Outcome Measures Pilot: SDQ and CASII

Report for Minnesota Mental Health Action Group (MMHAG)

Children's Mental Health Division
Minnesota Department of Human Services

July 2008

Outcome Measures Pilot: SDQ and CASII

Executive Summary

As part of an effort to improve delivery of services to children (and adults) with mental illness through the use of a coordinated statewide evaluation system across public and private payers and all providers, the Minnesota Mental Health Action Group (MMHAG) reviewed outcome measures for children's mental health. The goal of this review was to identify instruments that would provide useful, cost-efficient data and a common evaluation framework used statewide without increasing burden on providers or consumers.

Two instruments were identified to measure outcomes: The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) and the Child and Adolescent Service Intensity Inventory (CASII; American Academy of Child and Adolescent Psychiatry, 2005). These instruments were chosen for their ability to assess effectiveness of services and improved outcomes. The CASII is an 8-item instrument, completed by mental health providers and case managers, designed to objectively determine the service needs of children and adolescents. The SDQ is a behavioral screening questionnaire completed by parents, teachers, and/or self (children/adolescents ages 11 and older).

The SDQ and CASII were piloted by mental health providers throughout Minnesota from March 2006 through April 2008. In all, thirty counties, agencies, and collaboratives provided data on 4,239 children and adolescents. The goal of this pilot was to assess if these instruments demonstrated similar validity and reliability to comparable standardized instruments, were able to detect changes in mental symptoms and functioning, were related to providers' service recommendations, and were viewed by providers as clinically useful and appropriate for their populations.

Pilot highlights are as follows:

- CASII Scores ranged across all seven levels of care and decreased significantly from initial administration to 6-month follow-up and from 6-month follow-up to 12-month follow-up, suggesting that the instrument discriminates between services needs and assesses changes in service needs and functioning over time.

- On average, children/adolescents that were identified by the CASII as having higher services need were recommended more hours of services and more hours of services by providers.

- CASII test-retest validity was similar to previous studies and comparable instruments.

- Parent, Teacher, and Self SDQ scores decreased significantly from initial administration to 6-month follow-up and from 6-month follow-up to 12-month follow-up, suggesting that the instrument assesses changes in mental health symptoms over time.

- Inter-rater reliability between SDQ respondents was similar to other omnibus behavioral checklists.

- On average, cases with higher parent, teacher, or child-rated mental health symptom severity (SDQ scores) were also recommended more services by providers.

- All outcome measures (CASII, Parent SDQ, Teacher SDQ, and Self SDQ) related significantly to each other. CASII scores related significantly to all types of providers' service recommendations. Parent and Teacher SDQ's related most strongly to the number and hours of therapy services (individual, family, and group psychotherapy and skills). Self SDQ scores related most strongly to the number of case management hours recommended.

- Clinician satisfaction surveys indicated that the majority of providers agreed that these measures adequately measured emotional distress as well as strengths and assets; measured what they proposed to measure; were appropriate for children of most ages and in most settings; and required a reasonable amount of time to administer. Most providers indicated that the SDQ was helpful in treatment planning and was positively received by clients and families.

The following report details the results of the Children's Mental Health Outcome Measures Pilot.

TABLE OF CONTENTS

Page

I. Background Information 1

II. Composition of Data 1
.....A. Counties, Agencies, and Collaboratives Submitting Data 1
.....B. Demographics 1

III. Overview of CASII (Child and Adolescent Service Intensity Instrument) Data 2
.....A. Description 2
.....B. CASII: Average Scores 2
.....C. CASII: Level of Care Recommendations 3
.....D. Change in CASII Levels 3
.....E. CASII: Comparisons between Administration Times 3
.....F. CASII: Demographic Differences 3
.....G. CASII: Comparisons with Providers' Service Recommendations 4
.....H. Summary of CASII Findings 5

IV. Overview of SDQ (Strengths and Difficulties Questionnaire) Data 5
.....A. Description 6
.....B. SDQ: Average Scores 6
.....C. SDQ Probability Levels 7
.....D. Changes in SDQ Total Score 7
.....E. SDQ: Comparisons between Respondents and Administration Times 8
.....F. SDQ: Demographic Differences 8
.....G. SDQ: Comparisons with Providers' Service Recommendations 9
.....H. Summary of SDQ Findings 11

V. Overview of Recommended Mental Health Services 11
.....A. Description 11
.....B. Services Recommended Overall 11
.....C. Services Recommended: Demographic Differences 12
.....D. Summary of Services Recommended 13

VI. Relations between CASII, SDQ, and Services 13
.....A. Correlations between CASII, SDQ, and Services 13

VI. Clinician Satisfaction Survey 14
.....A. CASII: Clinician Satisfaction 14
.....B. SDQ: Clinician Satisfaction 15
.....C. Summary of Clinician Satisfaction 15

VII. Summary: Review of Results 16

Appendix A. Additional Results 17

I. Background Information:

As part of an effort to improve delivery of services to children (and adults) with mental illness through the use of a coordinated statewide evaluation system across public and private payers and all providers, the Minnesota Mental Health Action Group (MMHAG) reviewed outcome measures for children's mental health. The goal of this review was to identify instruments that would provide useful, cost-efficient data and a common evaluation framework used statewide without increasing burden on providers or consumers. Two instruments were identified to measure outcomes: The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) and the Child and Adolescent Service Intensity Inventory (CASII; American Academy of Child and Adolescent Psychiatry, 2005). These instruments were chosen for their ability to assess effectiveness of services and improved outcomes.

Many counties and agencies across Minnesota volunteered to pilot these instruments within their own mental health systems beginning in Spring 2006. They have submitted data including consumer demographics, SDQ and CASII scores, and recommended monthly services. Ideally, data was obtained at intake, follow-up assessments every 6 months, and at discharge. Additional data was collected through the School-Linked Mental Health Services Grant that was awarded to collaboratives throughout the state. Data collection procedures for the collaboratives were similar to that arranged for the pilot project and began in January 2007. The following is a report on the outcome measures data received from these counties, agencies, and collaboratives from the beginning of the pilot project through April 2008.

II. Composition of Data:

A. Counties, Agencies, and Collaboratives Submitting Data

1. Initial data (Time 1 data) for 4239 children and adolescents was submitted by 30 counties, agencies, or collaboratives. These participating groups submitted data on 1679 children and adolescents at 6-month follow-up (Time 2). Twelve-month follow-up data was available for 435 children/adolescents from 16 participating groups. Three counties/agencies sent 18-month follow-up data on 23 children/adolescents.

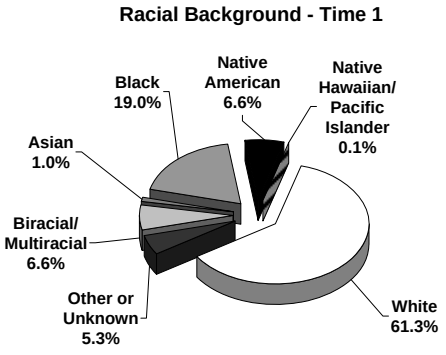
2. A list of the participating agencies, counties, and collaboratives and the number of measures received at each administration time (initial or follow-up data collection) is available in Appendix A (page 18 - 19).

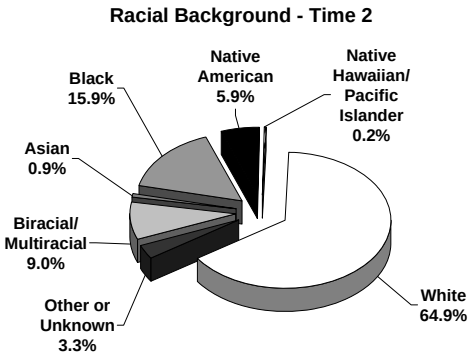
B. Demographics

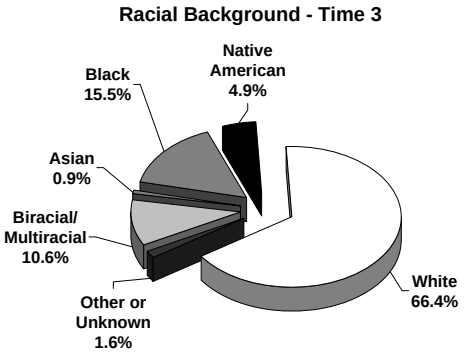
1. Gender: Approximately 61% to 62% of clients were male for the first three administration times and 38% to 39% were female. At Time 4, 73.9% of clients were male and 26.1% were female.

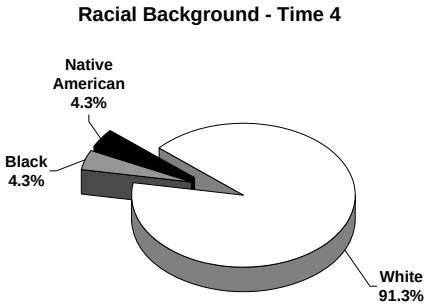
2. Age: The average age of children served at Time 1 was 12.1 (SD = 3.56). The average age at Time 2 was 12.4 (SD = 3.51). The average ages at Times 3 and 4 were 12.9 (SD = 3.26) and 13.3 (SD = 2.77) respectively.

3. Racial Background: The following racial backgrounds were represented:









- The percentage of children with Latino/a ethnic background was ranged from 4.3% to 9.7% across administration times.

- According to 2006 U.S. Census Bureau estimates, the racial background for the state of Minnesota is 3.5% Asian, 4.5% Black, 89.3% White, 1.2% Native American, 0.1% Native Hawaiian/Pacific Islander, and 1.5% Bi-racial or Multiracial. 3.8% identify themselves as having Latino ethnic background. A higher percentage of Native American, Black, Latino/a, and Biracial/Multiracial children were represented in this sample compared to statewide estimates and a lower percentage of White children (source: http://quickfacts .census.gov/qfd/states/27000.html, printed 1/09/2008).

- Additional information is available in Appendix A (pages 18-19).

III. Overview of CASII (Child and Adolescent Service Intensity Instrument) Data:

A. Description:

- The CASII (American Academy of Child and Adolescent Psychiatry, 2005) is an 8-item instrument designed to objectively determine the service needs of children and adolescents.

- Mental health providers rate clients on 8 dimensions: Risk of Harm, Functional Status, Comorbidity, Environmental Stress, Environmental Support, Resiliency, Child/Adolescent's Acceptance and Engagement in Treatment, and Parent's Acceptance and Engagement in Treatment.

- Each dimension has five levels that form scales from 1 (low or minimum problem area) to 5 (extreme problem area). Higher numbers indicate higher levels of problems or lower levels of strengths.

- In addition to ratings on each dimension, the CASII provides a Composite Score and Level of Care recommendation. The CASII's recommendations for level of care range from 0 (Basic Services for prevention and maintenance) to 6 (Secure, 24-hour services with psychiatric management).

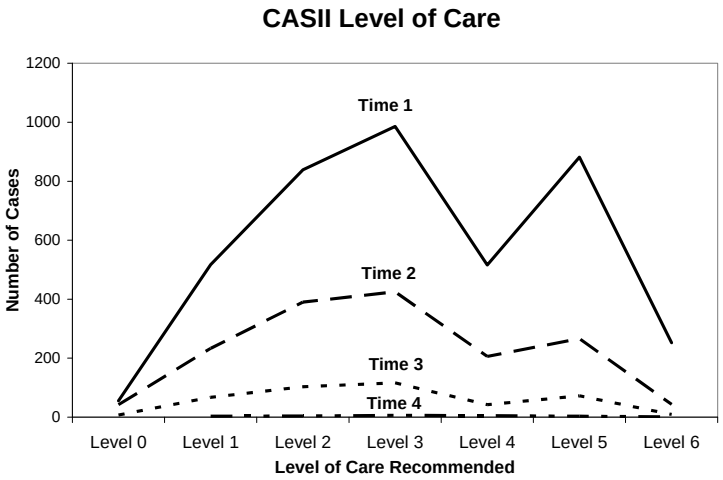
B. CASII: Average Scores

- Average CASII scores decreased from Time 1 (Mean = 18.3, SD = 4.56) to Time 2 (Mean = 17.5, SD = 4.25) and to Time 3 (Mean = 17.2, SD = 4.04). Average scores at Time 4 were based on a small sample of children and increased to 18.1 (SD = 4.17).

- Errors in calculating CASII Total Score were identified in 14.6% of CASII's submitted. Errors in calculating Level of Care were identified in 21.3% of CASII's.

- Due to concerns regarding errors in scoring, computer calculated CASII Total Score and Level of Care score are used for all analyses.

C. CASII: Level of Care Recommendations



D. Change in CASII Levels

- The CASII Level of Care decreased for approximately one-fifth to one-third of cases. Level of Care stayed the same for two-fifths to over two-thirds of cases. Level of Care increased in approximately one-tenth to one quarter of cases. Changes in Level of Care between specific administration times are listed in Appendix A (page 20).

E. CASII: Comparisons between Administration Times

1. Correlations between CASII administrations

- CASII Total Scores were moderately correlated between administration times (r = .50 to r = .65).
- CASII Levels of Care also were moderately correlated between administration times (r = .48 to r = .60).
- There were not enough CASII's at Time 4 to reliably assess the relation between this administration time and other administration times.

2. Differences in average scores

- Paired T-tests were used to compare average CASII Total Scores and average CASII Level of Care recommendations between administration times. CASII Total Score and Level of Care at Time 1 was significantly higher than at Time 2 or at Time 3. CASII Total Score at Time 2 was significantly higher than the Total Score at Time 3. Results are available in Appendix A (pages 20 - 21).

