



**Foundation 2**  
**DISCOVERIES**  
**Gender-Specific**  
**Day Treatment Program**

**Evaluation Final Report**



**Service Evaluation & Development Associates, Inc.**  
**Mount Vernon, IA**



**Foundation 2**  
**DISCOVERIES**  
**Gender-Specific**  
**Day Treatment Program**

**Evaluation Final Report**

**May, 2003**

**Anita M. Patterson, M.S.W., SEDA, Inc.**  
**Robert A. Hintz, Jr., Ph.D., Foundation 2**



**Service Evaluation & Development Associates, Inc.**  
**Mount Vernon, IA**

**Discoveries Gender-Specific Day Treatment Program  
Evaluation Final Report**

**TABLE OF CONTENTS**

Table of Contents .....	i
Table of Figures .....	ii
Table of Tables .....	ii
Preface and Acknowledgements .....	iv
Executive Summary .....	v
I. Background .....	1
II. Project Description .....	3
A. Program Overview and Activities .....	3
B. Evaluation Design .....	5
C. Sampling & Data Collection.....	12
III. Study Participants .....	15
A. Project Participation .....	15
B. Baseline Descriptive Data .....	19
IV. Quantitative Results .....	23
A. School Domain .....	24
B. Empowerment & Self-Efficacy .....	26
C. Prosocial Attitudes & Behavior .....	29
D. Relationships .....	35
E. Health Domain .....	38
F. Program & Services Data .....	50
V. Qualitative Results .....	53
A. Summary of Youth Interviews .....	53
B. Summary of Staff Interviews .....	58
VI. Conclusions & Recommendations .....	62
VII. References .....	71

APPENDICES

**Discoveries Gender-Specific Day Treatment Program  
Evaluation Final Report**

**TABLE OF FIGURES**

Figure 1. Referral Source .....	19
Figure 2. Adjudication Status .....	19
Figure 3. Single Parent Home .....	20
Figure 4. Attended Church .....	20
Figure 5. Lived out of Home.....	20

**TABLE OF TABLES**

Table 1. Data Collection Instruments .....	6
Table 2. Interviews by Time Point .....	16
Table 3. Interviews by Referral Source .....	16
Table 4. Sample Sizes .....	17
Table 5. Age at Intake .....	17
Table 6. Race/ethnicity .....	18
Table 7. County & City of Residence .....	18
Table 8. Current Living Arrangements .....	18
Table 9. Last Grade Completed .....	19
Table 10. Baseline Group Differences on MACI Personality Patterns .....	21
Table 11. Baseline Group Differences on MACI Expressed Concerns and Clinical Syndromes .....	22
Table 12. Differences Between Groups on Anxiety Subscale Measures .....	23
Table 13. Self-Report School Outcomes for Discoveries Clients .....	25
Table 14. Between Groups Comparison of Change in Self-Report School Truancy .....	25
Table 15. CAFAS School/Work Domain, Pre-Post Difference in Means .....	26
Table 16. T1-T3 Comparisons on Empowerment Subscales .....	27
Table 17. T1-T4 Significant Differences on Empowerment Subscales .....	28
Table 18. Between Groups Comparison of Change on Empowerment Subscales .....	28
Table 19. Self-Reported Changes in Empowerment-Related Areas .....	29
Table 20. T1-T3 Comparisons on Compliance Subscales .....	30
Table 21. T1-T4 Significant Differences on Compliance Subscales .....	31
Table 22. Between Groups Comparison of Change on Compliance Subscales .....	31
Table 23. CAFAS Total Score, Pre-Post Difference in Means .....	32
Table 24. CAFAS Community Domain, Pre-Post Difference in Means .....	32
Table 25. CAFAS Behavior Toward Others, Pre-Post Difference in Means .....	32
Table 26. CAFAS Thinking Scores .....	32
Table 27. Criminal Allegations Recorded by the Juvenile Court Pre-, During, and Post-Treatment .....	33
Table 27a. Months since Discharge .....	34
Table 28. Self-Reported Problems in Past Three Months at T1 and T3, Delinquency Issues .....	34
Table 29. Change in Self-Reported Criminal Activity .....	35
Table 30. T1-T3 Comparisons on Relationship Subscales .....	35
Table 31. T1-T4 Significant Differences on Relationship Subscales .....	36
Table 32. Between Groups Comparison of Change on Relationship Subscales .....	36
Table 33. CAFAS Home Domain, Pre-Post Difference in Means .....	36
Table 34. Family Cohesion scores .....	37

## TABLE OF TABLES, cont.

Table 35. Perceived Past 3 Month Improvement in Relationships .....	37
Table 36. T1-T3 Comparisons on Health Subscales .....	38
Table 37. Between Groups Comparison of Change on Health Subscales .....	39
Table 38. Changes in Suicidal Tendency by Sample .....	39
Table 39. T1-T4 Significant Differences on Health Subscales .....	40
Table 40. T1-T3 Comparisons on Total Testwell Scores by Sample .....	41
Table 41. Between Groups Comparison of Testwell Total Scores at all Time Points .....	41
Table 42. Discoveries vs. Comparison-services sample on Testwell Change Scores .....	42
Table 43. Discoveries vs. Comparison-no services sample on Testwell Change Scores .....	42
Table 44. Changes in Dieting and Safe Sex Measures by Sample .....	43
Table 45. T1-T3 Comparisons on Testwell Subscales .....	44
Table 46. Between Groups Comparison of Change on Testwell Subscales .....	45
Table 47. Correlations in MACI and Testwell Changes over Time .....	46
Table 48. Correlations of Changes in Identity Diffusion and Testwell Scores by Sample .....	47
Table 49. Self-Reported Problems in Past Three Months at T1 and T3, Health Issues .....	47
Table 50. Substance Abuse Measures at Baseline .....	48
Table 51. Discoveries sample, Changes in Substance Use Measures, T1 - T3 .....	48
Table 52. CAFAS Substance Use Domain, Pre-Post Difference in Means .....	49
Table 53. CAFAS Self-Harmful Behavior Scores .....	49
Table 54. CAFAS Moods/Emotions Domain, Pre-Post Difference in Means .....	49
Table 55. Correlations between Change in Testwell scores and Service Hours .....	50
Table 56. Correlations between Change in MACI subscale scores and Service Hours .....	51
Table 57. Progress Made During Treatment .....	52
Table 58. Reasons for Less than 50% Progress .....	52
Table 59. Placement At Discharge .....	52

**Discoveries Gender-Specific Day Treatment Program  
Evaluation Final Report**

**PREFACE & ACKNOWLEDGEMENTS**

This document represents the summative report for the evaluation study of the Discoveries Day Treatment Program in Cedar Rapids, Iowa, a nearly two-year project that was initiated in July, 2001 as a collaborative effort between Foundation 2, Inc., as host agency for the program and Service Evaluation & Development Associates, Inc., (SEDA) of Mount Vernon, Iowa, as the research contractor. The project was inspired and supported by the Iowa Gender-Specific Services Task Force, operating under the auspices of the Iowa Commission on the Status of Women and the Division of Criminal and Juvenile Justice Planning (CJJP), both of the latter housed within the Iowa Department of Human Rights. Funding for the project was allocated by CJJP and the Iowa Juvenile Justice Advisory Council, drawn from federal juvenile justice funds provided by the U.S. Office of Juvenile Justice & Delinquency Prevention.

Funding and support for the project does not constitute or imply endorsement of the project activities, its results, or the opinions expressed in the report. All materials, protocols, and contents of the report are solely the responsibility of the Research Director.

The project was supported at various times by the efforts of staff and administrators from the Foundation 2 Youth Shelter, Foundation 2 Family Services, and Linn & Johnson County Juvenile Court Services.

Special thanks are due Charlotte Bright, M.S.W., and the staff from Discoveries for their willingness to share their insight and experience with providing gender-specific services, and most especially to all of the girls in and outside of Discoveries who contributed their time and life experience, which was the heart of the project's success.

The following SEDA staff also provided vital assistance at key points during the course of the project:

Charlene Fisher, B.A., Research Assistant  
Christine Barton, Project Assistant  
Melissa Yoder, Data Entry Clerk  
Michele Gobush, Research Assistant  
Lindsay Patterson, Research Assistant  
Stephanie Swartz, B.A., Project Assistant

A copy of the Executive Summary of this report may be obtained by calling SEDA at (319) 895-8766, or Robert Hintz at Foundation 2 at (319) 366-8797. The Executive Summary and full Report may also be printed from SEDA's website at [www.sedainc.org](http://www.sedainc.org).

