhood for later dysfunction (violence, mental disorder and inability to cope) "appear to be adaptational failures, defined in age-appropriate terms". The underlying question in all functional assessments is "how well is the child adapting to his or her environment?"

The correlations between the Columbia Impairment Scale (CIS) and the CGAS scores with the other measures of psychological dysfunction provide support for the concurrent validity[11] of the impairment measures. Discriminant validity was demonstrated by the fact the scores were significantly different for referred (mean = 61.8 ± 13.5) and non-referred children (mean = 85.6 ± 6.6). A CGAS score>71 is probably "non-cases", CGAS ≥ 61 but < 71 is "probable cases" and CGAS<61 is "definite cases". To be certain of the clinical threshold for the CGAS, we established CGAS <50 as our boundary score for determining treatment success.

Predictive validity showed the CGAS is sensitive to change and predicts response to treatment. A study by Gold et al., reviewed in Steinhausen & Metxke (Ibid) found that CGAS alone accounted for 30.7% of the variance of the logarithm of length of stay in residential treatment.

The CGAS is now regarded as valid for clients of ages 0 to 23 years. Steinhausen & Metxke (Ibid) conclude:

"Findings indicate that a unidimensional scale like CGAS is a better measure of change and predictor of outcome than diagnoses and multidimensional scales."

---

[7] Haugen Schorre, Bjørg Elisabeth & Vandvik, Inger Helene (2004), "Global assessment of psychosocial functioning in child and adolescent Psychiatry: A review of three unidimensional scales (CGAS, GAF, GAPD)", *European Child and Adolescent Psychiatry*, 13:273-286

[8] Newman, F.L., Richard H. Hunter, R.H. & Irving, D. (1987) "Simple Measures of Progress and Outcome in the evaluation of Mental Health Services", *Evaluation and Program Planning*, 1987, vol 10, 209-218

[9] Guzder, J., Paris, J., Zelkowitz, P. & Marchessault, K. (1996), "Risk Factors for Borderline Pathology in Children", *Journal of the American Academy of Adolescent Psychiatry*, 35 (1), 26-33

[10] Sroufe, L. & Rutter, M (1984), "The Domain of Developmental Psychopathology", *Child Development*, 55, 17-29

[11] Steinhausen, Hans-Christoph & Metzke, Christa Winkler (2001), "Global measures of impairment in children and adolescents: results from a Swiss community survey, *Australian and New Zealand Journal of Psychiatry*, 35:282–286

Graph #4 displays how the children change over time in the performance of their social roles.

Graph #4   Changes in Social Functioning

Changes in Social Competence
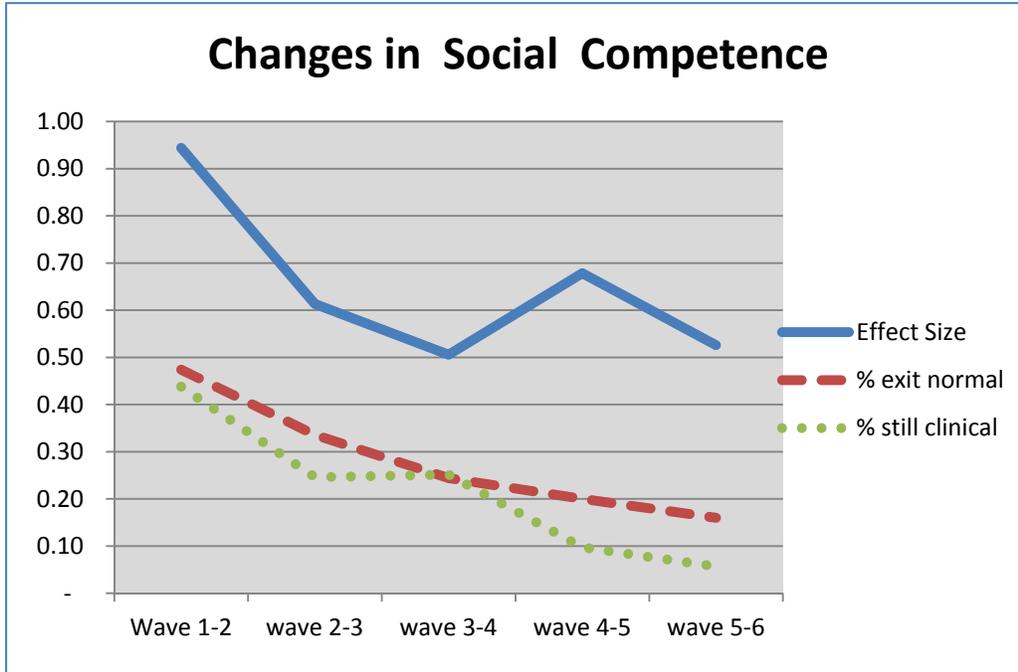
The solid line at the top indicates there was a strong treatment effect across four years of service with the largest improvement occurring the first year of treatment. 47% of the children tested in the wave 1-2 pair crossed the at risk threshold of 50 on this test. This is indicated by the heavy dashed line.

In a complementary fashion, the percentage of children who remained within or deteriorated into the at risk range fell significantly across four years of service, ending with 8% of the original cohort, still functioning very poorly at wave 5-6.

Some children did deteriorate over the four years; they represent 11% of the total group at risk on this measure. These new at risk cases, plus the original at risk cohort, total 350 children. The total percentage of the 350 children who remained in the clinical range equaled 33%. Conversely 67% of the at risk cohort experienced a positive outcome. The CGAS results at the end of four years are displayed in graph #4.1.