F. CASII: Demographic Differences.

- Gender: Providers endorsed higher CASII Total Scores for males at Time 2 and Time 3. Level of Care calculations were higher for males at Times 1, 2, and 3. More detailed statistics are available in Appendix A (pages 20 - 21).
- Age: CASII Total Score and Level of Care were significantly correlated with Age at Time 1 and Time 2. The degree of the relation was small (r's ranged between .11 to .18) and only evident with large sample sizes. However this indicates that there is a slight trend for older children/adolescents to have higher CASII Total Scores and Level of Care scores.
- Racial Background: Significant differences in CASII Total Score and Level of Care were identified at Time 1 and Time 2. Significant differences were also found at Time 3, although the sizes of the racial groups did not allow for reliable testing of which groups differed. For these analyses, children with

Native Hawaiian/Pacific Islander racial backgrounds were recoded into the Other category due to their small numbers in this sample. Significant racial differences are listed on next page.

- At Time 1, children in the "Other" category (Mean = 17.6, SD = 4.36) and children identified as White (Mean = 17.8, SD = 4.93) had lower CASII Total Scores than children with Native American (Mean = 19.4, SD = 5.19) and Black (Mean = 20.2, SD = 4.45) racial backgrounds.

- At Time 1, children in the "Other" category (Mean = 3.0, SD = 1.65) and children identified as White (Mean = 3.0, SD = 1.51) had lower CASII Level of Care scores than children with Black (Mean = 3.9, SD = 1.40) racial backgrounds.

- At Time 2, children in the "Other" category (Mean = 15.8, SD = 3.87) had lower CASII Total Scores than children with Asian (Mean = 18.4, SD = 4.57), Biracial (Mean = 18.3, SD = 3.64), Native American (Mean = 18.4, SD = 4.92), and Black (Mean = 19.6, SD = 4.07) racial backgrounds. Children identified as White (Mean 16.8, SD = 4.07) had lower scores than children with Native American or Black racial backgrounds.

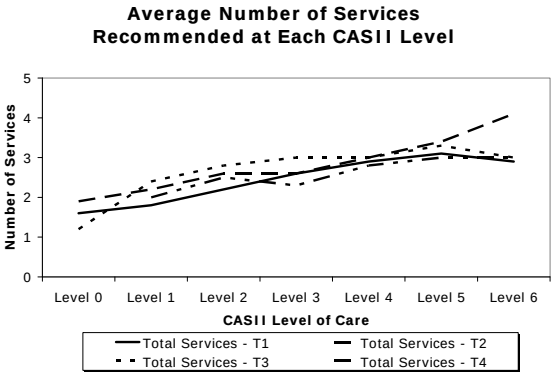
- At Time 2, children in the "Other" category (Mean 2.3, SD = 1.32) had lower CASII Level of Care scores than children with Biracial (Mean = 3.2, SD = 1.29) and Black (Mean 3.7, SD = 1.29) racial backgrounds. Children identified as White (Mean 2.7, SD = 1.43) had lower scores than children with Black racial backgrounds.

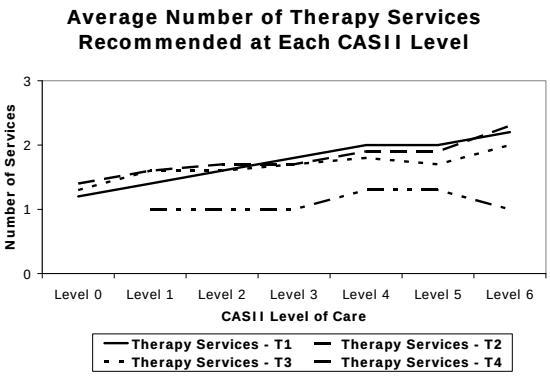
- There was no significant differences in CASII Total Score or Level of Care between children identified as Latino/a and children not identified with Latino/a ethnic background.

G. CASII: Comparisons with Providers' Services Recommendations

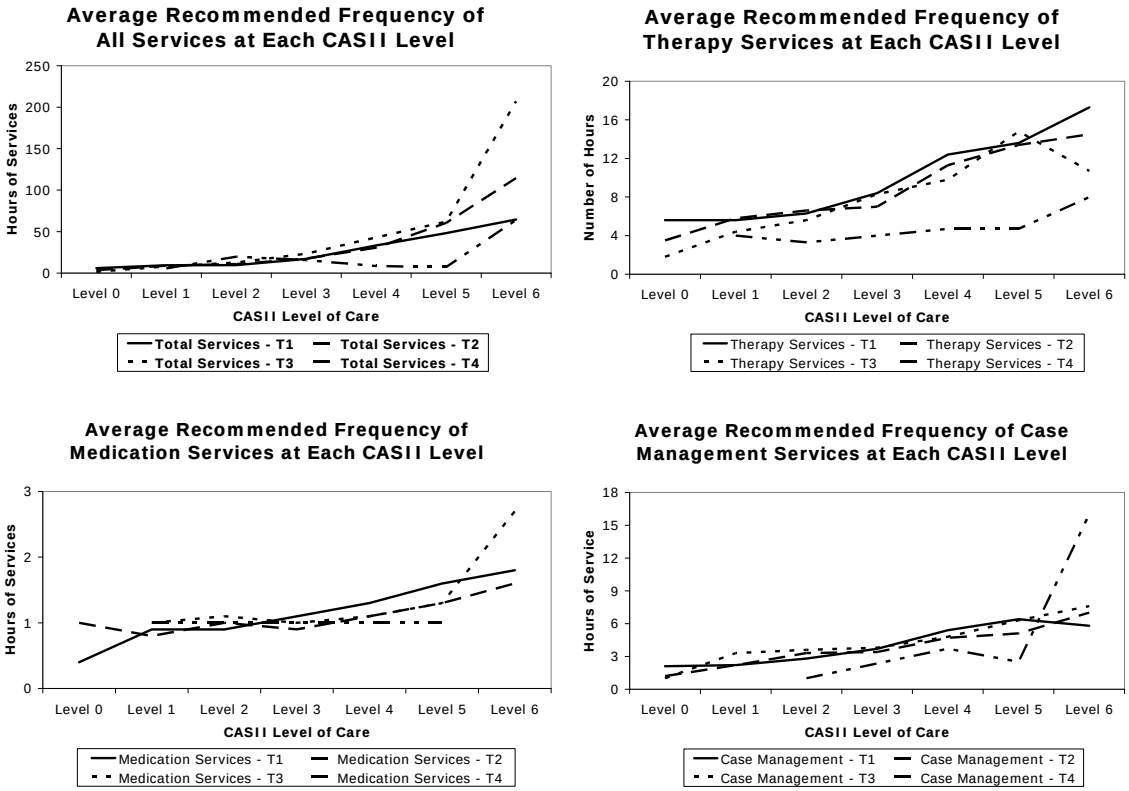
- In addition to outcome measures, providers listed the types of services they would recommend and the frequency of each service. The following figures show average number and frequency of service recommendations for each CASII Level. The average number of services recommended is plotted, which is the average number of all services providers recommended for children/adolescents whose scores fall at each level. The average number of therapy services recommended is the average number of individual, family, or group psychotherapy or skills training that providers recommended at each level. The average hours of services/month of total services, therapy services, medication services, or case management services recommended for children/adolescents are also plotted across CASII Levels. Ideally, the number and frequency of services should be higher for higher CASII levels.

CASII Levels and Number of Services Recommended





CASII Levels and Hours of Services Recommended



H. Summary of CASII Findings:

Average CASII Total scores decrease over the first 12 months of treatment and Level of Care recommendations decrease for one-fifth to one-third of cases indicating that the CASII is identifying a decrease in service needs over a one-year time period. Correlations between administration times indicate moderate test-retest reliability. Due to the nature of this instrument, high test-retest reliability is not expected. Providers tend to rate males and older children/adolescents higher on the CASII. Racial differences in scores indicate that Black and Native American children/adolescents tend to have higher CASII ratings than White or children in the "Other" racial category. The number of services recommended and number of therapy services recommended tend to be higher at CASII higher levels, as would be expected. The hours/month of all services, therapy services, case management services and medication services recommended tend to be higher at higher CASII levels.

IV. Overview of SDQ (Strengths and Difficulties Questionnaire) Data

A. Description:

- The SDQ (Goodman, 1997) is a brief behavioral screening questionnaire that is separated into two sections.
- The first section has 25 items listing 25 attributes, some positive and some negative, which are divided into five scales of five items each. The five scales include Emotional Symptoms, Conduct Problems, Inattention-Hyperactivity, Peer Problems, and Prosocial Behavior. A Total Score is comprised of the Emotional Symptoms, Conduct Problems, Inattention-Hyperactivity, and Peer Problems subscales.
- The second section is comprised of 7-9 questions and creates an Impact score that assesses the impact of symptoms on the child and the child's family or school environment. Earlier versions of the

SDQ forms and scoring materials did not include the Impact scale. Thus the Impact Score was not included in much of the data.

- The SDQ can be completed by parents, teachers, or the child and there are separate versions for each. There are also different SDQ forms based on the child's age. The same attributes are measured on each form, although the wording and examples of behaviors vary.

- The SDQ has been standardized on several populations, allowing scores to be classified into categories by the probability that a significant problem exists in a specific area. Scores are categorized into three levels of probability: Normal (score falls in the 0-79th percentile), Borderline (score falls in the 80th -89th percentile), and Abnormal (score falls in the 90th -100th percentile).

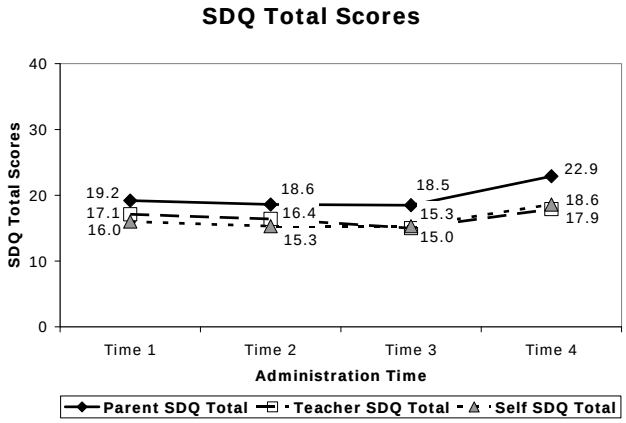
B. SDQ: Mean (Average) Scores

- Parent SDQ: Total Scores decreased from Time 1 (Mean = 19.2, SD = 6.98) to Time 2 (Mean = 18.6, SD = 6.65) and to Time 3 (Mean = 18.5, SD = 6.65). Average scores at Time 4 were based on a small sample of children and increased to 22.9 (SD = 5.33).

- Teacher SDQ: Total Scores decreased from Time 1 (Mean = 17.1, SD = 6.96) to Time 2 (Mean = 16.4, SD = 6.86) and to Time 3 (Mean = 15.0, SD = 5.71). Average scores at Time 4 were based on a small sample of children and increased to 17.9 (SD = 5.88).

- Self SDQ: Total Scores decreased from Time 1 (Mean = 16.0, SD = 5.94) to Time 2 (Mean = 15.3, SD = 6.42) and to Time 3 (Mean = 15.3, SD = 5.36). Average scores at Time 4 were based on a small sample of children and increased to 18.6 (SD = 5.76).

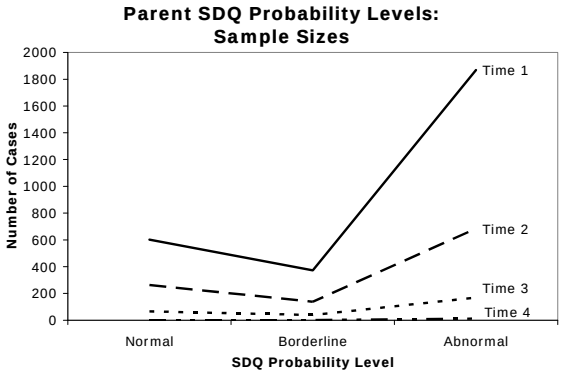
- Figure of SDQ Total Score

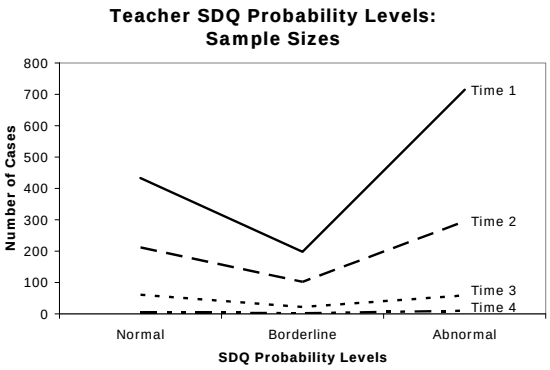


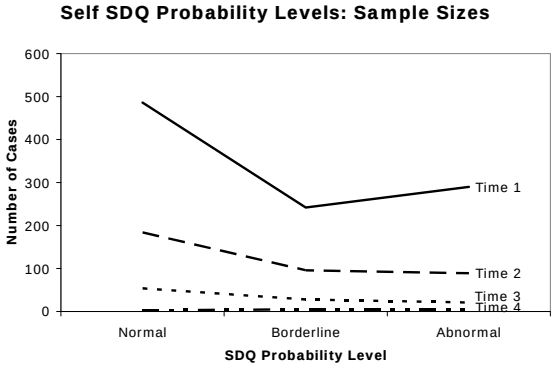
- Figures of SDQ Subscale Scores are available in Appendix A (page 22).