## **Discoveries Gender-Specific Day Treatment Program Evaluation Final Report**

### **EXECUTIVE SUMMARY**

The Discoveries Gender-Specific Day Treatment Program Evaluation was a nearly two-year project that was initiated in July, 2001 as a collaborative effort between Foundation 2, Inc., as host agency for the program and Service Evaluation & Development Associates, Inc., (SEDA) of Mount Vernon, Iowa. The project was inspired and supported by the Iowa Gender-Specific Services Task Force, operating under the auspices of the Iowa Commission on the Status of Women and the Division of Criminal and Juvenile Justice Planning (CJJP), as part of a larger effort to better meet the needs of girls in the juvenile justice system in Iowa.

Discoveries is a day treatment program operated by Foundation 2 in Cedar Rapids, Iowa that is designed as a gender-specific approach to providing therapeutic services and supports to girls between the ages of 14 and 17 who are at risk of being placed out of the community because of emotional and/or behavioral problems. Most clients are referred by the Juvenile Court Office and/or the Department of Human Services. They typically attend the program for 25 hours per week after school Monday through Friday for a 12-week intensive phase of treatment and then participate in aftercare sessions for another six weeks. Clients are admitted in cohorts, every 12 weeks, comprised of 5-8 girls each. In-house program staff include the Coordinator, two case managers, the school liaison, the aftercare program coordinator, and a part-time counselor. The Coordinator and two case managers design and implement most of the direct program services each evening, and in addition, the case managers maintain contact with each client and her family over the weekend. The program also employs part-time youth trackers and part-time therapists who meet with the girls individually.

The research design for the Discoveries evaluation was quasi-experimental, using a pre-post treatment design with comparison groups. Pretreatment measures were administered at intake for all clients who consented to participate in the evaluation. The same measures were administered again at the end of their intensive phase of treatment (typically 12 weeks), at the end of their aftercare phase of treatment which also marked the end of their treatment episode altogether (typically 6 weeks after the end of intensive), and again at 8-10 weeks follow-up. For the treatment group, the interview points were referred to as intake, discharge, aftercare, and follow-up, or Time1, Time2, Time3, and Time4. The same measures were administered to comparison subjects 4 times corresponding to the same time intervals typically experienced by the clients in treatment.

Research questions focused on the extent to which clients experienced positive changes, as compared to girls who did not attend the program, in: a) school and/or job performance? b) prosocial attitudes and behavior, including reduced involvement with Juvenile Court? c) perceptions of self-efficacy, assertiveness, and empowerment? d) satisfaction in relationships with family members, other adults, and peers?; and e) health and self-care, including avoidance and/or reduction in substance use & unsafe sex?

Secondary data was collected from the following sources during the evaluation: program admission and discharge forms, group and individual session forms, the Juvenile Court database, and the Child and Adolescent Functional Assessment Scale (CAFAS), the latter an instrument completed by program staff for each client at intake and discharge. In addition, three instruments were administered by research staff

to all treatment and comparison subjects: the Millon Adolescent Clinical Inventory (MACI), a validated clinical and personality assessment tool that measures change over time; the Testwell Wellness Inventory for Adolescents, a comprehensive measure of health-related behaviors; and a Youth Questionnaire, an instrument adapted for this study that covers a variety of personal and family concerns. Qualitative interview formats were also developed and used to guide in-depth interviews with program staff and former clients regarding their perspectives and experiences with the Discoveries program.

Both treatment and comparison subjects were recruited for the evaluation using detailed protocols specific to each sample and referral source. Written consent was sought from parents prior to inviting participation from youth. Informed written consent from youth was sought by research assistants in private interviews; participation in the study was not a requirement for participation in treatment. All subjects received gift certificates for local stores and restaurants of \$ 5 - \$ 25 as compensation for each completed interview.

The project had contact with and/or received data in regard to 102 adolescents during the 15-month data collection period. Of the 102 girls, 45 were from Discoveries, and of those, 35 participated in research interviews. The remaining 57 girls were comparison subjects, and of those, 52 participated in research interviews. In total, 257 interviews were conducted with 87 interview participants. Comparison subjects were drawn from several sources, including the Juvenile Court Office in Linn and Johnson Counties, Foundation 2 family services and Youth Shelter, and through networking among youth in the Cedar Rapids (CR) and Mount Vernon (MV) areas.

For some analyses, comparison subjects were assigned to one of two subsample comparison groups. Subjects assigned to the “comparison-services” group were those who had a case record with Juvenile Court Services (regardless of the recruitment category) and/or who reported having received any kind of services during the three months prior to their baseline interview. The latter included anyone who reported being involved in self-help groups or personal counseling, or who reported having been in a less than positive residential situation, such as juvenile detention, on the run, or homeless. Subjects assigned to the “comparison - no-services” group appeared to have been living in a stable home situation for the most part with reportedly no services or system involvement. The criteria used were later supported by attitudinal and other differences in the samples.

The primary focus of the data analysis for this report was on identifying and testing the significance of any observed changes on key measures for the Discoveries girls between T1 (intake) and T3 (end of aftercare, essentially the end of treatment contact), and occasionally, when the results were interesting or significant for those, T1 and T4 (follow-up). A second focus was on between-groups means testing. In some cases, independent samples tests were conducted simply comparing the Discoveries and comparison girls on their scores at T3. More often, change scores for each subject were calculated, and then the independent samples tests were conducted comparing the differences in change score means (average amount of change). Small sample sizes were sometimes a concern in the analyses.

Results showed that the Discoveries clients made statistically significant improvements in a number of areas while they were in treatment, in contrast to the comparison groups over the same time periods. There were also a few areas in which no change was apparent or the girls appeared to lose ground. The primary client outcomes in the area of school performance were: a) improved school attendance during treatment; and b) temporary focus on school as important/meaningful. In the area of empowerment: a) increase in understanding of personal identity, goals and values; and b) improvement in self-confidence

and self-image. In the area of prosocial behavior: a) decrease in impulsivity; b) increase in overt compliance with rules and the law; c) less rebellious or oppositional behavior; and d) little or no improvement in traditional “prosocial” attitudes. In the area of relationships: a) little change in peer relationships, possibly more insecurity; b) little change in romantic relationships; and c) somewhat less conflict with parents and other family members. In the area of health: a) temporary restriction in alcohol use, increase in tobacco use; b) no change in attitudes toward substance abuse or sexuality; c) no change in physical activity, safety, or self-care activities; d) improved goal-setting & decision-making; e) increased knowledge of strengths, skills, career choices; f) enhanced body image & reduced eating dysfunctions; g) reduced depression & hopelessness; and f) enhanced self-confidence & personal belief system.

Overall, the study demonstrates that the Discoveries program is producing good outcomes and should retain and build on its strongest program components, including: a) individual clients assigned to specific full-time staff member with 24 hr-7 day access by cell phone; b) intensive group sessions with daily/weekly topics planned for each cohort of girls; c) use of outside speakers and organizations from the community; d) strong, trained, committed predominately female staff; e) therapeutic issues of specific relevance to girls; e) relationship-building within the group, among/between girls and staff; f) full-time school liaison staff position maintains communications with the schools, ensuring successful experience for each girl in her home school while in treatment; g) special policies regarding clothing, personal space, and interpersonal communication; h) family involvement; and g) behavioral monitoring and consequences on and off site.

Key recommendations for enhancing the program are to: a) reinstitute regular spirituality components; b) add a fine arts component as an aspect of spirituality enhancement; c) require training in mental health diagnoses and treatment for staff; d) purchase on-site psychiatric nurse visits for staff consultation and diagnostic purposes; e) offer a physical self-defense class once in each cohort; f) experiment with a new component that involves girls in gender- or race-based community action (i.e., Women for Peace, or a project to change a school policy on athletics) that gives them experience in community activism (appropriate “rebelliousness”), volunteerism, group belonging, etc., while working to end discrimination or some other just cause; g) develop an *ongoing* (long-term/two years) weekly aftercare support group for all former clients, with incentives to attend (i.e., paid stipends and/or court orders) to encourage retention of program outcomes; and h) to establish new program development mechanisms, such as: 1) integrated planning sessions; and 2) annual staff retreats, the first one to be facilitated by an outside person, agenda to be “visioning”, program planning, staff renewal, etc.