GRAPH #4.1   4-YEAR ACCUMULATED OUTCOMES (CGAS)

**4- year accumulated outcomes
Children's Global Assessment Scale**

23%

33%

67%

- % became normal
- % still clinical
- % of drop outs
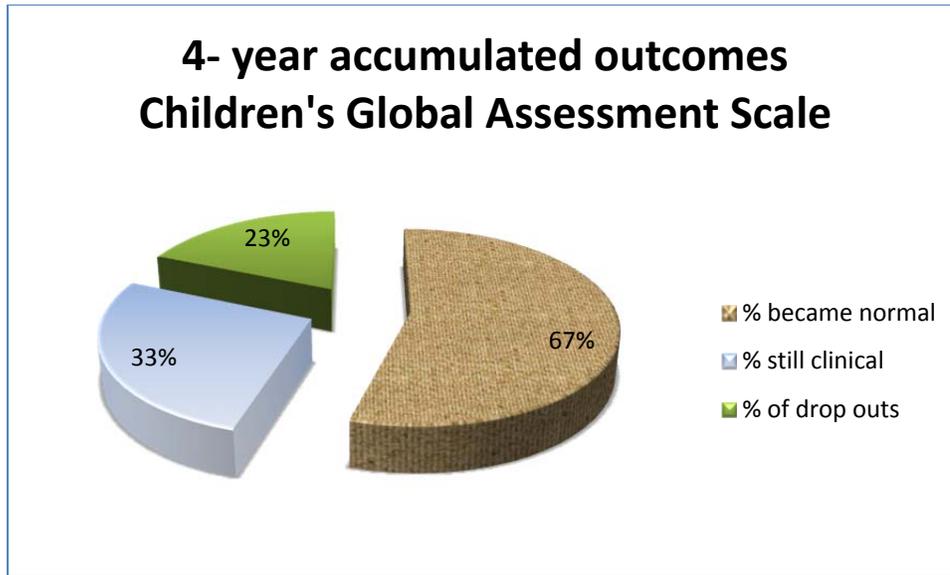
Two thirds of the children in the clinical range did not complete the four years of testing. The research drop-outs represent 23% of all children in the at risk group. There is no data to show what would have happened if the children remained in treatment for a longer period of time.

Children with neuro-behavioural impairment, such as FASD, autism and dual diagnosis are significantly more impaired than other children (neurologically impaired = 36.56, other children = 40.64, F = 6.777, df = 409, sig = .01). The difference in subsequent waves of testing disappears for the two groups of clients.

We did not identify any significant relationships between the changes in the CGAS scores and other variables, such as the agency, gender of the youth, history of abuse and history of parental criminality or mental illness.

Level of Assistance

The full title of this instrument is the Level of Assistance Required for Basic Functioning. It is a measure of adult support required to ensure completion in 22 daily living requirements for the child to participate in life. As children get older, they become more and more independent; although, some daily living skills take longer to acquire full independence. Daily living skills that may require daily supervision by a care provider until adolescence include:

- Manages own behaviour without instruction from others
- Follows rules, guidelines and routines of activities
- Cleaning up after an activity
- Going from place to place in the community
- Using Community Recreation Facilities

The full set of daily living requirements include those pertaining to basic self care (eating, grooming, bathing, toileting), self-regulation (following rules and routines, coping with negatives) communication (providing positive and negative feedback, making choices), social interaction (joining in a conversation, offering and accepting assistance) and participating in the community (going to and from places, attending community recreation, going to school).

Children who require more support from adults to perform these tasks require a lot of the parent's time. The higher the number the harder the job is for the parent. This is especially true for children from 10 years of age and up. Each skill is rated on a five point scale from independent, worth 0 points, to total care, worth 4 points.
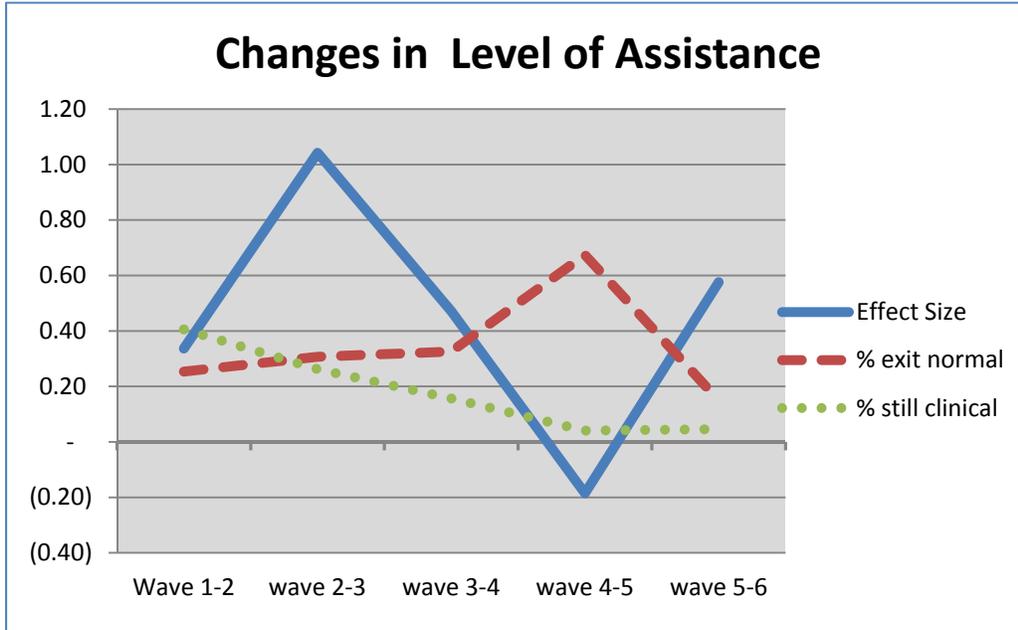
Validity of the Level of Assistance scale:

There were significant differences in means between different groups of children item by item and in the total score:

1. autism compared with mild ID + serious behaviour problems (dual diagnosis)
    a. significant differences in 12 of 22 domains
    b. t-score of difference in means for total support = 2.499, sig = .013
    c. autism average score = 56%, dual diagnosis = 46%
2. FASD compared with emotional and behaviour problems
    a. significant differences in 18 of 22 domains
    b. t-score of difference in means for total support = 2.971, sig = .003
    c. FASD average score = 38%, EB/BD = 28%
3. Placement in treatment foster care vs. group care
    a. significant differences in 22 of 22 domains
    b. F-ration (ANOVA) of difference in means for total support = 18.319, sig = .000
    c. TFC average score = 33%, GC = 44%

These differences demonstrated the discriminate validity of the LOA. Convergent validity was demonstrated by the correlation between the LOA and the CGAS (r = .751, sig = .000).

Graph #5 displays what happened to the children who were at risk on the LOA measure.

Graph #5 Changes in Level of Assistance

## Changes in Level of Assistance

The total number of children who were tested multiple times using the level of assistance scale was 389. The average score was 27.5 and the standard deviation was 12.9. This is typical of a population of children with diagnosed psychiatric disorders and serious emotional and behavioural needs. 68% of clients required more assistance than the threshold set for at risk youth (LOA >20).