C. SDQ Probability Levels

- Sample Sizes of SDQ Probability Levels based on Total Scores: Probability Levels that a significant problem exists are assigned to the SDQ total and subscale scores based on the United States standardization sample obtained by SDQ authors. "Normal" probability level is equivalent to the 0 to 79th percentile range. "Borderline" probability level is equivalent to the 80th to 89th percentile range. "Abnormal" probability level is equivalent to the 90th percentile and above. The number of cases at each SDQ Probability Level are depicted on the next page.







- Tables listing sample sizes of subscale scores are in Appendix A (page 23).

D. Changes in SDQ Total Score

- Parent SDQ: Parent SDQ Total Scores decreased from Time 1 to Time 2 and from Time 2 to Time 3 in approximately half of cases. Parent SDQ Total Scores remained the same in 11% to 17% of cases from Time 1 to Time 2 and Time 2 to Time 3 respectively. Parent SDQ Total Scores increased for 39% of cases from Time 1 to Time 2 and for 34% of cases from Time 2 to Time 3.

- Teacher SDQ: Teacher SDQ Total Scores decreased from Time 1 to Time 2 and from Time 2 to Time 3 in 51% to 57% of cases. Teacher SDQ Total Scores remained the same in 9% and 15% of cases from Time 1 to Time 2 and Time 2 to Time 3 respectively. Teacher SDQ Total Scores increased for 40% of cases from Time 1 to Time 2 and for 28% of cases from Time 2 to Time 3.

- Self SDQ: Self SDQ Total Scores decreased from Time 1 to Time 2 and from Time 2 to Time 3 in 57% to 65% of cases. Self SDQ Total Scores remained the same in 9% to 15% of cases from Time 1 to Time 2 and Time 2 to Time 3 respectively. Self SDQ Total Scores increased for 33% of cases from Time 1 to Time 2 and for 28% of cases from Time 2 to Time 3.

- Changes between other administration times are available in Appendix A (page 24).

E. SDQ: Comparisons between Respondents and Administration Times

1. Correlations between SDQ respondents at the same administration time.

- Parent, Teacher, and Self SDQ Total Scores indicated low to moderate correlations (r = .20 to r = .49).

- There were not enough SDQ's at Time 4 to reliably assess the relation between respondents' endorsements.

2. Correlations between SDQ administrations

- Parent SDQ Total Scores were moderately to highly correlated between administration times (r = .57 to r = .86).
- Teacher SDQ Total Scores were moderately correlated between administration times (r = .49 to r = .54). There were not enough Teacher SDQ's at Time 4 to reliably assess the relation between this administration time and other administration times.
- Self SDQ Total Scores were moderately correlated between administration times (r = .52 to r = .57). There were not enough Self SDQ's at Time 4 to reliably assess the relation between this administration time and other administration times.

3. Differences in average scores

- Paired T-tests were used to compare average SDQ Total Scores between respondents and between administration times (See Appendix A, page 25).
- Comparisons between administration times indicated that Parent SDQ Total Scores were significantly higher at Time 1 than at Time 2 or Time 3. Parent SDQ Total Scores at Time 2 were significantly higher than at Time 3. Teacher SDQ Total Scores at Time 1 were significantly higher than at Time 2 or Time 3. Teacher SDQ Total Scores at Time 2 were significantly higher than at Time 3. Teacher SDQ Total Scores at Time 3 were significantly higher than at Time 4. Self SDQ Total Scores at Time 1 were significantly higher than at Time 2 or Time 3. Self SDQ Total Scores at Time 2 were significantly higher than at Time 3. Results are available in Appendix A.
- Average SDQ scores were compared between respondents within each administration time. At Time 1, Parent ratings were significantly higher than Teacher or Self ratings. Teacher ratings were significantly higher than Self ratings. At Time 2 and Time 3, Parent ratings were significantly higher than Teacher or Self ratings. At Time 4, Parent ratings were significantly higher than Self ratings. Results are available in Appendix A (page 25).

F. SDQ: Demographic Differences.

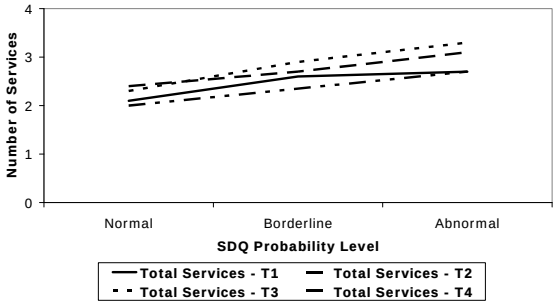
- Gender: Parents endorsed higher scores for males at Time 1 and Time 3. Teachers endorsed higher scores for males at all administration times (Times 1-4). Girls' ratings were higher than boys' at Time 1 and Time 2. More detailed statistics are available in Appendix A (pages 25-26).
- Age: Teacher and Self SDQ were negatively correlated with Age at Time 1. This indicates that there is a small trend for younger children/adolescents to have higher Teacher Total Scores and Self SDQ Total Scores. The degree of the relation was small (r = -.07) and only evident at the first administration time.
- Racial Background: For these analyses, children with Native Hawaiian/Pacific Islander racial backgrounds were recoded into the "Other" category due to their small numbers in this sample. Significant racial differences in Parent SDQ Total Scores were identified at Time 1, Time 2, and Time 3. Significant racial differences in Teacher SDQ Scores were identified at Time 1. The small sample size of some racial groups did not allow for reliable testing of which groups differed at Time 3.
  - At Time 1, parents of Black children (Mean = 20.9, SD = 7.39) endorsed higher SDQ scores than parents of children in the "Other" category (Mean = 17.8, SD = 6.99), Native American children (Mean = 18.2, SD = 7.05), or White children (Mean = 18.8, SD = 6.79). Parents of biracial children Mean = 20.2, SD = 6.47) endorsed higher ratings than parents of children in the "Other" category.
  - At Time 2, parents of Black children (Mean = 18.5, SD = 4.92) endorsed higher SDQ scores than parents of White children (Mean = 18.2, SD = 6.81).
  - At Time 1, teachers endorsed higher SDQ scores for Native American children (Mean = 19.8, SD = 6.68) than biracial children (Mean = 14.5, SD = 6.79), children in the "Other" category (Mean = 15.1, SD = 6.70), and White children (Mean = 16.9, SD = 6.96).
  - There was no significant differences in any SDQ Total Scores between children identified as Latino/a and children not identified with Latino/a ethnic background. There were no racial differences in Self SDQ Total Scores.

G. SDQ: Comparisons with Providers' Services Recommendations

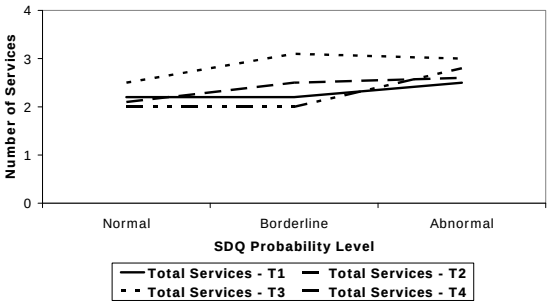
- In addition to outcome measures, providers listed the types of services they would recommend and the frequency of each service. The follow figures show average recommendations for each SDQ Probability Level. The average number of services recommended is plotted, which is the average number of all services providers recommended for each level. The average number of therapy services recommendations is the average number of individual, family, or group psychotherapy or skills training services providers recommended at each level. The frequency of each service is also plotted across SDQ Probability Levels. This is the average hours of services/month of total services, therapy services, medication services, or case management services recommended at each SDQ Probability Level. Ideally, the number and frequency of services should increase for higher SDQ Probability Levels.

SDQ Probability Levels and Number of Services Recommended

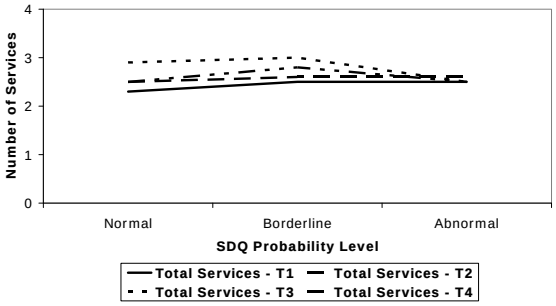
Parent SDQ and Average Number of Services Recommended



Teacher SDQ and Average Number of Services Recommended

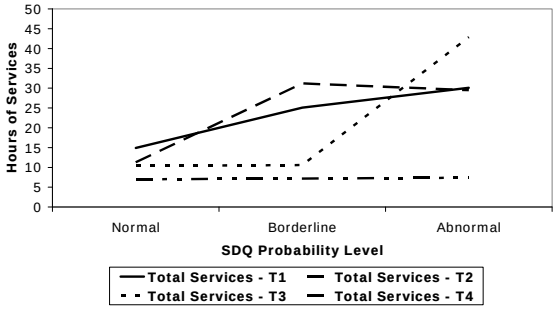


Self SDQ and Average Number of Services Recommended

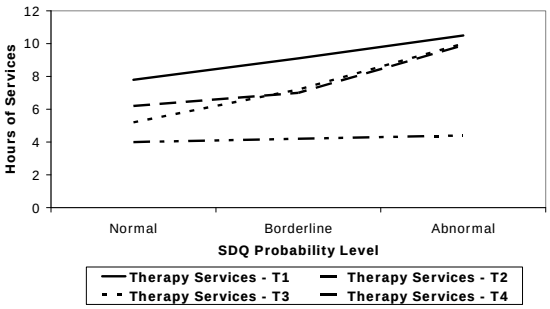


SDQ Probability Levels and Frequency of Services Recommended: Parent SDQ

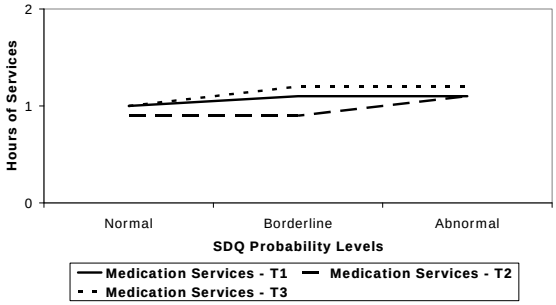
Parent SDQ and Average Frequency of Recommended Services: All Services



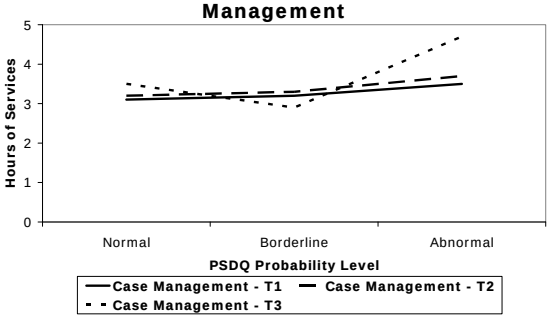
Parent SDQ and Average Frequency of Recommended Services: Therapy



Parent SDQ and Average Frequency of Recommended Services: Medication

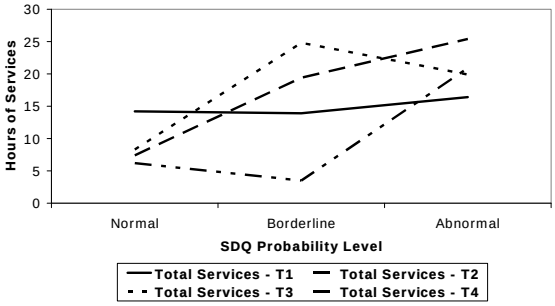


Parent SDQ and Average Frequency of Recommended Services: Case Management

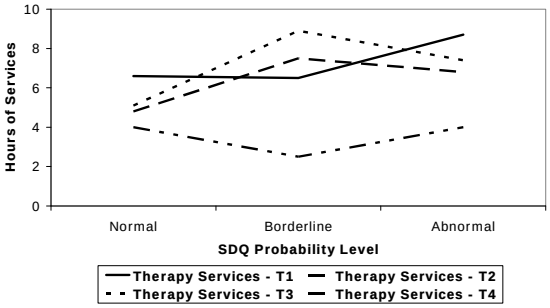


SDQ Probability Levels and Frequency of Services Recommended: Teacher SDQ

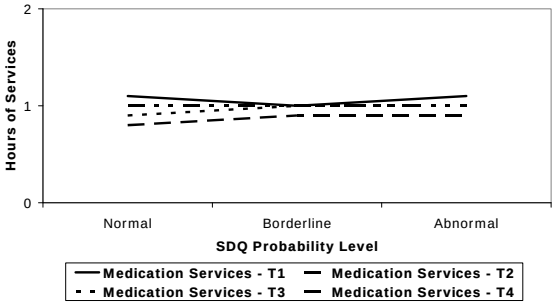
Teacher SDQ and Average Frequency of Recommended Services: All Services



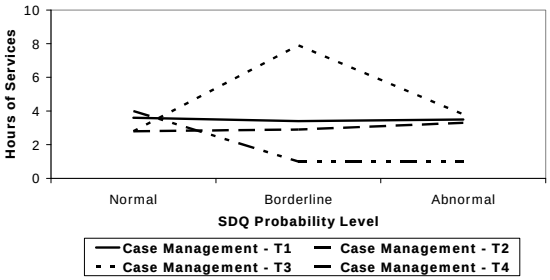
Teacher SDQ and Average Frequency of Recommended Services: Therapy