Other program planners and policy-makers are encouraged to review the program components and recommendations offered for Discoveries in facilitating the development of new programs and services for girls involved in the juvenile justice system. Other recommendations in general include to attend in a new way to sexuality and more to mental health issues for teen girls, to move forward with promoting and developing gender-specific approaches to meeting the needs of girls before and after they enter the system, to plan for the next phase of organizational change after new non-traditional programs have been in place for a couple of years, and to avoid seeing compliance as a primary goal for girls growing up in today’s society.

## **Discoveries Gender-Specific Day Treatment Program Evaluation Final Report**

### **I. BACKGROUND**

Great strides have been made in recent years to finally address the gender-specific needs of girls and women in the health, human service, and justice systems around the country. The federal government has at times been a leader in this process, through its Women's Health Initiatives at the NIH, the Office of Women's Health at the Department of Health and Human Services, the cross-agency Health and Justice funding of the GAINS Center for Co-Occurring Disorders in the Justice System, and other avenues. Among these steps forward, the Office of Juvenile Justice and Delinquency Prevention's (OJJDP's) State Challenge Grant Program, which provides incentive funding to States for the improvement of their juvenile justice systems, now offers states the option to use these grants to combat gender bias against girls and otherwise improve gender-specific policies and programs in the system. With more females than ever before entering justice systems that were set up and continue to be structured toward males, the need is clear and growing.

Iowa was among the States that opted to use Challenge Grant funding to address the specific needs of girls in the juvenile justice system beginning in 1995. The Iowa Juvenile Justice Advisory Council, working through the Division of Criminal and Juvenile Justice Planning (CJJP) and the Iowa Commission on the Status of Women, initiated and uses Challenge Grant funding to support the Iowa Gender-Specific Services Task Force as well as several other program, training, and awareness activities related to improving the system's response to female juvenile offenders in Iowa. Among those activities, CJJP offered funding in 2001-2002 for the evaluation of one or more established gender-specific treatment programs for adolescent females in the state. A collaborative application from Foundation 2 and Service Evaluation & Development Associates, Inc. (SEDA) was selected for funding to carry out an evaluation of the Discoveries Day Treatment Program in Cedar Rapids, Iowa.

Discoveries is a day treatment program operated by Foundation 2 in Cedar Rapids that is designed as a gender-specific approach to providing therapeutic services and supports to girls between the ages of 14 and 17 who are at risk of being placed out of the community because of emotional and/or behavioral problems. Most clients are referred by the Juvenile Court Office and/or the Department of Human Services. They typically attend the program for 25 hours per week after school Monday through Friday for a 12-week intensive phase of treatment and then participate in aftercare sessions for another six weeks. The program has been operational since 1997. Charlotte Bright, MSW, is the current Program Coordinator. Ms. Bright assisted in the design and implementation of the evaluation, providing consistent support and cooperation to the research effort throughout the project.

Foundation 2 is a not-for-profit human service agency serving people of all ages in Cedar Rapids and surrounding counties since 1970. The agency offers a 24-hour crisis hotline, a youth shelter, individual and family counseling, and several outreach and after-school programs for vulnerable youth. Discoveries is one of two day treatment programs operated by Foundation 2; the Crossroads program is available as well, serving both boys and girls, ages 11-15. Transitional and independent living programs, street outreach to teens, and runaway assistance to families are among the services provided to help keep kids safe and off the streets. Support groups of various kinds, including for suicide bereavement, for

depression and anxiety, and for partners of men in prison, are also offered. Foundation 2 has a tradition of partnering with other organizations to accomplish its work and has close ties with many other community groups and agencies in the area.

Robert Hintz, Ph.D., Program Director, oversees Discoveries and other programs for the agency. As such, he held the lead administrative role in the evaluation project, providing primary oversight for the grant funding and project implementation. In addition, Dr. Hintz supported the project in many ways, by participating in the evaluation design, serving as a liaison with Juvenile Court, brainstorming solutions to technical difficulties, and demonstrating a consistent interest in the research and its potential for improving the program over time.

The research organization, Service Evaluation & Development Associates, Inc. (SEDA), is a small consulting firm based in Mount Vernon that provides evaluation, grant-writing, and related services to agencies across the state in human services, substance abuse, mental health, and corrections. SEDA is involved in a variety of activities, including ongoing evaluation planning and assistance for the Iowa Federation of Families for Children's Mental Health, a federally-funded Drug Free Communities program, and a substance abuse county jail diversion program. Evaluation of a community tobacco prevention project was recently completed.

SEDA was founded by Anita M. Patterson, MSW, in 1999. Ms. Patterson served for nine years as the Associate Director for the Iowa Consortium for Substance Abuse Research and Evaluation, a statewide organization based at the University of Iowa that promotes collaboration among academia, state agencies, and the field in conducting interdisciplinary research and evaluation in substance abuse and related areas. As part of her role at the Consortium, Ms. Patterson interacted with state agencies and local providers on a regular basis, including youth-serving agencies and CJJP. Among the research projects that she was involved in were an evaluation of women's specific aftercare services in 13 substance abuse treatment programs, evaluation of two substance abuse treatment programs for mothers and their children, and NIJ-funded research on the associations between addiction, partner violence, and child abuse. She also participated in the design and the evaluation of the 99-00 Iowa Youth Survey project.

As Research Director for this project, Ms. Patterson's role was to lead the design and implementation of the evaluation, involving the agency and other partners in various aspects of the project as appropriate. Ms. Patterson held the primary responsibility for the research design, instrument development, data collection, data management, data analysis, and interpretation. She supervised SEDA staff in carrying out these functions, and she was the lead author for all of the quarterly progress reports and for this final report. The entire final report and executive summary is posted on SEDA's web site at [www.sedainc.org](http://www.sedainc.org).

Foundation 2 was encouraged to apply for the evaluation grant by Juvenile Court Services, as an agency that makes use of the Discoveries program and has staff that hold an interest in gender-specific programming. Staff from JCS met with the evaluation team early in the project to offer feedback on the research design, assisted at times with contact information on lost research subjects, and later provided appropriate access to juvenile court data for analysis. For a short time, the Circle Program operated by the Juvenile Court Office in Johnson County was also involved in the project as a source of comparison subjects, before funding was cut for the staff member who facilitated the program. The Foundation 2 Youth Shelter was a third supporting entity in the project, referring potential comparison subjects to the study throughout most of the past year.

The following narrative provides an overview of the Discoveries program, describes the research design and protocols used for the evaluation, and describes the instruments, data sources, and subject samples. Baseline descriptive data for the Discoveries and comparison samples is provided, followed by the results of the statistical analyses undertaken to assess program outcomes. The latter are organized according to the domains embedded in the research questions originally posed for the evaluation: school/work, empowerment/self-efficacy, prosocial attitudes & behavior, relationships, and health/mental health. Conclusions and recommendations are offered by way of summary at the close of the document. Appendices include a summary chart of study outcomes as well as various research materials used in the project.

## II. PROJECT DESCRIPTION

### A. Program Overview & Activities

The Foundation 2 Discoveries Girls Day Treatment Program was opened in 1997, and is currently located in a small house on 1<sup>st</sup> Avenue NE in Cedar Rapids. The program includes a 12-week intensive component followed by at least six weeks of aftercare; occasionally girls may repeat the intensive or after care program phases if doing so is deemed most appropriate to meet her needs. During the intensive phase, the girls are transported to the house after school and remain until approximately 8:30 pm, each weekday. During the aftercare phase, they attend group sessions at the house two or three afternoons each week. Clients are admitted in cohorts, every 12 weeks, comprised of 5-8 girls each. The program accepts high school girls, ages 14-17, who are typically referred by Juvenile Court Services and/or the Department of Human Services. Referrals are generally girls who would otherwise require residential treatment or are in transition back to the community and their homes following residential placement.

In-house program staff include the Coordinator, two case managers, the school liaison, the aftercare program coordinator, and a part-time counselor. The Coordinator and two case managers design and implement most of the direct program services each evening, and in addition, the case managers maintain contact with each client and her family over the weekend. The program also employs part-time youth trackers and part-time therapists who meet with the girls individually.