Four of the five pairs of tests show a strong positive treatment effect for the LOA, especially at the end of the first year of service. The at risk group deteriorated on this measure after 3.25 years of service, showing a medium negative treatment effect (Cohen's d = -.23). The average score at the end of test pair wave 4-5 = 34.13 (n = 46), compared to 31.87 for the same child nine months earlier at the end of test pair, wave 3-4. After three years of service, the at risk cohort had been reduced by the number of children who improved to the normal range or were discharged. This results in a distillation effect; the average score at the start of each pair of later waves is worse than the score for the whole group who were at risk in the beginning. The at risk cohort started to improve with a strong positive effect at wave 5-6 (mean = 32.5, n = 20).

The dashed line indicates that a high percentage of children were exiting at each wave as they tested in the normal range, providing further evidence of clinical improvement. The final position is displayed in graph #5.1

G_RAPH_ #5.1   4-_YEAR ACCUMULATED OUTCOMES_ (LOA)
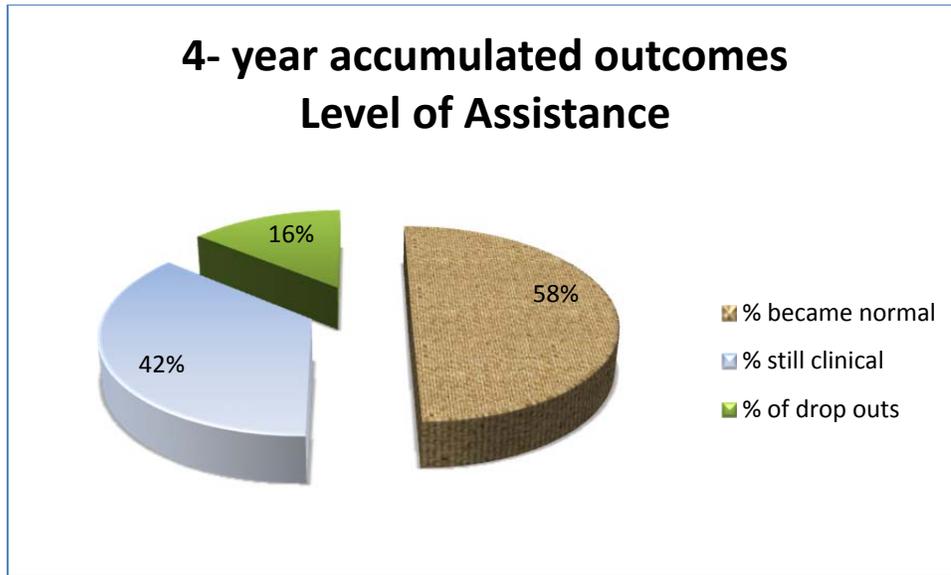
Graph #5.1 illustrates that more than a third of the children in the clinical range did not complete the four years of testing. A substantial proportion (58%) of the at risk group were in the normal range.

The Level of Adversity

The level of adversity instrument, also known as the Sociodemographic Checklist, is a 15 item checklist of serious risk factors in the social background of the children. The actual number of risk factors is a measure of lifetime adversity. A review of the literature for this checklist revealed that having four or more of any single item in the background means that the child had better than a 66% probability of being unable to function independently by age 18 years (Werner, 1989 and 1992[12]). Although each single stressor (such as the experience of sexual abuse) has consequences to the child's mental health and social functioning, it is the combination of stressors which overwhelms the child's ability to adapt and develop into a functioning young adult (Bronfenbrenner, 1979[13]).

The Sociodemographic Checklist was compiled from those risk factors that met the following criteria:

{a}   The presence of the stressor or condition is universally toxic to all children regardless of ethnic background, age or gender.

---

[12] Werner, E. (1989), "High Risk Children in Young Adulthood: A longitudinal study from birth to 32 years", *American Journal of Orthopsychiatry*, 59(1), 72-81

Werner, Emmy, E & Smith, Ruth, S. (1992), *Overcoming the Odds: high risk children from birth to adulthood*, Ithica, N.Y., Cornell University Press, 280 pages

[13] Bronfenbrenner, U., 1979, *The Ecology of Human Development: Experiments by Nature and Design*, Harvard University Press, Cambridge, Mass

{b}   The destructiveness of these factors remain active throughout the life span of childhood in the most common disorders.

{c}   There is a fairly high prevalence for each of these risk factors (between 1% and 5%), so that referrals from CAS agencies would likely display at least one of them.

This list does not include risk factors that are active during the early childhood years, e.g., difficult temperament, birth weight under 2500 grams, low Apgar score, duration of breastfeeding, maternal smoking during pregnancy and childhood lead intake. These risk factors are of great importance in primary prevention, but have little discriminating power in assessing the clinical significance of a referral for clinical services (Dishion, 1995[14]; Fergusson, 1995; Kazdin, 1995; O=Donnell, 1995; Rutter, 1985; Andrews, 1992).

Other examples of unstable risk markers include age of onset (Tolan, 1995), single parent, family income level and urban residence (Dishion, 1995; Guerra, 1995; Vaden-Kiernan, 1995) and family dysfunction and reading problems (Rutter, 1980). Physical and sexual abuse and learning problems starting at elementary school are also persistent and reliable risk markers (Tremblay, 1995[15]).

Sameroff (1995[16]) reviewed 14 longitudinal studies as well as his own research and found a continuity of risk factors, such as:

- abusive and unemotional parent-child interactions
- family conflict and distorted levels of communication
- lack of social support
- stressful life events
- poverty
- disorganized neighbourhoods
- inadequate school systems

These risk factors are much stronger than the continuity of any behavioural domain within the child. Given the transactional nature of the risk mechanism between the child and his social context, the overwhelming continuity of risk factors drives the correlations between early adaptive failure and adult psychopathology.