Teacher SDQ and Average Frequency of Recommended Services: Medication

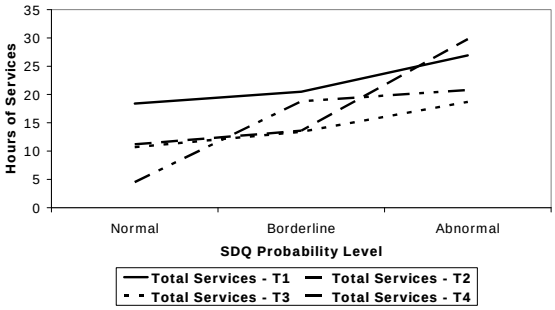


Teacher SDQ and Average Frequency of Recommended Services: Case Management

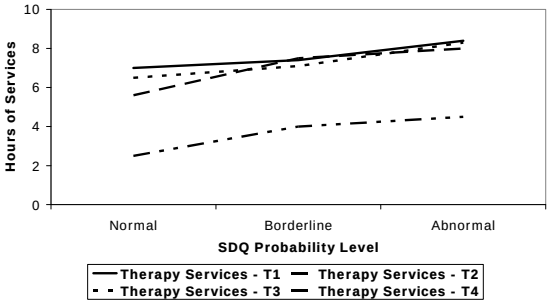


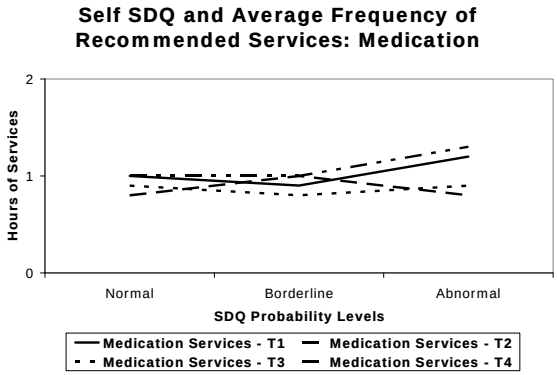
SDQ Probability Levels and Frequency of Services Recommended: Self SDQ

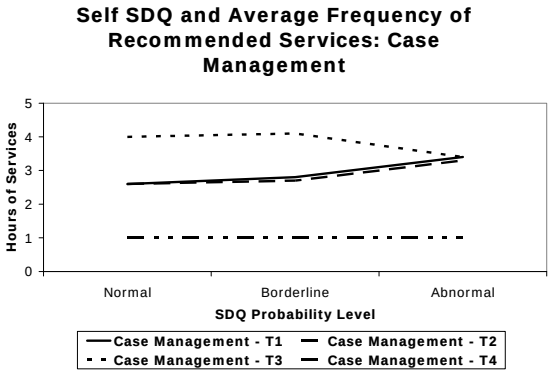
Self SDQ and Average Frequency of Recommended Services: All Services



Self SDQ and Average Frequency of Recommended Services: Therapy







G. Summary of SDQ Findings:

Average SDQ Total scores (parent, teacher, and self) decrease over the first 12 months of treatment. Total Parent SDQ scores decreased over 12 months in approximately half of cases with multiple parent reports. Teacher and Self Total SDQ scores decreased over the first 12 months of treatment for over half of cases with multiple SDQ reports. Parent, Teacher, and Self SDQ scores had low to moderate correlations which is similar to other behavioral checklists using multiple informants. SDQ responses were moderately correlated between administration times, although this cannot be considered test-retest reliability for Parent and Teacher SDQ because it is unknown if the same respondent completed the form each administration time. Parent ratings tended to be higher than Teacher and Self ratings and Teacher ratings tended to be higher than Self ratings. On average, boys received higher ratings from parents and teachers; whereas girls rated themselves higher. Parents of Black and biracial children tended to endorse higher ratings than parents of children in other racial groups. The number of services recommended tended to be higher at higher Parent and Teacher SDQ probability levels, but not at higher Self SDQ ratings. The hours/month of total services and therapy services recommended are higher at higher SDQ probability levels for parent, teacher, and self-reports indicating that children ratings as having more symptoms are also those recommended more services.

V. Overview of Recommended Mental Health Services

A. Description of Recommended Services

- Along with SDQ and CASII data, providers were asked to submit a list of mental health services they would recommend for each case. For each service, providers were asked to list the number of hours of that service they would recommend for the next month. They were also asked to submit this information (i.e., recommended monthly services and service frequencies) at each 6-month review.

B. Services Recommended Overall

- Providers recommended between 1 and 10 mental health services per month at Time 1 (Mean = 2.5, SD = 1.66). Providers recommended between 0 and 850 hours of services each month (Mean = 25.1, SD = 77.70). There are between 672 and 744 possible hours in a month. Those recommending several hundred hours of service in a month included hours of foster care or residential treatment in this total.

- At Time 2, providers recommended between 1 and 10 mental health services for the next month. The mean number of services was 2.7 (SD = 1.58). The average service frequency at Time 2 was 24.6 hours/month (SD = 80.63).

- At Time 3 between 1 and 10 services were recommended by providers (Mean = 2.9, SD = 1.55) with a mean frequency (hours/month) of 28.3 (SD = 92.03).

- At Time 4 between 1 and 4 services were recommended by providers (Mean = 2.5, SD = 0.81) with a mean frequency (hours/month) of 14.8 (SD = 20.98).

- Therapy services recommended include Individual, Family, and Group Psychotherapy as well as Individual, Family, and Group Skills Training.
  - Providers recommended from 1 to 6 therapy services at Time 1 (Mean = 1.8, SD = 0.98). They recommended between 0 to 166 hours/month in therapy services (Mean = 9.5, SD = 12.92).
  - At Time 2 the average number of therapy services recommended was 1.7 (SD = 0.97) with a mean of 8.3 hours/month in therapy (SD = 11.03).
  - At Time 3 an average of 1.6 services (SD = 0.92) were recommended for an average of 7.8 hours of therapy per month (SD = 9.83).
  - At Time 4 an average of 1.1 services (SD = 0.31) were recommended for an average of 4.3 hours of therapy per month (SD = 1.29).

- A list of recommended services and average hours of services recommended/month are listed in Appendix A (page 27).

C. Services Recommended: Demographic Differences

- Gender: Providers recommended more hours/month of medication services for girls at Time 1 and Time 2. Girls were recommended more hours/month of case management services at Time 1. More detailed statistics are available in Appendix A (page 28).

- Age:
  - Time 1: Overall number of services, overall frequency of services, frequency of therapy services, frequency of medication services, and frequency of case management services were positively correlated with Age (r's range from .06 to .15). This indicates that there is a small trend for older children/adolescents to be recommended more services and more hours of services at Time 1.
  - Time 2: Overall number of services, overall frequency of services, frequency of therapy services, frequency of medication services, and frequency of case management services were positively correlated with Age (r's range from .09 to .14). This indicates that there is a small trend for older children/adolescents to be recommended more services and more hours of services at Time 2.
  - Time 3: Overall number of services and overall frequency of services were positively correlated with Age (r's range from .11 to .14). This indicates that there is a small trend for older children/adolescents to be recommended more services and more hours of services at Time 3.

- Racial Background: For these analyses, children with Native Hawaiian/Pacific Islander racial backgrounds were recoded into the "Other" category due to the small sample size. Significant racial differences in recommended services were identified at Time 1, Time 2, and Time 3. The small sample size of some racial groups did not allow for reliable testing of which groups differed in some analyses at Times 2 and 3.

  - Number of Services Overall: At Time 1, providers recommended more services overall for Black children (Mean = 3.0, SD = 2.06) than for children in the "Other" category (Mean = 2.3, SD = 1.71) or White children (Mean = 2.4, SD = 1.49). At Time 2, providers recommended more services overall for Black children (Mean = 3.3, SD = 1.89) than for children in the "Other" category (Mean = 2.4, SD = 1.36) or White children (Mean = 2.5, SD = 1.45). Providers also recommended more services overall for Native American children (Mean = 3.2, SD = 1.73) than for White children. Although significant differences between racial groups in the number of services recommended were detected at Time 3, follow-up (post hoc) analyses did not identify these differences.

  - Frequency of Services Overall: At Time 1, providers recommended more hours/month of services for Black children (Mean = 41.6, SD = 94.02) than for White children (Mean = 20.7, SD = 74.11). At Time 2, significant racial differences in overall service frequency were detected, but post hoc analyses did not identify these differences. At Time 3, providers recommended more services overall for Black children (Mean = 71.7, SD = 179.82) than for White children (Mean = 18.7, SD = 56.75).

  - Number of Therapy Services: At Time 1, providers recommended more therapy services for Black children (Mean = 2.1, SD = 1.17) than for children in the "Other" category (Mean = 1.6, SD = 1.00) or White children (Mean = 1.7, SD = 0.90). At Time 2, providers recommended more therapy services for Black children (Mean = 1.9, SD = 1.25) and Native American children (Mean = 2.1, SD = 0.94) than for White children (Mean = 1.7, SD = 0.91). Although significant racial differences in number of therapy services recommended were detected at Time 3, follow-up (post hoc) analyses did not identify these differences.

- Frequency of Therapy Services: At Time 1, providers recommended more hours/month of therapy services for Black children (Mean = 16.6, SD = 18.28) than for White children (Mean = 7.5, SD = 10.67), children in the "Other" category (Mean = 8.6, SD = 9.51), Native American children (Mean = 10.5, SD = 12.17), and Biracial children (Mean = 10.8, SD = 13.75). At Time 2, providers recommended more hours/month of therapy services for Black children (Mean = 14.1, SD = 15.74) than for children in the "Other" category (Mean = 5.6, SD = 4.98), White children (Mean = 6.6, SD = 8.41), and Biracial children (Mean = 9.4, SD = 8.67).

- Frequency of Medication Services: At Time 1, providers recommended more hours/month of medication services for Black children (Mean = 1.7, SD = 1.57) than for White children (Mean = 1.1, SD = 0.88) and Biracial children (Mean = 1.1, SD = 0.93). At Time 2, providers recommended more hours/month of medication services for Black children (Mean = 1.4, SD = 1.39) than for White children (Mean = 0.9, SD = 0.55). At Time 3, providers recommended more hours/month of medication services for Black children (Mean = 1.8, SD = 1.84) than for White children (Mean = 0.9, SD = 0.51).

- Frequency of Case Management Services: At Time 1, providers recommended more hours/month of case management services for Black children (Mean = 6.3, SD = 6.69) than for White children (Mean = 3.6, SD = 4.06) and Biracial children (Mean = 3.7, SD = 3.93). At Time 2, providers recommended more hours/month of case management services for Native American children (Mean = 6.0, SD = 6.04) and Black children (Mean = 5.0, SD = 5.32) than for White children (Mean = 3.3, SD = 3.18).

➢ Ethnicity (Latino/a Background): Children with Latino/a ethnic background were recommended fewer services overall and fewer therapy services than children who did not have Latino/a backgrounds at Time 1. Providers recommended fewer hours of case management services for children with Latino/a backgrounds than other children at Times 1, 2, and 3. More information is available in Appendix A (page 28).

D. Summary of Services Recommended:

Although boys were rated as having more service needs and having higher symptom severity, girls tended to be recommended more medication and case management services. Providers tended to recommend more services for Black and Native-American children than for children/adolescents with other racial backgrounds which is consistent with CASII and SDQ ratings that are higher for these racial groups.

VI. Relations between CASII, SDQ, and Services

A. Correlations between CASII, SDQ, and Services

- Total scores of all instruments (Parent SDQ, Teacher SDQ, Self SDQ, and CASII scores) related significantly to each other (Appendix A, page 29).

- CASII scores related significantly to all service variables, with the exception of frequency therapy services and medication services at Time 3.

- Parent and Teacher SDQ's related most strongly to the number and hours of total services and therapy services recommended. Self SDQ scores related most strongly to the number of case management hours recommended.

IV. Clinician Satisfaction Survey:

- Providers completed an online survey about their satisfaction with the CASII and SDQ. Providers were requested to complete this survey at 2 times, once in March 07 and once in February 08.

A. CASII: Clinician Satisfaction

Item	March 2007 (N = 68)		February 2008 (N = 120)	
	Percent Disagree	**Percent Agree**	**Percent Disagree**	**Percent Agree**
Tool adequately assesses:				
Emotional Distress	8.9%	82.3%	10.5%	78.3%
Strengths/Assets	11.8%	75.0%	17.5%	70.0%
Tool provides clinically meaningful information regarding symptoms related to:				
Risk of Harm	5.9%	92.6%	7.5%	90.0%
Functional Status	5.9%	89.7%	9.2%	85.0%
Co-morbidity	10.3%	82.4%	13.4%	76.7%
Recovery Environment	5.9%	89.7%	8.4%	86.7%
Resiliency and Treatment History	3.0%	85.3%	7.5%	84.2%
Child Acceptance and Engagement	1.5%	88.3%	4.2%	89.2%
Parent Acceptance and Engagement	3.0%	86.8%	5.0%	88.3%
Tool seems appropriate for children ages:				
6 – 10	7.4%	60.3%	10.9%	60.0%
11-15	1.5%	91.2%	3.3%	90.0%
16-18	1.5%	88.2%	3.3%	87.5%
Tool appropriate for the:				
Outpatient treatment setting	4.4%	82.4%	6.6%	80.9%
In-home treatment setting	2.9%	86.8%	4.1%	85.0%
Residential treatment setting	13.2%	61.8%	14.1%	59.2%
Tool seems to:				
Require reasonable amount of time/effort to administer	13.2%	76.5%	14.1%	76.7%
Be positively received by clients	5.9%	48.5%	9.1%	44.2%
Be positively received by families	5.9%	48.6%	8.3%	45.0%

Note. Clinician's endorsed one of six options for each statement: "Strongly Disagree," "Disagree," "Undecided," "Agree," "Strongly Agree," or "Does not Apply."