The design, philosophy and components of Foundation 2 Discoveries were developed according to "Providing Gender-Specific Services for Adolescent Female Offenders: Guidelines and Resources" (1999), published by the Iowa Gender-Specific Services Task Force. Discoveries' structure was designed to address the 11 Components Of Gender-Specific Programming For Girls recommended in that publication. Program materials denote the following service components, activities, and principles:

To provide positive female role models and mentors:

- Outside speakers

- Youth conferences

- Women's history activities

- Posted quotes from positive women

- All Discoveries staff and mentors are women

To address abuse issues:

- Boundaries recognition activities
- Respect for personal space
- Assertiveness skills training to help girls gain more control and confidence in acknowledging their right to their own bodies
- Individual therapy and skill development
- Referrals and transportation to and support during health care appointments

To address sexuality, including pregnancy and parenting:

- All information is provided in a way that does not assume heterosexual relationships
- Candid discussions about sex, pregnancy, risks, etc. in individual and group sessions
- Speakers from area agencies, including Planned Parenthood, to discuss HIV/AIDS and other STDs, birth control, general sexual health and body development
- Young Parent's Network presents a 4-6 session series focusing on postponing pregnancy
- Sexual harassment and gender role expectations are openly discussed

To promote physical and emotional safety:

- Criminal background and child abuse registry checks
- Staff does not physically restrain a girl unless she poses an immediate physical danger to herself or others
- All staff members are trained in Crisis Prevention Intervention (CPI) techniques
- Supervision is provided throughout the program to assure physical safety, including keeping sharp objects in the staff office
- Confidentiality and respectful behavior are mandatory
- Group sessions focus on appreciation of everyone's differences as well as developing sensitivity to others' interpretations of comments, jokes, etc.
- Posters that support safe, open-minded, respectful behavior

Other program components:

- Skills groups that focus on topics such as seeking, interviewing for, and maintaining employment
- Transportation to assist the girls in seeking job opportunities
- Exercise such as walking, aerobics, outdoor games, yoga, tai chi, etc.
- Menu planning and shopping for healthy foods
- Journaling and inspirational reading
- Spirituality group
- Relaxation time for participants to unwind and regroup
- Monthly Family Nights including supper and fun activities
- Family therapy is provided upon request
- Diversity activities
- Discoveries newsletter, *Seasons*, written by the girls

Helping participants understand and overcome stereotypes is an important element of the Discoveries philosophy. Through both formal and informal discussions (some with outside speakers) societal stereotypes are examined, including how girls may have been victims of stereotypes, how this victimization may affect them now and in the future, and how to constructively overcome stereotyping. All discussions focus on helping the girls empower themselves.

Each day when girls arrive at the Discoveries program, they must remove all make-up and jewelry and change into program-provided clothes. The philosophy behind the development of this component was to ensure that the girls are not distracted by their own stereotypes of each other based on appearance, to remove their figurative “masks,” and to see themselves as an individual member of a greater sisterhood.

The core of Foundation 2 Discoveries is relationship-based services. The program is structured so that each rotation of girls is able to develop a positive, healthy, supportive relationship with each other and with the staff. Each girl is assigned an individual case manager with whom she meets once a week or whenever she needs to talk. Staff members carry their own program cell phones in order to deal with crisis situations after hours or on the weekends. To ensure the privacy and confidentiality of conversations only digital cellular phones are used.

Group discussions focus on relationships, including dating, family, peer and community. Specific groups guide girls in learning skills of assertiveness, appropriate expression of anger and selection of healthy relationships.

Further detail on program design and components is offered from the perspective of individual staff and clients in Section V.

## **B. Evaluation Design**

The **research design** for the Discoveries evaluation was quasi-experimental, using a pre-post treatment design with comparison groups. Pretreatment measures were administered at intake for all clients who consented to participate in the evaluation. The same measures were administered again at the end of their intensive phase of treatment (typically 12 weeks), at the end of their aftercare phase of treatment which also marked the end of their treatment episode altogether (typically 6 weeks after the end of intensive), and again at 8-10 weeks follow-up. For the treatment group, the interview points were referred to as intake, discharge, aftercare, and follow-up, or Time1, Time2, Time3, and Time4. In the event a client repeated a given phase of the treatment, the interview for that point was repeated, and the original interview data for that point was discarded. (This means, for example, that the period between T1 and T2 for a few clients is actually 18 weeks, rather than the 12 typically completed in the intensive phase.) The same measures were administered to comparison subjects 4 times corresponding to the same time intervals typically experienced by the clients in treatment. The T2 interview was completed approximately 12 weeks after the initial interview, T3 approximately 6 weeks after that, and T4 approximately 8-10 weeks after T3.

As stated in the original grant application, *“the overall goal of the research project is to determine the impact that the Discoveries program has had on the girls it serves. **Research questions** will be as follows: As compared to girls who do not complete Discoveries or attend other programs, to what extent do the clients graduating from Discoveries experience a positive change between placement date, interim/discharge, and follow-up in: a) school and/or job performance? b) prosocial attitudes and behavior, including reduced involvement with Juvenile Court? c) perceptions of self-efficacy, assertiveness, and empowerment? d) satisfaction in relationships with family members, other adults, and peers?; and e) health and self-care, including avoidance and/or reduction in substance use & unsafe sex? The project will also attempt to answer the question: To what extent were program factors associated with observed client outcomes?”*

Some modifications in the original design and protocol occurred as the project evolved. For example, the number of girls who dropped out of the program during the study period was very small and therefore was not used as a comparison group in the analysis. However, for girls who did complete treatment and two untreated comparison samples, quantitative data is available for every question to a large extent, as outlined in Section IV below. The extent to which outcome results can be attributed to the program was also investigated to some extent, in part by documenting the number of individual and group sessions each client participated in during her involvement in Discoveries. This data is also reported in Section IV. In addition to outcome evaluation and quantitative process evaluation, the project included a qualitative process component that attempted to document from staff and client perspectives the extent to which the program's actual day-to-day operations matched the recommended guidelines for gender-specific programs. These results are reported in Section V.

The **instruments and measures** selected to answer the research questions reveal how the questions were operationalized during the first two months of the project. The Research Director led this process, working with program administrators and JCS, and interviewing staff about program objectives, in order to end up with instrument(s) that would measure what the program actually was trying to accomplish, do so in a valid and reliable manner, and do so without undue burden on the program staff, the clients, or the research assistant(s). An overview of the instruments, domains, and sources is provided in Table 1, followed by additional details for each research instrument and/or data collection form.

**Table 1. Data Collection Instruments**

<b>Instrument or Source</b>	<b>Domain</b>	<b>Author</b>
Admission form & spreadsheet	Demographics, service dates	Discoveries
Session forms	Counseling hours & topics	Patterson
Coalition Admit & Discharge forms	Personal & family information	Coalition for Family & Children's Services in Iowa
Child & Adolescent Functional Assessment Scale (CAFAS)	Current functional abilities & needs	Kay Hodges, Ph.D.
JCS database	Criminal activity & system involvement	Juvenile Court office
Millon Adolescent Clinical Inventory (MACI)	Personality & clinical change	Theodore Millon, PhD & NCS Assessments
Testwell Wellness Inventory for Adolescents	Health, sexuality, spirituality, nutrition	National Wellness Institute
Youth Questionnaire	Miscellaneous, inc. drug use, crime, services involvement, personal issues	Patterson
Qualitative Interview Formats	Program assessment by staff & clients	Patterson & SEDA staff

Completed copies of the first four instruments listed on the chart were made available to the evaluation team in hard copy form by the Discoveries Program Coordinator for all of the girls who entered the program during the study period, whether or not the girls were study participants. Since subject consent is

not required for secondary analysis of program data by qualified researchers, provided that precautions are taken for the protection of the subjects and the data, this data was transferred for all clients in the hope that analyses could be undertaken comparing program drop-outs to completers. While the eventual sample size did not allow this level of analysis to occur, the treatment sample size in general for certain variables was a little larger as a result of having included all clients, which was helpful.

The second four instruments/sources on the chart provided data for both the treatment sample and the comparison sample. This allowed analyses to be undertaken comparing the two groups on the same measures over time. In the case of the JCS database, specific data was requested and provided by the JCS office data manager to the Research Director for those girls in the study who had records in the system. The other three instruments were administered multiple times to the treatment and comparison girls who were recruited to participate in the study by SEDA Research Assistants.

The final item listed on the chart is Qualitative Interview Formats, which are the two sets of questions that were used to guide qualitative interviews with the program staff and with clients, respectively. These were not instruments to be completed by subjects, with the exception of a one-page segment of the interview format for staff.