OARTY's research includes data on 2,081 unique children served over a 15 year period; these children resided in 82 separate OARTY agencies. This is much more extensive than the longitudinal test data. The graph of adversity is as follows:

---

[14] Dishion, T.J. & Andrews,D.W. (1995), "Preventing Escalation in Problem Behaviors With high-Risk Young Adolescents: Immediate and 1 year Outcomes", *Journal of Clinical and Consulting Psychology*, 63, (4) 538-548

[15] Tremblay, R.E., Masse, B., Perron, D., LeBlanc, M. (1992), "Early Disruptive Behavior, poor School Achievement, Delinquent behavior and Delinquent Personality: Longitudinal Analysis", *Journal of Consulting and Clinical Psychology*, 60 (1), 64-72

[16] Sameroff, Arnold (1995), "General Systems Theories and Developmental Psychopathology", in *Developmental Psychopathology*, Dante Cicchetti & Donald Cohen (eds), J. Wiley, New York, 659-695

Graph #6 Adversity in the Child's Background

Types of Adversity

Category	Value (%)
suicide fam membr	~4%
Sexually assaultive Fam Mbr	~6%
aboriginal	~8%
Fam Mbr victim/rape	~11%
child on drugs	~12%
born to teen mom	~14%
organic brain damage	~15%
Fam Mbr mentally ill	~15%
Violent Fam Mbr	~15%
Fam Mbr has intellectual deficit	~16%
Fam Mbr jailed	~22%
Sex Abuse	~26%
Fam Mbr addicted	~37%
Physical Abuse	~41%
Deprivation	~41%
chronic school failure	~55%
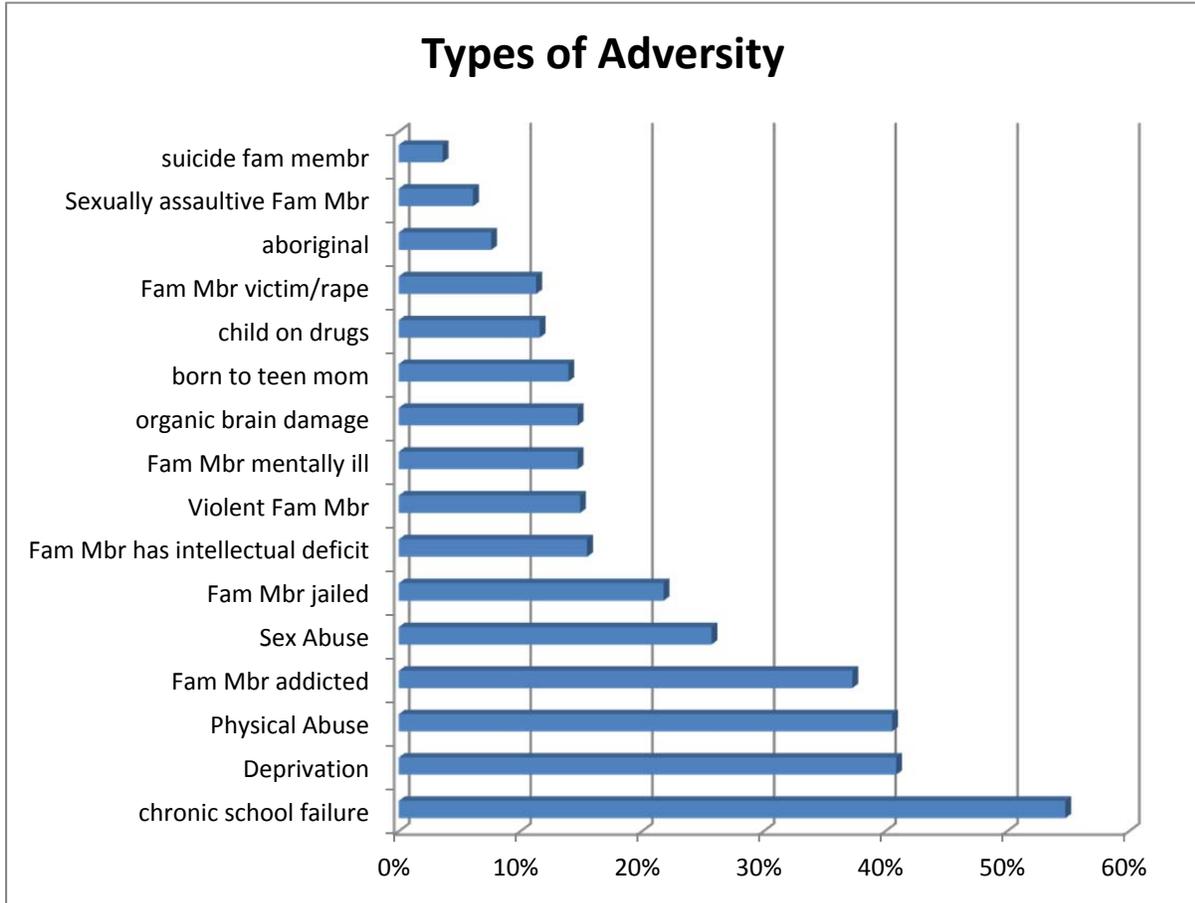
The adversity data is taken from four separate surveys dating from 2002, based on random selection each time. Each type of adversity has been found to be a potent risk factor predicting lifelong problems. As noted in the research cited above, the number of different types of adversity is the best predictor of poor adjustment across the lifespan. The average number of risk factors is four; this threshold was established by Emily Werner (ibid) and linked to poor lifespan outcomes for 66% of subjects.

Graph #7 displays the primary clinical issue that is the focus of treatment.

Graph #7 The Primary Clinical Issue Driving Service Provision

Primary issue for treatment or service
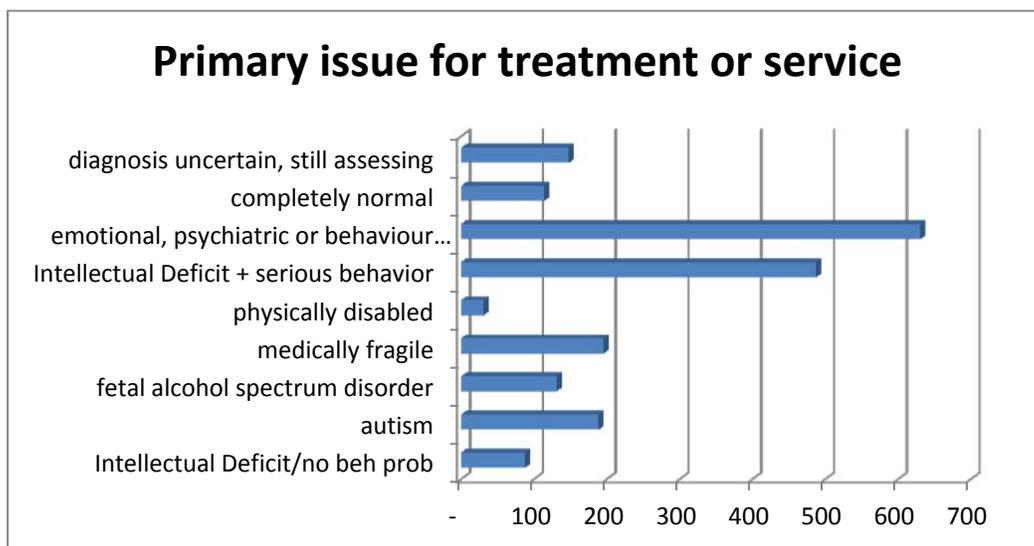
This data shows that 31% of the children served by OARTY members are in treatment due to emotional and behavioural disorders. The second largest group (24%) are children with intellectual deficit and serious behavioural problems, referred to as dual diagnosis. The children considered as dual diagnosis have a complex neuro-developmental disorder. 56% of the children served have complex neuro-developmental disorders.