B. SDQ: Clinician Satisfaction

Item	March 2007 (N = 82)		February 2008 (N = 148)	
	Percent Disagree	Percent Agree	Percent Disagree	Percent Agree
Tool adequately assesses:				
Emotional Distress	9.8%	79.3%	12.2%	79.0%
Strengths/Assets	14.6%	74.4%	15.6%	74.4%
Tool provides clinically meaningful information regarding symptoms related to:				
Depression	9.8%	79.2%	10.1%	77.0%
Trauma	25.7%	43.9%	29.8%	41.2%
Anxiety	7.3%	82.9%	11.5%	77.7%
Inattention/Hyperactivity	3.7%	87.8%	6.8%	87.1%
Disruptive Behavior	7.3%	85.4%	8.1%	84.5%
Tool seems appropriate for children ages:				
3 – 5	24.4%	24.4%	28.4%	20.9%
6 – 10	6.1%	73.2%	6.1%	72.3%
11-15	2.4%	93.9%	2.0%	93.9%
16-18	3.7%	87.8%	4.1%	85.1%
Tool appropriate for the:				
Outpatient treatment setting	1.2%	89.0%	3.4%	85.1%
In-home treatment setting	4.9%	83.0%	4.1%	83.8%
Residential treatment setting	12.2%	54.9%	13.2%	54.0%
Tool seems to be:				
Helpful in informing treatment planning	7.3%	71.9%	13.5%	70.3%
Sensitive to changes in symptoms over time	13.4%	54.9%	12.4%	61.5%
Good overall measure of strengths/difficulties	12.2%	71.9%	13.5%	72.3%
Tool seems to:				
Require reasonable amount of time/effort to administer	14.7%	79.3%	20.3%	75.0%
Be positively received by clients	7.3%	74.4%	8.2%	68.9%
Be positively received by families	6.1%	76.8%	8.8%	74.3%

Note. Clinician's endorsed one of six options for each statement: "Strongly Disagree," "Disagree," "Undecided," "Agree," "Strongly Agree," or "Does not Apply."

C. Summary of Clinician Satisfaction:

Overall, most clinicians reported satisfaction with both the CASII and SDQ. While most clinicians felt the CASII was appropriate for ages 11 to 18, fewer (approximately 60%) agreed that it was appropriate for children ages 6 to 10. Most clinicians felt the CASII was appropriate for outpatient and in-home treatment settings, while fewer (approximately 60%) agreed that it was appropriate for residential treatment settings. From 44% to 49% of providers reported that CASII results were favorably received by clients. The lower satisfaction rating on these items might indicate that providers did not share the results with clients and families. Most clinicians agreed that the SDQ assessed symptoms related to depression, anxiety, inattention/hyperactivity, and disruptive behavior; while fewer clinicians agreed that the SDQ assessed symptoms of trauma. Since the SDQ is not designed to assess trauma symptoms, it is not surprising that clinicians do not feel that this area is adequately assessed. Most clinicians agreed that the SDQ was and appropriate tool for children ages 6 to 18, but not ages 3-5. Clinicians report that the SDQ is an appropriate tool for outpatient and in-home settings, whereas fewer clinicians agree that it is appropriate for residential treatment settings. Most clinicians agree that the SDQ is helpful for treatment planning, a good overall measure of strengths/difficulties, requires a reasonable time to administer and is positively received by clients and families. Fewer clinicians agreed that the SDQ is sensitive to changes over time.

VII. Summary: Review of Results

A. Overview

- MMHAG identified two outcome measures of children's mental health: The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) and the Child and Adolescent Service Intensity Inventory (CASII; American Academy of Child and Adolescent Psychiatry, 2005). These instruments were chosen for their ability to assess effectiveness of services and improved outcomes.

B. Child and Adolescent Service Intensity Inventory (CASII)

- CASII (Child and Adolescent Service Intensity Instrument) Scores ranged across all seven levels of care and decreased from Time 1 to Time 2 and from Time 2 to Time 3. This suggests that the instrument discriminates between services needs and assesses changes in service needs and functioning over time.

- Scoring errors on the CASII were identified in 15% to 21% of all CASII's. This might indicate a need for additional training or for providers to review the scoring criteria.

- A comparison of CASII Levels and providers' recommended services indicates that providers tend to recommend more services and more hours of services at higher CASII levels, supporting the CASII as an instrument that discriminates service needs.

C. Strengths and Difficulties Questionnaire (SDQ)

- Average SDQ (Strengths and Difficulties Questionnaire) scores decreased significantly from Time 1 to Time 2 and from Time 2 to Time 3 for all respondent types (Parent, Teacher, and Self). Scores ranged across SDQ Probability Levels.

- Inter-rater reliability between SDQ respondents ranged from .20 to .49. This is similar to other omnibus behavioral checklists.

- The number of services recommended by providers was higher at higher Parent and Teacher SDQ probability levels. The frequency of total recommended services and recommended therapy services was higher at higher Parent, Teacher, and Self SDQ probability levels. These trends were in the direction expected. On average, providers recommended more services for cases that had higher parents, teachers, or child-rated symptom severity.

D. Relation between Measures

- All outcome measures (CASII, Parent SDQ, Teacher SDQ, and Self SDQ) related significantly to each other. CASII scores related significantly to all service variables. Parent and Teacher SDQ's related most strongly to the number and hours of therapy services recommended and overall service recommendations. Self SDQ scores related most strongly to the number of case management hours recommended.

E. Consumer Satisfaction

- Clinician satisfaction surveys indicated that the majority of providers agreed that these measures adequately measured emotional distress and strengths and assets. Most providers agreed that the tools measured what they proposed to measure, that they were appropriate for children of most ages and in most settings, and that they required a reasonable amount of time to administer. Most providers indicated that the SDQ was helpful in treatment planning and was positively received by clients and families.

F. Summary Statement

- Overall, outcome measures demonstrated the ability to discriminate strengths and difficulties, relate to providers' treatment recommendations, and relate to each other as would be expected for such measures.

APPENDIX A: TABLE OF CONTENTS

Page

I. Background Information 18
.....A. Counties, Agencies, and Collaborative Participants..... 18
.....B. Age 18
.....C. Gender and Racial/Ethnic Background 19
.....D. Measures Received..... 19

II. CASII 19
.....A. CASII: Mean (Average) Scores 19
.....B. CASII: Level of Care Recommendations..... 20
.....C. Change in CASII Level of Care 20
.....D. CASII Errors 20
.....E. Correlations between CASII Administrations 20
.....F. Differences between Average CASII Scores..... 20

III. SDQ 22
.....A. SDQ: Mean (Average) Total and Subscale Scores..... 22
.....B. Figures for Parent, Teacher, and Self SDQ Subscale Scores 22
.....C. SDQ Probability Levels 23
.....D. Change in SDQ Total Scores..... 24
.....E. Correlations between SDQ Respondents and Administrations..... 24
.....F. Differences between SDQ Average Scores 25

IV. Services Recommended 27
.....A. Description 27
.....B. Services Recommended: List..... 27
.....C. Services Recommended: Mean Hours/Month Recommended..... 27
.....D. Services Recommended: Demographic Differences 28

IV. Relations between CASII, SDQ, and Recommended Services 29
.....A. Correlations between CASII, SDQ, and Services 29

Appendix A: Additional Results

This section contains additional tables and information from the SDQ and CASII data. SDQ, CASII, and services recommended are organized into four different administration times:

Time 1 = initial information received for a case, regardless of the quarter in which the case was seen
Time 2 = First follow-up information for a case number
Time 3 = Second follow-up information received for a case number
Time 4 = Third follow-up information received for a case number

Results are in five sections: Background Information, CASII Results, SDQ Results, Services Recommended, and relations between variables.

I. Background Information: Includes Participants, Gender and Racial/Ethnic Background, Age, and Measures Received

A. County, Agency, Collaborative Participants

County/Agency/Collaborative	Time 1		Time 2		Time 3		Time 4	
	N	%	N	%	N	%	N	%
Becker Co. [a]	159	3.8%	33	2.0%	2	0.5%	--	--
Beltrami Area Service Col. [c]	118	2.8%	17	1.0%	2	0.5%	--	--
Benton Co. Children's MH Col. [c]	110	2.6%	45	2.7%	--	--	--	--
BRIDGES Col. [c]	28	0.7%	4	0.2%	--	--	--	--
Cass Co./Leech Lake Reservation Children's Initiative [c]	39	0.9%	13	0.8%	--	--	--	--
Clay Co./Lakeland [a]	221	5.2%	180	10.9%	89	20.6%	--	--
Clearwater Co. Children's MH Col. [c]	13	0.3%	7	0.4%	--	--	--	--
Crow Wing Co. Family Services Col. [c]	194	4.6%	27	1.6%	11	2.5%	--	--
Dakota Co. Col. [c]	27	0.6%	15	0.9%	--	--	--	--
Faribault-Martin Co. [a]	62	1.5%	36	2.2%	4	0.9%	--	--
Generations [a]	254	6.0%	25	1.5%	3	0.7%	--	--
Hennepin Co. [a]	1498	35.3%	574	34.9%	156	36.0%	2	8.7%
Hubbard Co. Family Services Col. [c]	26	0.6%	9	0.5%	--	--	--	--
Jackson Co. [a]	73	1.7%	45	2.7%	22	5.1%	--	--
Lincoln/Lyon/Murray Co. [a]	22	0.5%	3	0.2%	--	--	--	--
Mahnomen Co. Interagency Col. [c]	22	0.5%	10	0.6%	--	--	--	--
Marshall County [a]	28	0.7%	21	1.3%	10	6.6%	--	--
Northern St. Louis Co. Family Services Col. [c]	293	6.9%	208	12.6%	--	--	--	--
Otter Tail Family Services Col. [b]	200	4.7%	130	7.9%	67	44.4%	--	--
PACT 4 Families Col. [c]	28	0.7%	3	0.2%	--	--	--	--
Pine Co. Family Services Col. [c]	38	0.9%	26	1.6%	--	--	--	--
Polk Co. [a]	118	2.8%	24	1.5%	8	1.8%	2	8.7%
Redwood Co. Col. [b]	91	2.1%	14	0.9%	1	0.2%	--	--
Scott Family Net [c]	83	2.0%	37	2.2%	6	1.4%	--	--
Sibley Co. [a]	20	0.5%	4	0.2%	--	--	--	--
Stearns Co. Family Services Col. [c]	154	3.6%	62	3.8%	2	0.5%	--	--
Suburban Ramsey Family Col. [c]	118	2.8%	9	0.5%	--	--	--	--
Three Counties for Kids Children's MH Col. [c]	29	0.7%	1	0.1%	--	--	--	--
Washington Co. Child and Family Council [b]	97	2.3%	78	4.7%	48	11.1%	19	82.6%
Wright Co. Family Services Col. [c]	76	1.8%	19	1.2%	4	0.9%	--	--
Total	**4239**	**100%**	**1679**	**100%**	**435**	**100%**	**23**	**100%**

Note. Co. = County, Col. = Collaborative, MH = Mental Health, N = Number of cases, % = Percentage of all cases that administration time. Time 1 = initial data, Time 2 = 6-month follow-up, Time 3 = 12-month follow-up, Time 4 = 18 month follow-up
[a] Data submitted for pilot project, [b] Data submitted for both pilot and SLMHSG (School-Linked Mental Health Services Grant),
[c] Data submitted for SLMHSG

B. Age

	Mean (Average)	Standard Deviation	Minimum Age	Maximum Age	Sample Size
Time 1	12.1	3.56	0	20	4175
Time 2	12.4	3.51	2	20	1667
Time 3	12.9	3.26	4	19	434
Time 4	13.3	2.77	10	18	23

C. Gender and Racial/Ethnic Background

	Time 1		Time 2		Time 3		Time 4	
	N	%	N	%	N	%	N	%
Gender								
Male	2555	60.8%	1029	62.4%	271	62.3%	17	73.9%
Female	1650	39.2%	619	37.6%	164	37.7%	6	26.1%
Total	**4205**	**100.0%**	**1648**	**100.0%**	**435**	**100.0%**	**23**	**100.0%**
Race/Ethnicity								
Asian	41	1.0%	14	0.9%	4	0.9%	0	0.0%
Black	783	19.0%	260	15.9%	67	15.5%	1	4.3%
Native American	270	6.6%	96	5.9%	21	4.9%	1	4.3%
Native Hawaiian/Pacific Islander	6	0.1%	3	0.2%	0	0.0%	0	0.0%
White	2523	61.3%	1062	64.9%	287	66.4%	21	91.3%
Other or Unknown	219	5.3%	54	3.3%	7	1.6%	0	0.0%
Biracial or Multiracial	273	6.6%	147	9.0%	46	10.6%	0	0.0%
Total	**4115**	**100.0%**	**1636**	**100.0%**	**432**	**100.0%**	**23**	**100.0%**
Latino/a (Y,N)								
Latino/a	272	6.7%	115	7.2%	41	9.7%	1	4.3%
Not Latino/a	3475	86.0%	1446	90.4%	141	89.8%	22	95.7%
Unknown	293	7.3%	38	2.4%	2	0.5%	0	0.0%
Total	**4040**	**100.0%**	**1599**	**100.0%**	**421**	**100.0%**	**23**	**100.0%**

Note. N = Sample size. % = percentage of sample at that administration time. Due to incomplete data, sample sizes vary within administration times.