**Admission form & spreadsheet.** The Admission form is the existing format used by the Discoveries Program Coordinator to document information she collects during the intake process for each new client. The instrument is in hard copy only and most of the information collected is in a qualitative format, except the face page data, which is largely demographic, including family names and residency information. Each girl's form was copied and mailed to the Research Director (in groups of forms on a periodic basis). The program spreadsheet is a basic services dataset maintained by the Program Director that includes admission/readmission and discharge dates for each client, along with discharge status (successful, unsuccessful, partial). This file was transferred electronically to the Research Director as needed.

**Session forms.** The Research Director developed Individual and Group Session forms (included in the Appendix) for the program to use as a measure of the nature and volume of services provided to clients during their time in treatment. While the agency already recorded the number of service hours by treatment level for cost reimbursement purposes, the majority of treatment at Discoveries was delivered either in individual or group sessions which varied in nature from one day, one week, and one cohort to the next. The session forms therefore were created so the Program Director could track the number of minutes/hours clients spent in each session each evening and the topics covered. This data was then used by the study as the primary measure of treatment volume for each subject.

**Coalition Admit & Discharge forms.** Foundation 2 is among the agencies in Iowa that participates in a statewide data collection activity overseen by the Coalition for Family & Children's Services in Iowa. The data collection involves completion of standardized admission and discharge forms that include questions on referral source, adjudication status, child characteristics/ problems, family characteristics/ problems, prior services and placements, and discharge status and placement. Many of the questions, such as those regarding child and family characteristics, are formatted as standardized lists of items from which multiple answers can be checked. The forms are completed by staff based on their knowledge of the child and the family at the time of admission or discharge. Copies of these forms were provided periodically to project staff along with the other client records.

**Child and Adolescent Functional Assessment Scale (CAFAS; Hodges, 1997).** The CAFAS is an instrument that allows staff to record their impressions of the current functional abilities and needs of youth. Some of the Discoveries staff have been trained in its use, and the instrument is included in the Coalition data collection process. Staff complete the CAFAS instrument at admission and discharge and then the scores are recorded on the Coalition Admit and Discharge forms. CAFAS scores can be used to measure client change over time and thus were included as outcome measures for the evaluation.

The CAFAS is the most widely used and well-known functional assessment instrument for children, having been adopted by more than twenty states for use in service and funding decisions. The instrument is intended to be completed by a trained worker selecting a series of statements describing various levels of functioning in several domains: Role Performance (school, home, community), Behavior Toward Others, Moods/Self-Harm, Substance Use, and Thinking, leading to five or eight scale scores (depending on the version used) as well as a total score. Caregiver resources and problems can also be assessed in the areas of Material Needs and Family/Social Support. Optional items are available for noting strengths and goals, although some of the strengths are defined in terms of lacking a given problem.

Some studies have been undertaken assessing the psychometric properties of the CAFAS, including studies investigating concurrent validity (with the Child Behavior Checklist and other tools), criterion-related validity, validity in predicting services cost and utilization, test-retest reliability, and interrater reliability (Hodges & Wong, 1996; Quist, 2000). However, the evidence supporting the instrument is modest and no studies of the content and structural validity of the tool have been published (Bates, 2001). The widespread use of the instrument in spite of the modest empirical support can be explained by the field's increasing interest in and desperate need for functional assessment tools that are easy to administer, specific to youth, cover multiple life dimensions, and can be used to guide and evaluate service decisions.

**Juvenile Court Services (JCS) database.** The JCS database was the source of independent measures of criminal activity and system involvement for subjects involved in the study. After proper permissions were obtained, the data manager in the Cedar Rapids JCS office provided the Research Director with printed hard copy Case Summaries for all research subjects on file. In addition to basic family and residency information, the Case Summaries included police reports, charges, and dispositions, including placements and placement dates on record. The primary concern with the JCS data was how current it was, because it depended on updates from busy JCOs, which in fact proved to be a problem in using contact data for subject recruitment and tracking. However, it was important to have measures of criminal involvement, such as number and dates of charges filed, that did not rely on subjects' self-reporting.

**Millon Adolescent Clinical Inventory (Millon & NCS Pearson, Inc., 1993).** The Millon Adolescent Clinical Inventory (MACI) is a 160-item self-report personality and clinical assessment instrument that served as the core outcome tool for the evaluation. The instrument is grounded in a 30+ year history of scientific theory and research on adolescent development and behavior (Millon, 1969). Its primary author, Theodore Millon, Ph.D., has been among the foremost authorities involved in the classification of personality types and psychiatric diagnoses, including the DSM-III and -IV (Millon, 1991). The original Millon Adolescent Inventory was published in 1974, followed by the Millon Adolescent Personality Inventory (MAPI) in 1982. The MAPI was subsequently divided into two forms, the MAPI-C(linical), designed for mental health workers, and the MAPI-G(uidance), for use in school settings. Following 10 years of feedback from clinicians using the MAPI-C, an extensive MACI development project was launched in the early 1990's that included several phases of instrument construction/revision, scale definition and base rate transformation development, and cross-validation testing. A total of 1,107

adolescents and their clinicians from the U.S. and Canada participated in the original validation studies for the MACI.

The results from those initial studies showed internal consistency reliability (alpha) coefficients that tended to range from around .70 to .90 for each of the subscales. Test-retest correlations ranged from .57 to .92 (most in the .80 range), with retests occurring 3-7 days apart. Validity was also tested in various ways, primarily by comparing MACI scale scores with clinician judgments and to scores on other well-known instruments such as the Beck Depression Inventory and the Problem Oriented Screening Instrument for Teens. In the first sample of clinician judgment comparisons, 14 of the 25 correlation coefficients were statistically significant ( $p \leq .05$ ) and in the second sample, 20 of 25 were. There was a very wide range of results in comparing other instruments to the various MACI scales, but some were quite high for those that measured very similar constructs (such as depression and eating dysfunctions).

The MACI has a number of distinguishing features that made it on balance the best instrument for use in this evaluation. First and foremost, it is extremely important to have at least one well-validated instrument that measures some of the core changes expected of clients in the program being evaluated. Unless those changes are very straightforward (i.e., how many days did you use alcohol in the past 30 – versus - how self-confident are you feeling now?), the validity of the instrument will be crucial to showing any results at all, particularly if the sample size is small. Second, unlike many, in fact most, of the clinical assessment tools available today, the MACI was designed specifically for adolescents, and actually was normed on both boys and girls. It is not an adult instrument that was reworded for use with kids, or one that was normed on boys and then used for both. Third, the MACI is grounded in a comprehensive theoretical model, which guided its forerunner development and subsequent revision, as well as its scale development. A variety of professionals who work with adolescents were extensively involved in all phases of development. Fourth, unlike its forerunners, the MACI is designed for use with troubled adolescents in clinical, residential, and correctional settings; and it may be used for diagnosis, treatment, and most importantly here, for outcomes monitoring. Fifth, item content and scoring takes into account potential sources of respondent bias as well as the actual rates of distribution of various personality patterns and clinical disorders in the adolescent population. And finally, the MACI offers efficient administration and a broad range of content in a relatively brief format.

Of course, the MACI does have limitations as well. The instrument is not in the public domain and with a larger sample, the cost would have become prohibitive in a publicly-funded project of this sort. Secondly, although little training was needed for actual administration of the instrument, the scoring procedures were complicated and time-consuming. (Computerized scoring was available at additional cost.) The purchase of the instrument in fact requires documentation of professional qualifications. As far as the content of the instrument itself, the manual claims that the MACI can aid the clinician in assessing strengths as well as weaknesses, but it's not clear how the instrument itself is supposed to facilitate that because all of the scales and scale definitions are framed primarily from a deficit/dysfunctional perspective. Perhaps that information is provided in the computerized scoring and interpretations.

As noted, the MACI is a 160-item inventory that is structured into multiple subscales: 12 Personality Patterns, 8 Expressed Concerns, and 7 Clinical Syndromes, plus 4 Modifying Indices. The latter Modifying Indices include 3 Adjustment scales (Disclosure, Desirability, and Debasement) and 1 Reliability score, that are used in the scoring procedures to adjust scores for subjects who have a tendency to deny or exaggerate their problems or answer randomly. An Anxiety/Depression adjustment is also made if the subject scores in such a way as to suggest their test results may have been affected by anxiety

or depression that day. Scoring also involves taking raw scores and transforming them into base rate (BR) scores that differ according to the age and gender of the subject, which were derived from and use the distribution of each of the personality and clinical patterns in those populations of troubled adolescents. A final BR score of 60 or above is notable, with the greater the score magnitude above that, the greater probability being that the adolescent has the personality or clinical pattern measured by the scale. Cut points of 75 and 85 are used to indicate a high probability of “presence” and “prominence” respectively, and again, the higher the score, the greater the “intensity or severity” is expected to be. In a clinical situation, the scale scores are not intended to be reviewed independently for any given individual, but are to be interpreted as part of an overall profile, and then only in the context of an overall assessment. The MACI’s subscale structure is shown below.