We found that the number of types of adverse conditions is significantly correlated with the primary reason for service, as noted in the table below:

Table 1	number of adverse conditions		
Primary reason for service	N	Mean	Std. Deviation
Intellectual Deficit/no beh prob	88	2.91	2.14
autism	189	3.72	2.19
fetal alcohol spectrum disorder	131	4.34	3.25
medically fragile	196	3.69	1.92
physically disabled	31	2.42	1.75
Intellectual Deficit + serious behavior	488	4.90	2.97
emotional, psychiatric or behaviour problems	631	4.11	2.91
completely normal	114	2.11	2.45
diagnosis uncertain, still assessing	148	2.99	1.92
Total	2,016	3.97	2.78

The relationship was assessed by Analysis of Variance (F = 20.75, df = 2,015, sig = .000). Children with dual diagnosis have the highest number (4.9) of adverse conditions in their background followed by children with FASD who have 4.34 adverse conditions. It is significant to note the children who are regarded as normal have the lowest number (2.11) of adverse conditions, as predicted by the studies referenced.

The number of adverse conditions predicts the placement of children who require staff operated programs. The mean for staff operated programs is 4.03 and the mean for family operated services is 3.75 (F = 4.72, df = 2,080, sig = .03).

Remarks on Adversity

The degree of adversity in the background of the children is the dominant risk confronting the clients of OARTY member agencies. The degree of adversity and the individual adverse conditions are significantly correlated with the primary reason for service.

In general, adversity that is indicative of family pathology is much more prevalent with children whose reason for service is

- Emotional and Behavioural Disorders
- Dual Diagnosis
- Fetal Alcohol Spectrum Disorders

Clients with a high degree of adversity, especially related to family pathology, are likely to have very poor attachment experiences.[17] As these children grow up under adverse circumstances they create dysfunctional attitudes about themselves and others, hold negative memories and emotions, have poor problem solving skills, poor affect regulation, poor relationship patterns with peers and adults, emotionally coercive interactions, avoidance of social reciprocity, boundary problems, lack of trust, lack of empathy and inaccurate perceptions about their social environment. This set of attachment related patterns have been referred to as the inner working model[ii].

When attachment related clinical issues combine with intellectual limitation or psychiatric illness, the prognosis for improvement and a positive life span outcome is very poor. For this reason, we strongly recommend that OARTY member agencies provide all services and interventions through the meta-model of attachment.

Parental Bonding Instrument Caring Scale

The Parental Bonding Instrument (PBI) has been used extensively in cross sectional and longitudinal studies, globally. It is a norm-referenced test. The PBI measures how securely the

---

[17] Crittenden, P.M. (1985), "Social networks, quality of parenting, and child development", *Child Development*, 56,1299-313.

Crittenden, P. (2000), "A Dynamic-Maturational Approach to Continuity and Change in Pattern of Attachment", in *The Organization of Attachment Relationships*, Patricia Crittenden & Angelika Claussen (eds), Cambridge University Press: Cambridge, UK, pages 343-357.

child is bonded to a target person identified on the test. The target can include foster parent, child care worker or family of origin. The security of attachment is a critical protector of the life span outcome of many social and demographic risk factors and the prognosis for change in children with symptoms of emotional disturbance or behaviour problems (Mallincroft, 1992[18]; Werner, 1989 & 1992, Fonagy, 1993[19], Rutter, 1982; Rutter, 1995[20]; Costello & Angold, 1995[21]).

The scales were originally developed to measure the parent's ability to bond to their children based on how well they bonded to their own parents. There are four types of parental bonds or intergenerational relationships which are identified by the teenager or the adult reflecting on his or her childhood.

Pattern of Scores on PBI	Clinical Interpretation
high care-low over protection	optimal bonding
low care-low overprotection	absent or weak bonding
high care-high overprotection	affectionate constraint
low care-high overprotection	affectionless control

The PBI has been used with adolescents as a means for them to rate the quality of bonding with their parents (Rey, 1990[22]).

Canetti's study[23] on the internal reliability (using Cronbach's alpha coefficient) was computed. All four scales were found to be reliable with the following alpha values: maternal care = 0.75, maternal control = 0.82, paternal care = 0.80 and paternal control = 0.83.

Canatti's conclusion is as follows:

"The strong connection between parental bonding and psychopathology found in the present study provides support for Bowlby's attachment theory. Adolescents who reported high care and low control (optimal bonding) reported less psychiatric symptoms and distress. They had a positive feeling of well being and felt supported by family and friends. On the opposite end, the combination of low care and high control

[18] Mallinckroft, B. (1992), "Childhood Emotional Bonds With Parents, Development of Adult Social Competencies, and Availability of Social Support", *Journal of Clinical Psychology*, 39 (4), 453-461

[19] Fonagy, P., Steele, M., Steele, H., Higgitt, A. & Target, M. (1993), "The Emanuel Miller Memorial Lecture 1992 The Theory and Practice of Resilience", *Journal of Child Psychology and Psychiatry*, 35(2), 231-257

[20] Rutter, Michael (1995), "Developmental Psychopathology: concepts and prospects", in *Developmental Psychopathology*, Dante Cicchetti & Donald Cohen (eds), J. Wiley, New York, 209-237

[21] Costello, E. & Angold, A. (1995), "Developmental Epidemiology", in *Developmental Psychopathology: theory and methods*, D. Cicchetti & D. Cohen (eds), John Wiley & Sons, New York, 23-56

[22] Rey, J. (1990), "Quality of Perceived Parenting in Oppositional and Conduct Disordered Adolescents", *Journal of the American Academy of Child & Adolescent Psychiatry*, 29 (3), 382-385

[23] Canetti, L., Bacher, E. Galili-Weisstub, E., Kaplan De-Nour, A. & Shalev, A. (1997), "Parental Bonding and Mental Health in Adolescence", *Adolescence*, 32 (126), Summer,381-394

(affectionless control) gave rise to psychological symptomatology and a lesser-feeling of well being. These individuals also experience less support from their surroundings".