D. Measures Received

Measure	Time 1		Time 2		Time 3		Time 4	
	N	%	N	%	N	%	N	%
SDQ								
Parent	2843	67.1%	1086	64.7%	274	63.0%	14	60.9%
Teacher	1346	31.8%	609	36.3%	142	32.6%	17	73.9%
Child [a]	1018	37.4%	369	33.2%	103	32.2%	12	63.2%
CASII	4047	95.6%	1605	95.6%	416	95.6%	22	95.7%

Note. N = Number of cases providing data, % = percentage of sample (T1, T2, T3, or T4). Only complete data is included. Data was not included if SDQ's did not include respondent information (Parent, Teacher, or Child), if scores were outside the possible range, or if there was a problem with Client ID numbers.
[a] Only children age 11 and older can complete the SDQ. % = the percentage of children age 11 and older who completed the SDQ.

II. CASII: Includes Mean (Average) Scores, Level of Care Recommendations, Change in Scores, CASII Scoring Errors, Correlations, and Differences in Mean Scores

A. CASII: Mean (Average) Scores

CASII Items and Scales	Time 1		Time 2		Time 3		Time 4		Possible Range
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Item i: harm	2.3	1.11	2.2	0.89	2.2	0.88	2.4	0.95	1 - 5
Item ii: functioning	2.8	0.87	2.7	0.81	2.6	0.82	2.7	0.84	1 - 5
Item iii: comorbidity	2.1	1.05	2.0	0.98	2.0	0.98	1.8	0.96	1 - 5
Item iv-a: stress	2.8	1.00	2.7	0.97	2.6	0.98	2.7	1.17	1 - 5
Item iv-b: support	2.6	0.91	2.5	0.86	2.5	0.79	2.8	0.91	1 - 5
Item v: resilience	2.9	0.90	2.8	0.86	2.7	0.82	3.1	0.65	1 - 5
Item vi-a: child acceptance	2.5	0.82	2.4	0.78	2.4	0.74	2.2	0.80	1 - 5
Item vi-b: parent acceptance	2.4	0.86	2.3	0.81	2.4	0.71	2.4	0.81	1 - 5
CASII Composite Score - Provider	18.5	4.57	17.6	4.36	17.3	4.11	18.0	4.22	7 - 35
CASII Composite Score - Computed	18.3	4.56	17.5	4.25	17.2	4.04	18.1	4.17	7 - 35
Level of Care - Provider	3.2	1.44	2.9	1.36	2.8	1.31	3.2	1.26	0 - 6
Level of Care – computed	3.3	1.53	2.9	1.45	2.9	1.42	3.2	1.40	0 - 6

Note. Mean = Average score. SD = Standard Deviation.

B. CASII: Level of Care Recommendations

	Level 0		Level 1		Level 2		Level 3		Level 4		Level 5		Level 6	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Time 1	55	1.4%	517	12.8%	839	20.7%	986	24.4%	516	12.8%	882	21.8%	252	6.2%
Time 2	42	2.6%	233	14.5%	390	24.3%	425	26.5%	206	12.8%	266	16.7%	43	2.7%
Time 3	7	1.7%	67	16.1%	103	24.8%	116	27.9%	42	10.1%	72	17.3%	9	2.2%
Time 4	0	0.0%	3	13.6%	4	18.2%	6	27.3%	5	22.7%	3	13.6%	1	4.5%

C. Change in CASII Level of Care

Administration Times	Total Score Decreases		Total Score Stays Same		Total Score Increases	
	N	%	N	%	N	%
Time 1 to Time 2	570	36.5%	646	41.3%	347	22.2%
Time 1 to Time 3	179	43.6%	136	33.1%	96	23.4%
Time 1 to Time 4	5	22.7%	10	45.5%	7	31.8%
Time 2 to Time 3	129	32.2%	173	43.1%	99	24.7%
Time 2 to Time 4	5	22.7%	10	45.5%	7	31.8%
Time 3 to Time 4	5	22.7%	14	63.6%	3	13.6%

D. CASII Errors

	CASII Total Score			CASII Level of Care		
	# Cases	# Errors	% Errors	# Cases	# Errors	% Errors
Time 1	4043	653	16.2%	3937	883	22.4%
Time 2	1603	189	11.8%	1548	294	19.0%
Time 3	416	47	11.3%	399	77	19.3%
Time 4	22	1	4.5%	21	4	19.0%
Total	6084	890	14.6%	5905	1258	21.3%

Note. # Cases = Total number of cases for that administration time; # Errors = Total number of cases with errors for that administration time; % Errors = Percentage of cases that have errors. The Level of Care was left blank in some cases reducing the number of possible cases for comparison.

- Typical errors in computing the Total Score include adding both parent and child acceptance scores into the total, using the lower of the two acceptance scores (parent and child) instead of the higher score, and basic errors in addition or data entry.

- The most common error in determining the Level of Care in the CASII is the failure to use the independent criteria. The independent criteria states that if a provider rates a child at level 4 for Risk of Harm, Functioning, or Comorbidity the level of care recommended is automatically 5 or higher. If the provider rates a child at a level 5 on any of these scales the level of care recommended is automatically 6. These criteria can only be waived for Functioning and Comorbidity if the sum of Stress and Support equals 2.

- Other errors in determining CASII Level of Care include using miscalculated total scores to identify the level and listing the incorrect level for the listed total score.

E. Correlations between CASII Administrations

	Total Score				Level of Care			
	1	2	3	4	1	2	3	4
1. CASII - Time 1	--				--			
2. CASII - Time 2	.62**	--			.59**	--		
3. CASII - Time 3	.50**	.65**	--		.48**	.60**	--	
4. CASII - Time 4	.32	.64**	.77**	--	.28	.55**	.72**	--

Note. Sample sizes vary. Sample size for Time 4 correlations equals 22 which limits the power to identify significant relations.
* Significant at $p < .05$, ** Significant at $p < .01$

F. Differences between Average CASII Scores

- Paired T-tests were used to compare mean CASII Total Scores and mean CASII Level of Care recommendations between administration times. These tests assess if scores changed significantly overall from one administration time to another. The significant results are listed below.

Significant Differences between Administration Times

Comparison	Mean	SD	t	df	p
CASII Total Score Time 1 & Time 2					
Time 1	18.3	4.29			
Time 2	17.5	4.25			
Paired T-Test			8.428	1562	<.001
CASII Level of Care Time 1 & Time 2					
Time 1	3.21	1.45			
Time 2	2.93	1.45			
Paired T-Test			8.282	1562	<.001
CASII Total Score Time 1 & Time 3					
Time 1	18.3	4.00			
Time 3	17.2	4.03			
Paired T-Test			5.443	410	<.001
CASII Level of Care Time 1 & Time 3					
Time 1	3.23	1.42			
Time 3	2.90	1.42			
Paired T-Test			4.640	410	<.001
CASII Total Score Time 2 & Time 3					
Time 2	17.7	3.94			
Time 3	17.3	4.02			
Paired T-Test			2.343	400	.020

Note. t = T-test score, df = degrees of freedom in test, p = level of significance

- Paired T-tests were used to compare mean CASII Total Scores and mean CASII Level of Care recommendations between genders. These tests assess if scores changed significantly overall between genders. The significant results are listed below.

Significant Gender Differences in CASII Scores

Comparison	Mean	SD	t	df	p
CASII Level of Care Time 1			2.276	4023	.023
Male	3.3	1.52			
Female	3.2	1.54			
CASII Total Score Time 2			2.609	1595	.009
Male	17.7	4.18			
Female	17.1	4.31			
CASII Level of Care Time 2			3.444	1595	<.001
Male	3.0	1.45			
Female	2.8	1.43			
CASII Total Score Time 3			2.627	414	.009
Male	17.6	3.90			
Female	16.5	4.18			
CASII Level of Care Time 3			2.237	414	.026
Male	3.0	1.38			
Female	2.7	1.48			

- One-Way Analyses of Variance (ANOVA's) were run to explore racial differences in CASII Total Scores and Level of Care. Means and standard deviations are listed in the text. Significant overall F-tests are listed below. Sheffe' post hoc analyses were used due to uneven sample sizes.

Overall ANOVA Results for Differences between Racial Groups on CASII Scores

Comparison	F	df	p
CASII Total Score Time 1	45.992	5, 3962	<.001
CASII Level of Care Time 1	42.230	5, 3962	<.001
CASII Total Score Time 2	24.025	5, 1574	<.001
CASII Level of Care Time 2	22.348	5, 1574	<.001
CASII Total Score Time 3	3.546	5, 407	.004
CASII Level of Care Time 3	2.315	5, 407	.043

- T-tests were used to compare children with Latino/a ethnicity with those not endorsed as have Latino/a ethnicity. None of these comparisons were significant.

III. SDQ: Includes Mean (Average) Scores; Change in SDQ Total Scores; Figures for Parent, Teacher, and Self SDQ Subscale Scores; and SDQ Probability Levels

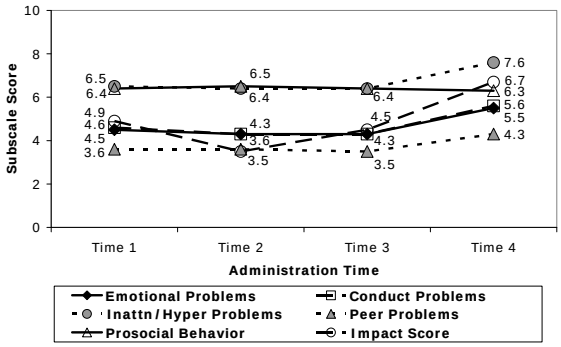
A. SDQ: Mean (Average) Total and Subscale Scores

SDQ Scales	Time 1		Time 2		Time 3		Time 4		Possible Range
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Parent Form	(N = 2843)		(N = 1086)		(N = 274)		(N = 14)		
Total	19.2	6.98	18.6	6.65	18.5	6.65	22.9	5.33	0-40
Emotional Problems	4.5	2.58	4.3	2.59	4.3	2.47	5.5	2.59	0-10
Conduct Problems	4.6	2.56	4.3	2.47	4.3	2.31	5.6	2.68	0-10
Inattn./Hyp. Problems	6.5	2.63	6.4	2.45	6.4	2.49	7.6	2.44	0-10
Peer Problems	3.6	2.41	3.6	2.14	3.5	2.07	4.3	1.86	0-10
Prosocial Behavior	6.4	2.26	6.5	2.21	6.4	2.03	6.3	1.94	0-10
Impact Score	4.9	3.29	3.5	2.96	4.5	3.01	6.7	3.07	0-10
Teacher Form	(N = 1347)		(N = 609)		(N = 143)		(N = 17)		
Total	17.1	6.96	16.4	6.86	15.0	5.71	17.9	5.88	0-40
Emotional Problems	4.0	2.64	3.7	2.57	3.5	2.40	4.3	2.52	0-10
Conduct Problems	3.4	2.61	3.2	2.53	2.5	2.14	3.5	2.27	0-10
Inattn./Hyp. Problems	6.0	2.87	5.9	2.87	5.4	2.46	5.8	2.39	0-10
Peer Problems	3.6	2.28	3.5	2.30	3.6	2.17	4.3	1.96	0-10
Prosocial Behavior	5.7	2.53	5.8	2.49	6.0	2.43	6.2	1.94	0-10
Impact Score	3.4	1.98	3.4	1.85	3.6	1.88	3.4	1.64	0-6
Self Form	(N = 1018)		(N = 369)		(N = 103)		(N = 12)		
Total	16.0	5.94	15.3	6.42	15.3	5.36	18.6	5.76	0-40
Emotional Problems	4.1	2.46	3.8	2.44	4.0	2.34	4.8	2.79	0-10
Conduct Problems	3.6	2.09	3.3	2.13	3.3	1.95	3.9	1.56	0-10
Inattn./Hyp. Problems	5.4	2.31	5.1	2.38	5.3	1.97	6.2	1.75	0-10
Peer Problems	2.9	2.00	3.0	2.07	2.8	1.79	3.7	1.23	0-10
Prosocial Behavior	7.0	2.03	7.3	2.04	7.4	1.75	7.0	2.00	0-10
Impact Score	2.2	2.68	1.7	2.38	2.8	2.76	2.6	1.80	0-10

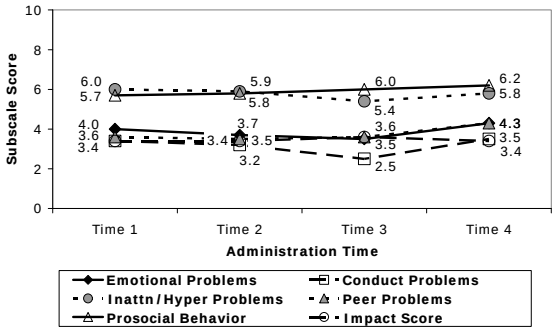
Note. Mean = Average score. SD = Standard Deviation. Impact scores were not available on all forms. Thus, the sample size for impact scores is lower than other subscales.