#### MACI Personality Patterns:

Introversive (apathetic, diminished capacity to be motivated by either pleasure or pain),  
Inhibited, Doleful (similar to depressed), Submissive, Dramatizing, Egotistic, Unruly, Forceful,  
Conforming, Oppositional, Self-Demeaning, and Borderline Tendency

#### MACI Expressed Concerns:

Identity Diffusion (unsure about self, goals, values), Self-Devaluation, Body Disapproval, Sexual  
Discomfort (sexual thoughts/feelings confusing or disagreeable), Peer Insecurity, Social  
Insensitivity, Family Discord, and Childhood Abuse.

#### MACI Clinical Syndromes:

Eating Dysfunctions, Substance Abuse Proneness, Delinquent Predisposition, Impulsive  
Propensity, Anxious Feelings, Depressive Affect, Suicidal Tendency.

Discussion of the individual scales will be offered as needed when results are presented in the report.

Unlike most of the other instruments used in the evaluation, the MACI will not be found in the appendix. None of the MACI materials can be reproduced because of copyright restrictions.

**Testwell Wellness Inventory for Adolescents (National Wellness Institute, 1999).** The Testwell is a self-scoring assessment tool based on a six-dimension theoretical model of wellness originally developed by Bill Hettler, M.D. in 1979. The six dimensions of wellness posited by Dr. Hettler are Occupational, Intellectual, Spiritual, Social, Physical, and Emotional, with the six areas being seen as interdependent, and balance among them as the goal. The National Wellness Institute offers a number of Testwell assessments for different target groups and settings, with paper and pencil versions as well as on-line administration and scoring. The 50-question (short) version for adolescents was selected for the study, which includes 5 questions in each of the following dimensions: Physical Activity, Nutrition, Self-Care, Safety, Social & Environmental Wellness, Emotional Awareness & Sexuality, Emotional Management, Intellectual Wellness, Occupational Wellness, and Spirituality & Values.

Some of the Testwell scales are a bit more self-explanatory than others. For example, depending on the version used, Self-care items on the Testwell may include getting enough sleep, using sun screen, and flossing your teeth, whereas safety items involve wearing seat belts, using safety equipment when needed, and not riding with drunk drivers. In the area of Social & Environmental Wellness, items query respondents about recycling and resolving conflicts, whereas in Emotional Awareness & Sexuality, questions focus on interactions with others, being tolerant, and safe sex. In the Emotional Management

section, on the other hand, subjects rate themselves on being able to set goals and the extent to which they enjoy life. Intellectual Wellness is about staying informed about current events, reading, and making decisions, and Occupational Wellness focuses on such things as being aware of one's own skills and having good work habits.

Subjects are asked to rate the frequency with which they engage in or experience each of the items on a scale of 1 to 5, 1 being Never or Almost Never and 5 being Always or Almost Always. All of the items are framed in an affirmative format such that scoring is straightforward and simply a matter of adding up the ratings for each subscale and then totaling the subscales. The Testwell is copyrighted, but sample versions can be found at [www.testwell.org](http://www.testwell.org).

If there are psychometric studies documenting the reliability and validity of the Testwell, they do not seem to be promoted by the Institute. However, the Testwell appears to be in popular use and more importantly, offered a number of advantages for this project other than scientific rigor in that it covered many of the key program areas in a brief instrument that could measure change over time, was relatively inexpensive, and was easily administered, scored, and analyzed. Although there were some communication problems with the company, overall the instrument proved to be a worthwhile component of the evaluation.

**Youth Questionnaire.** This is an instrument authored previously by the Research Director and modified for this project that gathers information from youth on their own perceptions of their current circumstances, strengths, and problems. An abbreviated scale measuring Family Cohesion is included, with items drawn from other validated instruments in the field measuring family cohesion and family conflict. In addition to asking youth how many hours per week they spend in school and work activities, they are also asked to rate how often their school work is meaningful and important, and how often they cut classes. From the perspective of the past 3 months, subjects check off which of a long list of stressful problems they have experienced, which living arrangements they have been in, and which activities and services they have been involved in, including church, self-help groups, and extramural activities. A set of questions patterned on the federal and state substance abuse surveys measures youth attitudes (how wrong it is for someone your age to use x), substance use (number of days in the past 30 for each), and peer association (number of friends who use x). Finally, a set of items developed for Discoveries uses a rating scale from much worse (1) to much better (5) for youth to assess the extent to which they experienced changes in each of seven areas during the past three month period. The seven items are intended to address key Discoveries treatment objectives in plain language: getting along with parents and teachers; expressing anger in healthy ways; taking care of self physically; feeling good about myself as a person, that I am worthwhile and capable; feeling good about myself as a girl, that I have the right to speak for myself and pursue my own goals; making/having close friends; and healthy romantic relationships. All girls in both the treatment and comparison groups answered these questions at all time points, not just the Discoveries clients.

**Qualitative Interview Formats.** In order to collect qualitative information about the program, interviews were conducted with program staff and clients using standardized formats developed for that purpose. The format for interviewing the girls was simply a set of interview questions designed to elicit information from the clients' perspectives and experience with various aspects of the program. For example, the girls were asked what their goals were and to what extent they were met, how issues of physical health, mental health, sexuality, and abuse were addressed, how feelings were expressed, how safe and supportive the environment was, how differences in background among the girls were addressed,

etc. All interviews were audiotaped (with permission) and transcribed for subsequent review by the Research Director.

The format for interviewing the staff was a bit more extensive, and included a one-page paper-and-pencil instrument form that each staff member completed as part of the interview. The general format was used as a guide for discussing the program with the interviewees, not as a list of questions to be asked, and therefore not all questions were asked of all participants. The questions included items pertaining to the goals and strengths of the program, ideas for improvement, staffing by gender, continuing education, relationship-building with clients and among the staff, family involvement, racial/ethnic diversity, self-disclosure, emotional and physical safety, and other program issues. The one-page instrument included three items: a) a list of program topics, which the respondents were asked to rank order according to the amount of time spent by the program on each; b) a list of theoretical approaches to treatment, also to be rank ordered according to how much it was used by staff with clients in the program; and c) a brief set of statements describing experiences or attitudes toward sexism and gender discrimination, for which respondents were asked to rank their level of agreement or disagreement. All interviews were audiotaped (with permission) and transcribed for subsequent review by the Research Director. The staff qualitative interview format is included in the appendix.

### C. Sampling & Data Collection

Detailed protocols for subject recruitment and data collection were developed and modified during the first few months of the project and beyond as needed to address varying circumstances. At Discoveries, recruitment of girls for the study began in September, 2001, with the first new cohort of clients to enter the program after the start of the project. The **Discoveries protocol** called for the Program Coordinator to present the study to each potential client's parent(s) at the intake interview. Written documents explaining the study were created for this purpose that are consistent with federal guidelines on the protection of human subjects in social science research. Among other things, the Information Summary explained that if the parent consented to the daughter's participation in the study, the girl would still simply be invited to participate, not required or compelled to do so. If the parent did consent, s(he) signed the Parent Consent form, the form was transferred to SEDA, and a SEDA Research Assistant (RA) made arrangements to visit Discoveries during the girl's first week of treatment to meet with her individually and invite her participation in the study. A similar Information Summary and Youth Consent form were used to explain the evaluation study to the girls and obtain their consent. If they agreed, the RA administered each of the instruments to each girl, typically requiring approximately 45 minutes per interview. During the cohort's last week of the intensive phase of treatment or their first week of aftercare, the RA returned and repeated the administration of the instruments with each girl who remained willing to do so. A third administration typically took place during their last week of aftercare or the first week thereafter.

For follow-up, the RA staff contacted the girls by telephone approximately 8-10 weeks after their aftercare interviews were completed to make arrangements for conducting the follow-up interview. A number of guidelines were used to structure these arrangements. For example, no interviews took place in the home; all were set up in a public setting, usually a restaurant during low-traffic hours or the library. A few were done at the F2 Youth Shelter or other agency locations. Transportation was provided directly by the RA staff; since the primary interviewer in Cedar Rapids was visually impaired, it became most convenient for a second RA to do client tracking and scheduling, and provide transportation for both the interviewer and the subject to and from their homes and the interview locations in that community. Girls

or parents who were uneasy with any of the arrangements had the option to negotiate alternatives, but none did so other than specifying a preference for a given time or location.