There are three scales produced by the PBI:

(1)   a measure of affection: i.e., the teenagers perception about how much affection the parent figure displayed;

(2)   a measure of neglect: i.e., the teenagers perception about how much he/she was left without sufficient structure in the home;

(3)   a measure of unfairness: i.e., the teenagers perception about how unfair the parent figure was in the setting rules and expectations.

Graph #8 displays the outcomes for children who are at risk on the affection scale

Graph #8   Changes in Affection
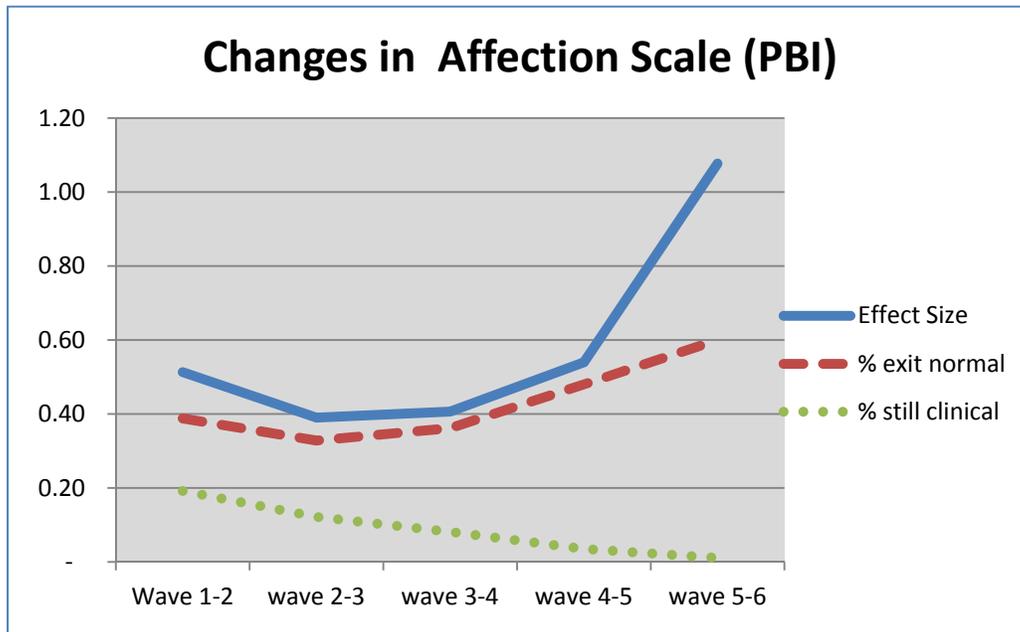
The total number of children who were tested multiple times on the affection scale was 366. The average score was 45 and the standard deviation was 14.62. The large standard deviation indicates there is a large proportion of children at the top and bottom ends of the distribution at the time of the first test.

Most of the clients who were tested initially (60%) were below the at risk threshold (affection t-score <45). The average score for this at risk group was 30.7. Less than 3% of the children in society have a score measuring affection that is this low or lower. The at risk cohort feels that no one likes them or shows them any affection.

The outcomes at the subsequent waves of testing produced a strong treatment effect on each occasion. The largest change occurred at the end of four years of treatment (Cohen's d = 1.08). In the final analysis, 55% of the at risk cohort crossed the threshold that indicates they feel their caregivers like them.

Some children refused to take a follow-up PBI test and others were discharged before they improved. The drop-outs represent 13% of the original cohort. It is important to note that the effect size on this variable increases significantly over time. In other words, it takes time for the children to allow their inner working model to change.

Graph #8.1 provides the accumulated outcomes across 4 years of service with 187 children who were followed after the 1st test.
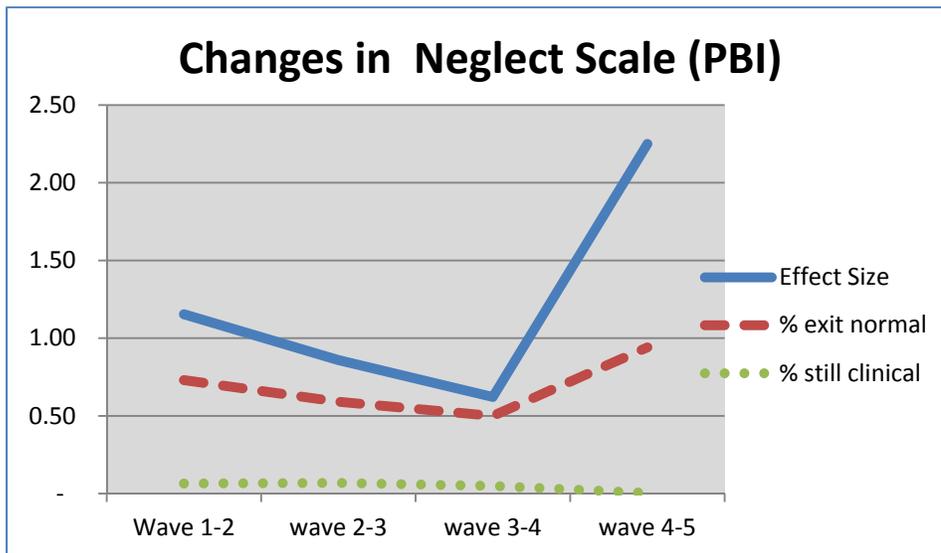
GRAPH #8.1   4-YEAR ACCUMULATED OUTCOMES (AFFECTION SCALE)

4- year accumulated outcomes Affection Scale (PBI)

13%

55%

45%

- % became normal
- % still clinical
- % of drop outs
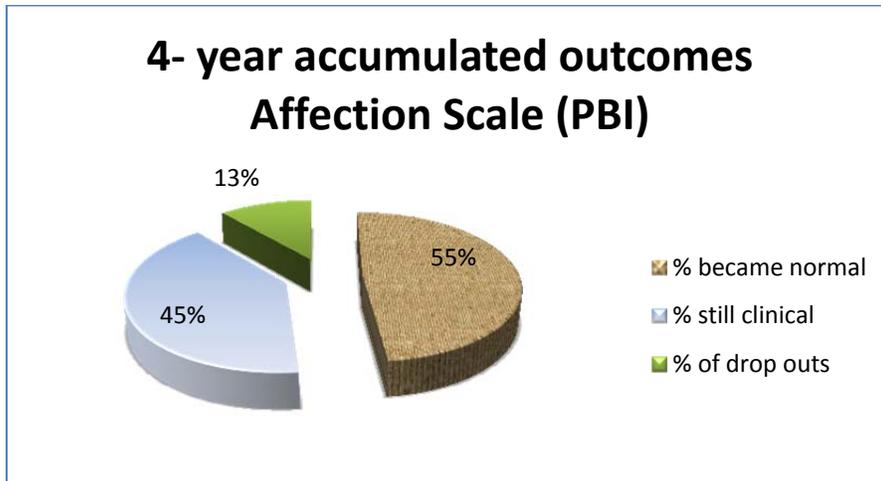
A large group of children, 45%, were left in the clinical range, indicating that they felt a deep sense that they were not liked. A third of this group did not complete the four years of testing.