B. Figures for Parent, Teacher, and Self SDQ Subscale Scores

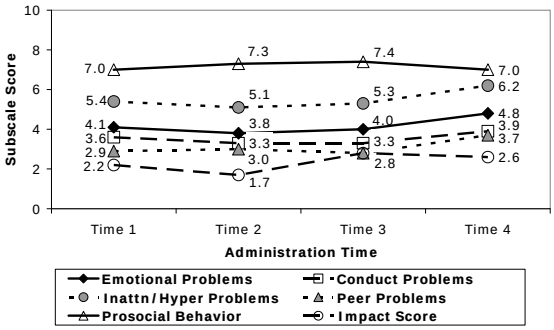
Parent SDQ Subscale Scores



Teacher SDQ Subscale Scores



Self SDQ Subscale Scores



C. SDQ Probability Levels

SDQ Scales	Time 1						Time 2					
	Normal[1]		Borderline[2]		Abnormal[3]		Normal[1]		Borderline[2]		Abnormal[3]	
	N	%	N	%	N	%	N	%	N	%	N	%
Parent Form												
Total Score	602	21.2%	372	13.1%	1869	65.7%	263	24.2%	138	12.7%	685	63.1%
Emotional Problems	1047	36.8%	376	13.2%	1423	50.0%	447	41.2%	148	13.6%	491	45.2%
Conduct Problems	659	23.2%	355	12.5%	1832	64.4%	282	25.9%	151	13.9%	655	60.2%
Inattention/Hyperact	1008	35.4%	372	13.1%	1464	51.5%	400	36.8%	155	14.2%	533	49.0%
Peer Problems	946	33.3%	463	16.3%	1435	50.5%	369	33.9%	183	16.8%	536	49.3%
Prosocial Behavior	1826	64.4%	475	16.8%	533	18.8%	706	65.1%	203	18.7%	175	16.1%
Impact Score	58	12.4%	35	7.5%	375	80.1%	33	25.4%	9	6.9%	88	67.7%
Teacher Form												
Total Score	433	32.2%	198	14.7%	715	53.1%	212	34.8%	102	16.7%	295	48.4%
Emotional Problems	621	46.1%	177	13.1%	549	40.8%	305	50.1%	80	13.1%	224	36.8%
Conduct Problems	556	41.3%	178	13.2%	613	45.5%	275	45.2%	80	13.1%	254	41.7%
Inattention/Hyperact	589	43.7%	138	10.2%	621	46.1%	284	46.6%	58	9.5%	267	43.8%
Peer Problems	460	34.1%	214	15.9%	673	50.0%	228	37.4%	102	16.7%	279	45.8%
Prosocial Behavior	666	49.5%	266	19.8%	413	30.7%	292	48.4%	140	23.2%	171	28.4%
Impact Score	30	15.0%	10	5.0%	160	80.0%	10	12.0%	3	3.6%	70	84.3%
Self Form												
Total Score	486	47.7%	242	23.8%	290	28.5%	184	49.9%	96	26.0%	89	24.1%
Emotional Problems	725	71.2%	105	10.3%	188	18.5%	281	76.2%	32	8.7%	56	15.2%
Conduct Problems	520	51.1%	157	15.4%	341	33.5%	204	55.3%	53	14.4%	112	30.4%
Inattention/Hyperact	531	52.2%	157	15.4%	330	32.4%	206	55.8%	56	15.2%	107	29.0%
Peer Problems	662	65.0%	240	23.6%	116	11.4%	228	61.8%	100	27.1%	41	11.1%
Prosocial Behavior	774	76.1%	141	13.9%	102	10.0%	291	78.9%	45	12.2%	33	8.9%
Impact Score	64	44.8%	16	11.2%	63	44.1%	23	56.1%	2	4.9%	16	39.0%

Note. Impact scores were not available for cases, thus numbers for this scale are lower than those of other scales.
[1] Normal = Approximately 0-79th percentile; [2] Borderline = Approx. 80-89th percentile; [3] Abnormal = Approx. 90-100th percentile

SDQ Scales	Time 3						Time 4					
	Normal[1]		Borderline[2]		Abnormal[3]		Normal[1]		Borderline[2]		Abnormal[3]	
	N	%	N	%	N	%	N	%	N	%	N	%
Parent Form												
Total Score	66	24.1%	39	14.2%	169	61.7%	1	7.1%	1	7.1%	12	85.7%
Emotional Problems	107	38.9%	48	17.5%	120	43.6%	3	21.4%	1	7.1%	10	71.4%
Conduct Problems	71	25.8%	36	13.1%	168	61.1%	2	14.3%	1	7.1%	11	78.6%
Inattention/Hyperact	99	36.0%	37	13.5%	139	50.5.%	4	28.6%	0	0.0%	10	71.4%
Peer Problems	94	34.3%	45	16.4%	135	49.3%	4	28.6%	1	7.1%	9	64.3%
Prosocial Behavior	180	65.5%	57	20.7%	38	13.8%	10	71.4%	2	14.3%	2	14.3%
Impact Score	3	10.3%	2	6.9%	24	82.8%	1	7.7%	0	0.0%	12	92.3%
Teacher Form												
Total Score	61	43.0%	22	15.5%	59	41.5%	5	29.4%	2	11.8%	10	58.8%
Emotional Problems	71	49.7%	21	14.7%	51	35.7%	7	41.2%	3	17.6%	7	41.2%
Conduct Problems	80	55.9%	25	17.5%	38	26.6%	5	29.4%	4	23.5%	8	47.1%
Inattention/Hyperact	77	53.8%	22	15.4%	44	30.8%	9	52.9%	3	17.6%	5	29.4%
Peer Problems	46	32.4%	23	16.2%	73	51.4%	3	17.6%	2	11.8%	12	70.6%
Prosocial Behavior	77	54.6%	32	22.7%	32	22.7%	8	50.0%	6	37.5%	2	12.5%
Impact Score	3	8.8%	3	8.8%	28	82.4%	0	0.0%	2	13.3%	13	86.7%
Self Form												
Total Score	54	52.4%	28	27.2%	21	20.4%	3	25.0%	5	41.7%	4	33.3%
Emotional Problems	78	75.7%	7	6.8%	18	17.5%	8	66.7%	1	8.3%	3	25.0%
Conduct Problems	58	56.3%	19	18.4%	26	25.2%	3	25.0%	5	41.7%	4	33.3%
Inattention/Hyperact	61	59.2%	11	10.7%	31	30.1%	3	25.0%	5	41.7%	4	33.3%
Peer Problems	66	64.1%	33	32.0%	4	3.9%	6	50.0%	5	41.7%	1	8.3%
Prosocial Behavior	89	86.4%	8	7.8%	6	5.8%	9	75.0%	2	16.7%	1	8.3%
Impact Score	6	23.1%	5	19.2%	15	57.7%	2	18.2%	1	9.1%	8	72.7%

Note. Impact scores were not available for cases, thus numbers for this scale are lower than those of other scales.
[1] Normal = Approximately 0-79th percentile; [2] Borderline = Approx. 80-89th percentile; [3] Abnormal = Approx. 90-100th percentile

D. Change in SDQ Total Scores

1. Change in Parent SDQ Scores

Administration Times	Total Score Decreases		Total Score Stays Same		Total Score Increases	
	N	%	N	%	N	%
Time 1 to Time 2	471	50.2%	101	10.8%	367	39.1%
Time 1 to Time 3	131	60.1%	13	6.0%	74	33.9%
Time 1 to Time 4	3	33.3%	1	11.1%	5	55.6%
Time 2 to Time 3	107	49.1%	37	17.0%	74	33.9%
Time 2 to Time 4	4	33.3%	1	8.3%	7	58.3%
Time 3 to Time 4	3	33.3%	0	0.0%	6	66.7%

2. Change in Teacher SDQ Scores

Administration Times	Total Score Decreases		Total Score Stays Same		Total Score Increases	
	N	%	N	%	N	%
Time 1 to Time 2	237	50.6%	43	9.2%	188	40.2%
Time 1 to Time 3	54	65.1%	2	2.4%	27	32.5%
Time 1 to Time 4	8	66.7%	0	0.0%	4	33.3%
Time 2 to Time 3	51	57.3%	13	14.6%	25	28.1%
Time 2 to Time 4	7	58.3%	2	16.7%	3	25.0%
Time 3 to Time 4	10	83.3%	1	8.3%	1	8.3%

3. Change in Self SDQ Scores

Administration Times	Total Score Decreases		Total Score Stays Same		Total Score Increases	
	N	%	N	%	N	%
Time 1 to Time 2	158	57.5%	26	9.5%	91	33.1%
Time 1 to Time 3	39	62.9%	4	6.5%	19	30.6%
Time 1 to Time 4	3	42.9%	1	14.3%	3	42.9%
Time 2 to Time 3	39	65.0%	4	6.7%	17	28.3%
Time 2 to Time 4	4	50.0%	0	0.0%	4	50.0%
Time 3 to Time 4	5	83.3%	1	16.7%	0	0.0%

E. Correlations between SDQ Respondents and Administrations

1. Correlations between different respondents at the same administration time.

Inter-Rater Reliability: Time 1 & Time 2

Instrument	1	2	3
1. Parent SDQ	--	.37**	.42**
2. Teacher SDQ	.33**	--	.20**
3. Self SDQ	.33**	.30**	--

Note. Time 1 results on bottom-left diagonal, Time 2 results on top right diagonal.
* p < .05, ** p < .01

Inter-Rater Reliability: Time 3 & Time 4

Instrument	1	2	3
1. Parent SDQ	--	.28	.62
2. Teacher SDQ	.32**	--	.00
3. Self SDQ	.42**	.49**	--

Note. Time 3 results on bottom-left diagonal, Time 4 results on top right diagonal. Correlations at Time 4 are based on a very small sample size (7-10 cases) resulting in difficulty identifying significant results.
* p < .05, ** p < .01

2. Correlations between the same respondent types (Parent, Teacher, Self) at different administration times.

Administration Time	Respondent		
	Parent	Teacher	Self
Time 1 to Time 2	.61**	.49**	.57**
Time 1 to Time 3	.57**	.49**	.54**
Time 1 to Time 4	.73*	-.03	.66
Time 2 to Time 3	.61**	.54**	.52**
Time 2 to Time 4	.73**	.27	.14
Time 3 to Time 4	.86**	.73**	.80**

Note. Correlations at Time 4 are based on a very small sample size (7-12 cases) resulting in difficulty identifying significant results.
* p < .05, ** p < .01

F. Differences between SDQ Average Scores

- Paired T-tests were used to compare mean SDQ Total Scores between respondents and administration times. Comparisons between respondents are only completed within the same administration time. Comparisons between administration times are only completed for the same type of respondent (e.g., Parent SDQ Time 1 with Parent SDQ Time 2, Teacher SDQ Time 1 with Teacher SDQ Time 2). These tests assess if scores differed significantly overall from one respondent to another or from one administration time to another. The significant results are listed below.

Significant Differences between Administration Times

Comparison	Mean	SD	t	df	p
Parent SDQ: Time 1 and Time 2					
Time 1	19.6	6.42			
Time 2	18.7	6.69			
Paired T-Test			4.642	938	<.001
Parent SDQ: Time 1 and Time 3					
Time 1	20.0	6.95			
Time 3	18.7	6.81			
Paired T-Test			3.054	217	.003
Parent SDQ: Time 2 and Time 3					
Time 2	19.7	6.52			
Time 3	18.9	6.49			
Paired T-Test			2.239	217	.026
Teacher SDQ: Time 1 and Time 2					
Time 1	17.4	6.85			
Time 2	16.6	7.02			
Paired T-Test			2.609	467	.009
Teacher SDQ: Time 1 and Time 3					
Time 1	17.8	6.42			
Time 3	14.9	6.92			
Paired T-Test			4.143	82	<.001
Teacher SDQ: Time 2 and Time 3					
Time 2	17.5	6.93			
Time 3	15.0	6.02			
Paired T-Test			3.739	88	<.001
Teacher SDQ: Time 3 and Time 4					
Time 3	20.9	5.82			
Time 4	15.9	5.30			
Paired T-Test			4.168	11	.002
Self SDQ: Time 1 and Time 2					
Time 1	16.6	5.75			
Time 2	15.1	6.29			
Paired T-Test			4.669	274	<.001
Self SDQ: Time 1 and Time 3					
Time 1	17.7	5.53			
Time 3	15.6	5.22			
Paired T-Test			3.200	61	.002
Self SDQ: Time 2 and Time 3					
Time 2	15.9	5.45			
Time 3	14.4	5.06			
Paired T-Test			2.280	59	.026

Note. t = T-test score, df = degrees of freedom in test, p = level of significance.

Significant Differences between Respondents: Part 1

Comparison	Mean	SD	t	df	p
Time 1: Parent and Teacher					
Parent SDQ Total Score	18.2	6.74			
Teacher SDQ Total Score	16.9	6.93			
Paired T-Test			5.019	938	<.001
Time 1: Parent and Self					
Parent SDQ Total Score	18.2	6.77			
Self SDQ Total Score	15.8	5.98			
Paired T-Test			8.755	696	<.001
Time 1: Teacher and Self					
Teacher SDQ Total Score	16.9	7.02			
Self SDQ Total Score	16.2	5.95			
Paired T-Test			2.266	581	.024

Note. t = T-test score, df = degrees of freedom in test, p = level of significance.