Initially, girls who participated in a research interview received a gift certificate to a local retail store or restaurant in an amount between \$ 5 and \$ 15. When concerns later arose about sample sizes, the amount of the gift certificates was raised to \$ 25 for all samples for each interview. In addition, when the interview took place in a restaurant, snacks and soft drinks consumed by the girls were paid for. (Thus, that was the location alternative they most often selected.)

Despite the incentives and the girls' involvement in the system, tracking for follow-up was a challenging business. Given the relatively small number of girls attending the treatment program and the short project period compared to the length of the treatment, the study could not afford to lose any subject to follow-up. Yet the reasons for attrition abound in any study of this nature. The most common appeared to be transience in family living situations (i.e., disconnected phones), combined with difficulties in communications with busy JCS staff. Foundation 2 staff assisted in the tracking process whenever possible. A few girls placed in long-term residential treatment out of the community completed the instruments verbally over the telephone with the RA staff, and then received their gift certificates through the mail. All in all, the decline and attrition rates for the study, although frustrating, were reasonable for a field study involving troubled families (see Section III).

**Study protocols for comparison subjects** were developed and/or modified as the referral source for each sample was identified. Initially, the project plan had projected a large number of comparison subjects being referred by the JCS office, either through a screening process to be set up at the Juvenile Court office, referrals from individual JCS staff, or even as a list of names of other girls with characteristics similar to those being referred for treatment at Discoveries. Eventually a list of all current cases ages 14-17 in Linn County was provided by the data manager, and the RA staff used this list to attempt to recruit comparison subjects over the telephone and/or through the mail. Unfortunately, this effort was largely unsuccessful. The listed information did not typically result in a current parent contact, and when it did, presenting the relatively complicated study concept to parents not always trusting of the system (or simply how the RA came to have their contact information) was challenging. It did not help that the RA staff were not experienced in cold-call sales or recruitment. When parents did consent verbally over the telephone, they seldom mailed the written consent form back despite reminder calls to do so. All in all, this resulted in a very small number of JCS comparison subjects.

Some of the JCS comparisons that were recruited for the study actually came from the Iowa City office through another process altogether. Early in the project, it had been suggested by JCS that the project consider expanding to include the Circle program in Johnson County, a girls weekly support group facilitated by a staff member at the Juvenile Court office in Iowa City. This staff member agreed and worked with SEDA to establish a protocol for recruitment and data collection of Circle clients and other JSC girls from Iowa City. Several Circle program girls were in fact recruited and parental consent obtained through the joint efforts of the facilitator and an Iowa City-based RA, before the facilitator was abruptly laid off due to budget cuts. A subsequent effort to reestablish the project in the Iowa City office was not successful.

The barriers that arose in the early months in regard to recruiting comparison subjects led the project team to brainstorm new recruitment methods, which ultimately proved more successful and in fact led to the recruitment of two different samples that allowed analyses that would not have otherwise been possible.

These were: a) a group of girls who were similar to the Discoveries clients, including having some involvement in services and/or criminal activity, but not necessarily Juvenile Court; and b) girls who did not appear to have any evidence of serious personal or family troubles, and had not recently been involved in any sort of services or criminal activity. In addition to identifying new recruitment sources and methods, as will be discussed further below, the decision was made to increase the stipend amount for interviews to \$ 25 and to extend the project period by six months, although both would eventually cause a major strain on the project budget.

The first of the new comparison group subjects began to appear after a study protocol was established at the Foundation 2 Youth Shelter. Shelter staff were oriented in the study and materials, and began to present the Information Summary and Consent Form to parents when a 14-17 year-old female entered the facility. If and when parental consent was obtained, staff faxed or mailed the consent form along with a referral form developed for that purpose to SEDA. The RA then made arrangements with shelter staff to visit with the girl privately at the shelter, invited her participation in the study, and if she agreed, proceeded with the initial interview. Subsequent interviews depended on how long the girl remained in shelter care and if not, where she was placed after leaving the shelter.

Tracking forms were completed by each girl at each interview, beginning with the first interview for comparison subjects. The girls were instructed to only list individuals (workers, family members, and friends) that they would be comfortable with research staff contacting for information on their whereabouts. No information about the girls was to be shared or ever was shared with the contacts, other than the fact they were in the Girls Youth study with SEDA and had provided that person's name for tracking.

The second comparison sample, and other members of the first comparison group, were recruited through a "snowball" sampling technique initiated by all of the SEDA project staff in various locations. Personal contacts were made with adolescent girls in a low-income housing project in Cedar Rapids, in the Mount Vernon High School, and with all of the girls involved in the study to that point. Any girl between the ages of 14-17 was invited to participate in the study, and in addition to the standard incentives for their own participation, the girls were offered a \$ 10 gift certificate for every other subject they successfully recruited to participate as well. This process began to bear fruit quickly once it was fully underway, although by that point the time frames were such that a much longer project period would have been required to take full advantage of the process. Nevertheless, a large percentage of the comparison sample was recruited via this method.

The protocols for working with the network samples, particularly in Mount Vernon, were modified when needed to take into account special circumstances. For example, because Mount Vernon is a small community, the girls' participation in the study feels a little less anonymous than it might with RA staff that they were not previously acquainted with. Therefore, when an RA interviewed a girl with whom she was familiar, and in all cases in Mount Vernon, the instruments were not reviewed by the RA when completed. Rather, the girl was instructed to seal them in an envelope pre-addressed to a data entry clerk in Iowa City, allowing the girls a modicum of privacy in those interviews. The data entry clerk then notified the Research Director when those packets were received and if/when an instrument had not been fully completed by a given subject.

**Data transfer and management** overall was a major focus throughout the project. Careful procedures were developed and used for maintaining the privacy of the data provided by the subjects and for

preventing loss. Following each interview, all instruments and instrument packets were maintained in locked boxes or filing cabinets until it was transferred by hand to the Research Director in Mount Vernon (except the few items directly mailed to the clerk). The Director maintained a master list of all subjects, interview dates, referral sources, etc., and provided general supervision to the tracking, scheduling, data collection, and transfer process. After basic review of the materials, she sent them on to the data entry clerk in Iowa City using an overnight tracking mail service. Data was then entered by the clerk into database programs developed by the Director for that purpose. All hard copy materials were stored in locked filing cabinets in all locations. Electronic files were maintained using code numbers and passwords, as appropriate. Most data was entered into Excel database programs, and then later downloaded into SPSS 11.0 for statistical analysis.

The **qualitative interviews** with staff and clients were conducted using the Qualitative Interview Formats, as described above. Nine clients were independently recruited and interviewed by the RA staff to participate. All were in aftercare or had already graduated from the program at the time of their interview. Only one of the clients was interviewed individually; the others participated in one of three small group discussions facilitated by the lead RA using the client interview format. Each subject was informed in advance that the interviews would be audiotaped as well as how the tapes and information would be handled by the research team. All of the interviews were held in a private room at the Cedar Rapids Public Library. Transportation was provided by the RA staff, and each girl received a \$ 10 gift certificate for her participation. When the interviews were concluded, the audiotapes were marked by number and hand delivered to the Research Director, who in turn transferred them by overnight mail to the data entry clerk for transcription. The typed transcriptions and audiotapes were then returned to the Research Director for analysis.

Staff interviews were conducted by the Research Director using a similar process, except that each of the staff interviews were done individually and of course followed the staff qualitative interview format. Four of the five full-time regular staff were interviewed, including the Program Coordinator; one staff member went on maternity leave during the time the interviews were taking place. Each of these interviews typically lasted 60-90 minutes, and were also audiotaped and transcribed for qualitative review.

### III. STUDY PARTICIPANTS

#### A. Project Participation

The project had contact with and/or received data in regard to 102 adolescents during the 15-month data collection period. Of the 102 girls, 45 were from Discoveries, and of those, 35 participated in research interviews. The remaining 57 girls were comparison subjects, and of those, 52 participated in research interviews. As shown in Table 2, a total of 257 interviews were conducted with the 87 interview participants.