Graph #9 displays the outcomes for children at risk on the neglect scale.

Graph #9   Changes in the Neglect Scale

Changes in Neglect Scale (PBI)

Wave 1-2   wave 2-3   wave 3-4   wave 4-5

—— Effect Size
--- % exit normal
··· % still clinical

The total number of children who were tested multiple times on the Over Protective Scale was 362. The average score was 58 and the standard deviation was 13.8. The large standard deviation means there is a large proportion of children at the top and bottom ends of the distribution at the time of the first test.

The Over Protective Scale measures the child's perception of how much supervision, control and protection his/her parents (foster parents or child and youth worker) provided. If they provided too little, i.e., Over Protective Scale <45, then the instrument implies the child feels he/she was neglected and unsafe. Using this criteria, 28% of the children felt they were neglected at wave #1. The average score for this at risk cohort was 40.4. In comparison with the normative reference group, the OARTY at risk group was below the 16th percentile.

The outcomes at the subsequent waves of testing produced a very strong treatment effect on each occasion. The largest change occurred at the end of 3.25 years of treatment (Cohen's d = 2.25). During the wave 4-5 test period, 94% of the at risk cohort had crossed the threshold into a score that indicates they feel safe. After four years of service, 20% of the at risk group were feeling unsafe in relation to the closest person in their life. 80% of those at risk on this measure returned to the normal range during the study. This is displayed in graph #9.1
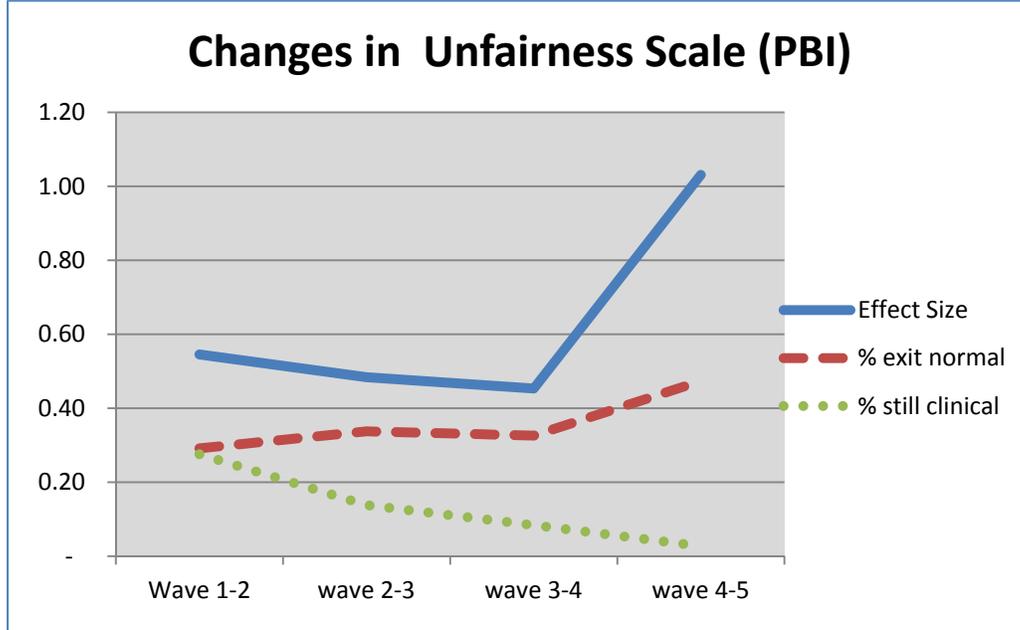
Some children refused to take a follow-up PBI test and others were discharged before they improved. The drop-outs represent 12% of the at risk group. It is important to note that the effect size on this variable increases significantly over time.

GRAPH #9.1   4-YEAR ACCUMULATED OUTCOMES (NEGLECT SCALE)
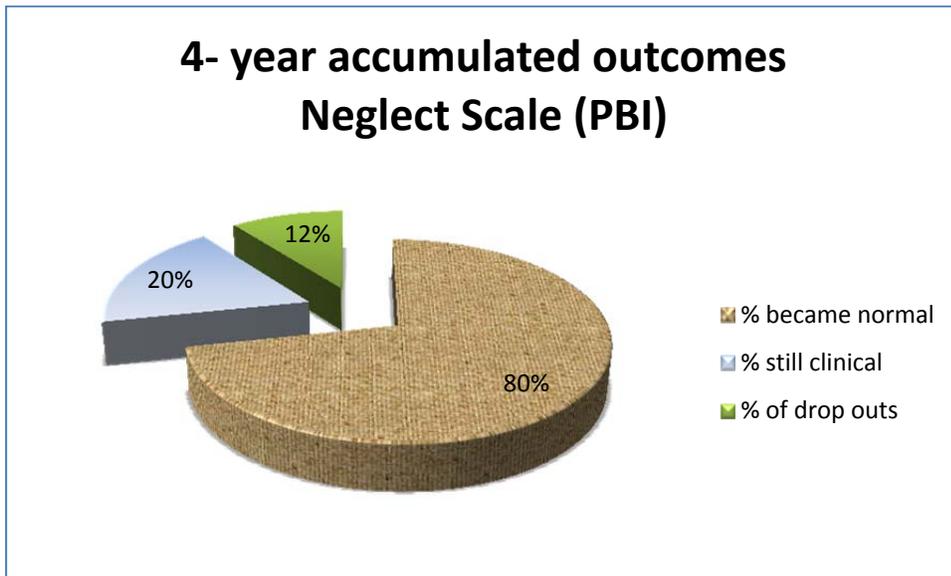
Graph #10 displays the outcomes for children at risk on the unfairness scale.