Significant Differences between Respondents: Part 2

Comparison	Mean	SD	t	df	p
Time 2: Parent and Teacher					
Parent SDQ Total Score	17.8	6.85			
Teacher SDQ Total Score	16.6	6.95			
Paired T-Test			3.155	403	.002
Time 2: Parent and Self					
Parent SDQ Total Score	18.0	6.97			
Self SDQ Total Score	15.6	6.51			
Paired T-Test			4.977	239	<.001
Time 3: Parent and Teacher					
Parent SDQ Total Score	17.2	6.58			
Teacher SDQ Total Score	15.3	5.65			
Paired T-Test			2.472	82	.015
Time 3: Parent and Self					
Parent SDQ Total Score	18.2	5.64			
Self SDQ Total Score	14.2	4.19			
Paired T-Test			5.984	66	<.001
Time 4: Parent and Self					
Parent SDQ Total Score	22.1	5.15			
Self SDQ Total Score	17.3	6.02			
Paired T-Test			2.598	6	.041

Note. t = T-test score, df = degrees of freedom in test, p = level of significance.

- Paired T-tests were used to compare mean SDQ Total Scores between genders. These tests assess if scores changed significantly overall between genders. The significant results are listed below.

Significant Gender Differences in SDQ Scores

Comparison	Mean	SD	t	df	p
Parent SDQ Time 1					
Male	19.6	6.90			
Female	18.7	7.08			
Paired T-Test			3.287	2807	.001
Parent SDQ Time 3					
Male	19.3	6.66			
Female	17.2	6.50			
Paired T-Test			2.501	272	.013
Teacher SDQ Time 1					
Male	17.8	6.72			
Female	15.8	7.13			
Paired T-Test			5.228	1336	<.001
Teacher SDQ Time 2					
Male	17.2	6.66			
Female	15.0	6.61			
Paired T-Test			3.923	602	<.001
Teacher SDQ Time 3					
Male	15.9	5.64			
Female	13.6	5.55			
Paired T-Test			2.464	140	.015
Teacher SDQ Time 4					
Male	20.5	4.56			
Female	11.6	3.36			
Paired T-Test			3.911	16	.001
Self SDQ Time 1					
Male	15.2	5.90			
Female	16.8	5.89			
Paired T-Test			-4.185	1014	<.001
Self SDQ Time 2					
Male	14.3	6.12			
Female	16.3	6.53			
Paired T-Test			-3.07	366	.002

Note. t = T-test score, df = degrees of freedom in test, p = level of significance.

- One-Way Analyses of Variance (ANOVA's) were run to explore racial differences in SDQ Total Scores. Due to the small sample size at Time 4, racial differences could not be analyzed at that administration time. Means and standard deviations are listed in the text. Significant overall F-tests are listed below. Sheffe' post hoc analyses were used due to uneven sample sizes.

Overall ANOVA Results for Differences Between Racial Groups on SDQ Scores

Comparison	F	df	p
Parent SDQ Total Score Time 1	10.518	5, 2734	<.001
Parent SDQ Total Score Time 2	4.104	5, 1047	.001
Parent SDQ Total Score Time 3	2.634	5, 266	.024
Teacher SDQ Total Score Time 1	5.018	5, 1322	<.001

- T-tests were used to compare children with Latino/a ethnicity with those not endorsed as have Latino/a ethnicity. None of these comparisons were significant.

IV. Services Recommended

A. Description: Along with SDQ and CASII data, providers were asked to submit a list of mental health services they would recommend for each case. For each service, providers were asked to list the number of hours of that service they would recommend for the next month. They were also asked to submit this information (i.e., recommended monthly services and service frequencies) at each 6-month review.

B. Services Recommended: List

Service Proposed	Time 1		Time 2		Time 3		Time 4	
	# Clients	% of Cases	# Clients	% of Cases	# Clients	% of Cases	# Clients	% of Cases
Case Management	2230	58.0%	902	59.4%	276	69.5%	7	33.3%
Day Treatment	393	10.2%	167	11.0%	44	11.1%	2	9.5%
Family Psychotherapy	718	18.7%	267	17.6%	58	14.6%	0	0.0%
Family Skills	562	14.6%	244	16.1%	51	12.8%	0	0.0%
Group Psychotherapy	360	9.4%	130	8.6%	25	6.3%	0	0.0%
Group Skills	608	15.8%	282	18.6%	58	14.6%	0	0.0%
Individual Psychotherapy	1756	45.7%	665	43.8%	181	46.5%	2	9.5%
Individual Skills	1248	32.5%	560	36.9%	148	37.3%	20	95.2%
Med – Primary Physician	260	6.8%	95	6.3%	29	7.3%	0	0.0%
Med – Psychiatrist	1127	29.3%	592	38.3%	203	51.1%	15	71.4%
Residential Treatment	230	6.0%	76	5.0%	18	4.5%	1	4.8%

Note. Cases listing recommended services: Time 1 = 3844, Time 2 = 1519, Time 3 = 397, Time 4 = 21

C. Services Recommended: Mean Hours/Month Recommended

Service Proposed	Time 1		Time 2		Time 3		Time 4	
	Average Hours	SD	Average Hours	SD	Average Hours	SD	Average Hours	SD
Case Management	4.4	5.17	4.0	4.27	4.4	4.04	4.7	5.79
Day Treatment	50.7	33.67	56.4	27.93	55.1	24.27	60.0	0.00
Family Psychotherapy	3.2	3.70	3.2	3.59	3.4	2.33	--	--
Family Skills	3.9	3.95	3.6	4.32	2.7	1.66	--	--
Group Psychotherapy	11.2	8.62	10.8	9.14	10.1	6.60	--	--
Group Skills	9.1	10.37	6.0	7.46	6.1	8.47	--	--
Individual Psychotherapy	3.6	3.13	3.4	2.85	3.0	1.84	2.0	0.00
Individual Skills	5.7	8.60	5.5	6.67	5.5	8.37	4.1	1.15
Med – Primary Care Physician	1.2	1.00	1.1	1.33	1.1	0.69	--	--
Med– Psychiatrist	1.2	1.01	1.0	0.62	1.1	0.85	1.0	0.00
Residential Treatment	133.6	231.16	196.2	284.63	284.6	345.35	40	--
Total # of services	2.5	1.66	2.7	1.58	2.9	1.55	2.5	0.81
Total Hrs of services	25.1	77.70	24.6	80.63	28.3	92.03	14.8	20.98
Total # of therapy services	1.8	0.98	1.7	0.97	1.6	0.92	1.1	0.31
Total Hrs of therapy services	9.5	12.92	8.3	11.03	7.8	9.83	4.3	1.29
Total # of medication services	1.0	0.17	1.0	0.17	1.0	0.17	1.0	0.00
Total Hrs of medication services	1.3	1.17	1.1	0.90	1.1	1.07	1.0	0.00

Note. Cases listing recommended services: Time 1 = 3844, Time 2 = 1519, Time 3 = 397, Time 4 = 21

D. Services Recommended: Demographic Differences

- Paired T-tests were used to compare the number and frequency of recommended services between genders. These tests assess if scores changed significantly overall between genders. The significant results are listed below.

Significant Gender Differences in SDQ Scores

Comparison	Mean	SD	t	df	p
Frequency of Medication Services: Time 1					
Male	1.2	1.08			
Female	1.4	1.29			
Paired T-Test			-2.584	1165	.010
Frequency of Case Management Services: Time 1					
Male	4.1	4.80			
Female	5.0	5.68			
Paired T-Test			-3.654	1765	<.001
Frequency of Medication Services: Time 2					
Male	1.0	0.72			
Female	1.2	1.17			
Paired T-Test			-3.113	554	.002

Note. t = T-test score, df = degrees of freedom in test, p = level of significance.

- One-Way Analyses of Variance (ANOVA's) were run to explore racial differences in the number and frequency of recommended services. Due to the small sample size at Time 4, racial differences could not be analyzed at that administration time. Means and standard deviations are listed in the text. Significant overall F-tests are listed below. Sheffe' post hoc analyses were used due to uneven sample sizes.

Overall ANOVA Results for Differences Between Racial Groups on Recommended Services

Comparison	F	df	p
Number of Services Overall: Time 1	15.350	5, 3780	<.001
Frequency of Services Overall: Time 1	7.147	5, 3297	<.001
Number of Therapy Services: Time 1	16.146	5, 2730	<.001
Frequency of Therapy Services: Time 1	37.448	5, 1315	<.001
Frequency of Medication Services: Time 1	14.513	5, 1143	<.001
Frequency of Case Management Services: Time 1	17.513	5, 1744	<.001
Number of Services Overall: Time 2	12.753	5, 1494	<.001
Frequency of Services Overall: Time 2	4.920	5, 1334	<.001
Number of Therapy Services: Time 2	5.754	5, 1210	<.001
Frequency of Therapy Services: Time 2	17.943	5, 1126	<.001
Frequency of Medication Services: Time 2	5.260	5, 546	<.001
Frequency of Case Management Services: Time 2	6.375	5, 732	<.001
Number of Services Overall: Time 3	2.800	5, 387	.017
Frequency of Services Overall: Time 3	3.392	5, 314	.005
Number of Therapy Services: Time 3	3.167	5, 308	.008
Frequency of Medication Services: Time 3	4.567	5, 230	.001
Frequency of Case Management Services: Time 3	3.108	5, 206	.010

- T-tests were used to compare children with Latino/a ethnicity with those not endorsed as have Latino/a ethnicity. The following are the significant findings.

Comparison	Mean	SD	t	df	p
Number of Services Overall: Time 1			-2.082	320	.038
Latino/a	2.4	1.36			
Not Latino/a	2.5	1.62			
Number of Therapy Services: Time 1			-3.220	202	.001
Latino/a	1.6	0.83			
Not Latino/a	1.8	0.97			
Frequency of Case Management Services: Time 1			-4.321	237	<.001
Latino/a	3.0	3.36			
Not Latino/a	4.2	5.01			
Frequency of Case Management Services: Time 2			-2.323	95	.022
Latino/a	3.0	3.40			
Not Latino/a	4.0	4.34			
Frequency of Case Management Services: Time 3			-2.888	72	.005
Latino/a	3.2	1.99			
Not Latino/a	4.6	4.23			

Note. t = T-test score, df = degrees of freedom in test, p = level of significance.

IV. Relations between CASII, SDQ, and Services

A. Correlations between CASII, SDQ, and Recommended Services:

- **Correlations at Time 1**

Instrument	1	2	3	4	5	6	7	8	9
1. CASII Total Score	--								
2. Parent SDQ	.33**	--							
3. Teacher SDQ	.36**	.33**	--						
4. Self SDQ	.18**	.33**	.30**	--					
5. Total # Services	.24**	.16**	.15**	.09**	--				
6. Total Hours Services	.24**	.07**	.03	.05	.36**	--			
7. Total # Therapy Services	.26**	.11**	.09**	.06	.76**	.26**	--		
8. Total Hours/Mo Therapy	.27**	.11**	.13**	.08*	.55**	.43**	.59**	--	
9. Total Hours/Mo Med	.29**	.04	.02	.08	.36**	.25**	.23**	.31**	--
10. Total Hours/Mo CM	.30**	.06*	-.05	.15**	.40**	.19**	.27**	.29**	.49**

* Significant at p < .05. ** Significant at p < .01.

- **Correlations at Time 2**

Instrument	1	2	3	4	5	6	7	8	9
1. CASII Total Score	--								
2. Parent SDQ	.30**	--							
3. Teacher SDQ	.34**	.37**	--						
4. Self SDQ	.15**	.42**	.20**	--					
5. Total # Services	.28**	.21**	.13**	.04	--				
6. Total Hours Services	.25**	.09**	.07	.14*	.34**	--			
7. Total # Therapy Services	.16**	.11**	.11*	.05	.72**	.23**	--		
8. Total Hours/Mo Therapy	.27**	.15**	.13**	.13*	.53**	.41**	.55**	--	
9. Total Hours/Mo Med	.23**	.06	.04	.06	.27**	.28**	.17**	.27**	--
10. Total Hours/Mo CM	.31**	.06	.13	.18*	.29**	.29**	.24**	.27**	.51**

* Significant at p < .05. ** Significant at p < .01.

- **Correlations at Time 3**

Instrument	1	2	3	4	5	6	7	8	9
1. CASII Total Score	--								
2. Parent SDQ	.43**	--							
3. Teacher SDQ	.28**	.32**	--						
4. Self SDQ	.46**	.42**	.49**	--					
5. Total # Services	.19**	.25**	.15	-.05	--				
6. Total Hours Services	.32**	.17**	.16	.19	.18**	--			
7. Total # Therapy Services	.08	.13	-.04	.02	.76**	.12	--		
8. Total Hours/Mo Therapy	.34**	.23**	.10	.11	.51**	.26**	.49**	--	
9. Total Hours/Mo Med	.13	.14	.13	.17	.02	.16*	.12	.08	--
10. Total Hours/Mo CM	.27**	.11	.00	-.04	.12	.14*	.06	.19*	.33**

* Significant at p < .05. ** Significant at p < .01.

- **Sample size at Time 4 was too small to reliably assess significant relations between variables.**