Graph #10   Changes in the Unfairness Scale

Changes in Unfairness Scale (PBI)

Legend:
- Effect Size (solid blue line)
- % exit normal (red dashed line)
- % still clinical (green dotted line)

X-axis: Wave 1-2, wave 2-3, wave 3-4, wave 4-5
Y-axis: -, 0.20, 0.40, 0.60, 0.80, 1.00, 1.20

The Over Protective Scale measures the child's perception of how much supervision, control and protection his/her parents (foster parents or child and youth worker) provided. If they provided too much structure and control, i.e., Over Protective Scale >60, then the instrument implies the child feels he/she was treated unfairly. Being "over-controlled" leaves the child with a feeling that his/her parents did not understand him/her. According to Fonagy (ibid), this undermines the quality of attachment.

Using this criteria, 52% of the children felt that they were treated unfairly at wave #1. The average score for this at risk cohort was 72.9. Less than 2% of teenagers in the normative reference group produced a score equal to or higher than the OARTY sample.

The outcomes at the subsequent waves of testing produced a very strong treatment effect on each occasion. The largest change occurred at the end of 3.25 years of treatment (Cohen's d = 1.03). During the wave 4-5 test period, 47% of the at risk cohort had crossed the threshold that indicates they feel their needs were understood and they were fairly treated. After four years of service, 58% of the at risk group were feeling unfairly treated by the closest person in their life. This is the highest rate of poor responders across all of the tests.

Some children refused to take a follow-up PBI test and others were discharged before they improved. The drop-outs represent 14% of the at risk group. We are uncertain of the outcome if they remained. It is important to note the effect size on this variable increases significantly over time.

Conclusions

The measures used in this longitudinal study are directly related to the most potent factors predicting life span outcomes for children with serious emotional and behavioural problems. The instruments identify children who are at risk of developing enduring mental illness or an inability to function as independent adults. The criteria for determining who is at risk were based on other longitudinal studies referenced in this report. The at risk threshold was set very high to minimize inflating the number of children truly at risk.

The risk factors measured were:

- hyperactivity, impulsiveness, inattention and emotionality
- broad psychiatric symptoms mapped to the DSM-IV R
- attitudes, feelings and behaviour
- ability to perform the social roles of student, family member and community member successfully
- degree of adult support required to perform 22 daily living tasks, such as eating and crossing the street, ranging from totally independent to fully dependent on the adult to do everything
- degree of adversity in the background before admission
- degree of affection felt by the child in relation to his/her closest caregivers
- degree of neglect felt by the child in relation to his/her closest caregivers
- degree of unfairness and over-control felt by the child in relation to his/her closest caregivers

The degree of adversity does not change over time, but the risk factors affect the life span outcomes. In all measures, the data shows a strong to very strong positive treatment effect. The majority of children who started the study in the at risk zone crossed over to the range of scores indicative of a normal child.

A strong positive outcome for this cohort of children is remarkable given the amount of adversity in the child's background. This much adversity is associated with poor outcomes. The simplest explanation of this paradox is stated below.

Although the adversities experienced in childhood cannot be changed, i.e., physical abuse, the appropriate treatment can change the child's perception of precursory events. This is also referred to as changing the inner working model, an application of attachment theory to residential care.

Discussion

There are two limitations with this longitudinal study:

(1) The data in the study is restricted to the risk factors listed in the conclusions; contextual data, such as what interventions were provided and in depth information about the clinical profile, was not available. This makes it impossible to explain why the effect sizes are so strong.

(2)   The five agencies who participated in the study have a fairly common target group of children with serious psychiatric, emotional and behavioural problems; other children served by the OARTY members, such as children with autism, complex medical needs and intellectual deficits, are not represented.

The Next Generation LS is designed to correct these limitations by collecting the measurement data within a relational database in which the outcomes are linked to comprehensive contextual data. Secondly, the Next Generation LS has expanded the number of measures to include areas that are the focus of treatment for children with significant developmental and medical needs.

The data gathered in this study suggests the OARTY member agencies that have adopted a theoretical framework, i.e., attachment, and employ a series of sociometric and psychometric pre-post measurements, produce excellent outcomes. The research literature cited in this report supports this model of best practice in the context of evidence based treatment.

The OARTY agencies that participated in this study have demonstrated good outcomes for the clients served. It is imperative for clients to remain in treatment over the prescribed period found in this study in order to achieve their potential positive outcome.

Endnotes

[i] The effect size of treatment is a standardized unit for measuring the amount of change between two scores. The formula for measuring treatment effect (Cohen's D) is

(mean 2 – mean 1)

    Divided by

pooled standard deviation of both wave 1 and wave 2.

The formula for computing the pooled standard deviation is

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2}},$$

[ii] Attachment is not something that exists within the child; it is a relationship between two people. The theoretical definition of attachment (security and comfort in a relationship providing the personal confidence to engage in other activities) is quite distinct from the inner working model defined above. The inner working model is a process of the mind. According to both the theory and primary research, the child creates a separate and distinct inner working model for each attachment relationship (Fonagy, 1993). If you encounter the child as a young adolescent, you might very well see a number of attachment relationships: with the mother and father of origin, a foster parent and an adoptive parent. The quality of love, safety, comfort and stability of the relationship might vary significantly across the four attachment figures. The child will encode significant components of each attachment figure in his mind; according to Fonagy, there be four separate inner working models; although this is just a metaphor, the child does have only one mind. The child will have four mental representations for storing the memories and feelings he associates with each parent; his behavioural strategies for keeping himself safe and making contact with his current relationship – the adoptive parent - may be a constantly evolving combination of a variety of strategies learned from all four parent figures. This observation was demonstrated by Victor Groze (1996) who tested both the current relationship (in terms of love, security, comfort) and the inner working model (in terms of the attachment behaviours) and found that the two concepts were independent on testing at time one and relative to changes over time. The majority of special needs children (Groze, 1996; Goldberg, 2000; Crittenden, 2000) can form secure attachments, but they often retain dysfunctional aspects of their inner working model. Crittenden (2000) also points out that no matter what happens, the attachment relationship and the inner working model change continuously over time with each maturational stage and with significant disruption and periods of stress.

Please address all inquires to the

Ontario Association of Residences Treating Youth (OARTY)

550 Alden Road, Suite 210
Markham, ON L3R 6A8
Phone: 905.475.KIDS (5437)

Fax: 905.475.5430

info@oarty.net

www.oarty.net