**Table 2. Interviews by Time Point**

<b>N of Unduplicated Subjects</b>	<b>102</b>
<b>N of Unduplicated Interviewees</b>	<b>92</b>
<b>N of Interviews</b>	
T1/Baseline/intakes	89
T2/discharge from intensive	75
T3/aftercare completion	61
T4/follow-up	32
<b>Total number of interviews conducted</b>	<b>257</b>

As discussed above, comparison subjects were drawn from several sources, including the Juvenile Court Office in Linn and Johnson Counties, Foundation 2 family services and Youth Shelter, and through networking among youth in the Cedar Rapids (CR) and Mount Vernon (MV) areas. Table 3 provides data on the number of interviews conducted with subjects from each source, as of the end of the study period.

**Table 3. Interviews by Referral Source**

	No parent consent	Parent yes Youth no	N of interviews				N Subjects
			1	2	3	4	
<b>Discoveries</b>	<b>3</b>	<b>7</b>	<b>6</b>	<b>7</b>	<b>11</b>	<b>11</b>	<b>45</b>
JCO Linn County		0	0	0	3	0	<b>3</b>
JCO Johnson County		3	0	1	0	4	<b>8</b>
F2 services/shelter care		2	5	1	4	3	<b>15</b>
CR network		0	1	6	0	5	<b>12</b>
MV network		0	1	0	13	5	<b>19</b>
<b>Total Comparisons</b>		<b>5</b>	<b>7</b>	<b>8</b>	<b>20</b>	<b>17</b>	<b>57</b>

Over half of the comparison subjects were still enrolled in the study at the end of the project period. The attrition rate for comparison subjects was also much lower than for the treatment sample, not surprisingly. Sixteen of the 35 Discoveries girls with consent dropped out of the study: eleven of them were lost during the treatment phase (31% drop-out rate), the other five after completing treatment and/or their 3<sup>rd</sup> interview at the end of aftercare (14% lost to follow-up). Of the 52 consenting comparison youth, 8 declined or were lost after completing only one or two interviews (15%) and 1 was considered lost to follow-up after completing 3 interviews (2%).

For some analyses, data on the comparison subjects was examined in order to assign them to one of two subsample comparison groups. Subjects assigned to the “comparison-services” group were those who had a case record with Juvenile Court Services (regardless of the recruitment category) and/or who reported having received any kind of services during the three months prior to their baseline interview. The latter included anyone who reported being involved in self-help groups or personal counseling, or who reported having been in a less than positive residential situation, such as juvenile detention, on the run, or homeless. Foster care was a judgment call, and was not included in this category. Living with

relatives was also not included. All comparison subjects from JCO and shelter sources were included in this “services” sample. Subjects assigned to the “comparison - no-services” group appeared to have been living in a stable home situation for the most part with reportedly no services or system involvement. The criteria used were later supported by attitudinal and other differences in the samples. The resulting sample sizes when aggregated by service type are shown in Table 4.

**Table 4. Sample Sizes**

Primary Samples	Records	Interview Participants
Discoveries	45	35
"Comparison-services" sample	35	30
"Comparison - no-services" sample	22	22
<b>Total</b>	<b>102</b>	<b>87</b>

The number of each sample having “records” refers to those having records in the master database for the evaluation even if they did not participate in the evaluation. The Discoveries girls, for example, had program data available even if they did not participate in interviews. The five comparison girls in the “services” sample

who had master records but did participate in interviews were those who were identified as potential participants for the study but did not consent to do so.

It is important to consider the extent to which the comparison sample is a good match for the treatment sample, because many of the analyses to follow will be based on the assumption that any differences observed in outcomes are in fact due to the treatment and not to differences between the samples. This is actually not scientifically accurate: theoretically there are many factors not even considered in this report that might be the cause of any observed differences in outcomes between the groups. However, given that we can only achieve that level of rigor with an experimental research design, which was not feasible in this case, we are left with making some assumptions and hopefully using some caution in doing so. With that in mind, the following tables offer some demographic data as a general overview of the three samples, and their differences and similarities.

**Table 5. Age at Intake**

Age at intake	Discoveries		Comparison-services		No Services	
	N	%	N	%	N	%
13	1	0.03	2	0.07	0	0.00
14	11	0.20	5	0.17	2	0.09
15	12	0.31	11	0.37	7	0.32
16	14	0.29	8	0.27	10	0.45
17	7	0.17	4	0.13	3	0.14
<b>Total</b>	<b>45</b>	<b>1.00</b>	<b>30</b>	<b>1.00</b>	<b>22</b>	<b>1.00</b>

**Table 6. Race/ethnicity**

Race/ethnicity	Discoveries		Comparison-services		No Services	
	N	%	N	%	N	%
African American	5	11.1	3	9.1	1	4.5
Biracial	7	15.6	1	3.0	0	0.0
Caucasian	33	73.3	26	78.8	19	86.4
Asian American	0	0.0	2	6.1	0	0.0
Hispanic	0	0.0	1	3.0	0	0.0
Lebanese	0	0.0	0	0.0	1	4.5
Native American	0	0.0	0	0.0	1	4.5
<b>Total</b>	<b>45</b>	<b>100</b>	<b>33</b>	<b>100</b>	<b>22</b>	<b>100</b>

**Table 7. County & City of Residence**

County	City of Residence	Discoveries		Comparison-services		No Services	
		N	%	N	%	N	%
Linn	Cedar Rapids	33	73.3	19	57.6	6	27.3
	Hiawatha	2	4.4	0	0.0	0	0.0
	Marion	8	17.8	1	3.0	0	0.0
	<b>Subtotal</b>	<b>43</b>	<b>95.5</b>	<b>20</b>	<b>60.6</b>	<b>6</b>	<b>27.3</b>
Johnson	Iowa City	0	0.0	5	15.2	0	0.0
	Coralville	0	0.0	1	3.0	0	0.0
	North Liberty	0	0.0	2	6.1	0	0.0
	<b>Subtotal</b>	<b>0</b>	<b>0.0</b>	<b>8</b>	<b>24.2</b>	<b>0</b>	<b>0.0</b>
Benton	Belle Plaine	1	2.2	1	3.0	0	0.0
Rural Linn	Center Point	1	2.2	0	0.0	0	0.0
Rural Linn	Mount Vernon	0	0.0	4	12.1	13	59.1
Cedar	Mechanicsville	0	0.0	0	0.0	2	9.1
Cedar	Stanwood	0	0.0	0	0.0	1	4.5
	<b>Subtotal</b>	<b>2</b>	<b>4.4</b>	<b>5</b>	<b>15.2</b>	<b>16</b>	<b>72.7</b>
<b>Total</b>		<b>45</b>	<b>100</b>	<b>33</b>	<b>100</b>	<b>22</b>	<b>100</b>

**Table 8. Current Living Arrangements**

Current Living Arrangements	Discoveries		Comparison-services		No Services	
	N	%	N	%	N	%
Parental home	31	0.89	17	0.61	20	0.95
Relatives' home	2	0.06	0	0.00	1	0.05
Shelter care	2	0.06	9	0.32	0	0.00
Own household	0	0.00	1	0.04	0	0.00
Other	0	0.00	1	0.04	0	0.00
<b>Total</b>	<b>35</b>	<b>1.00</b>	<b>28</b>	<b>1.00</b>	<b>21</b>	<b>1.00</b>

Overall, the comparison girls were more likely to be Caucasian and from a more rural area than the girls in treatment at Discoveries. The no-service comparison girls were also a little more likely to be living in their parents' home than the Discoveries clients, but a third of those in the comparison-services group

were referred by the Foundation 2 Youth Shelter. The extent to which these are meaningful differences is hard to say. Benton and Cedar are relatively rural counties, whereas Mount Vernon and Center Point are both about 15 minutes from Cedar Rapids. What will be more clear are some of the personality and clinical differences between the samples, which are detailed at the end of the next section.

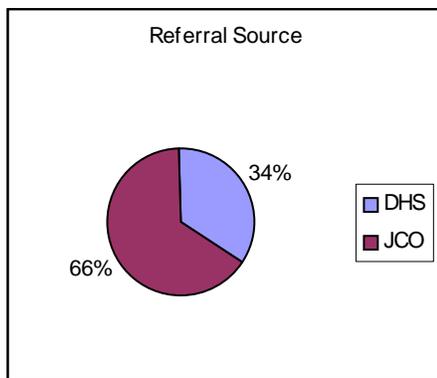
### C. Baseline Descriptive Data

First, following are some descriptive data at baseline for the Discoveries sample. Like Table 5 above, Table 9 below shows that the girls in treatment were distributed pretty evenly in terms of age and grade.

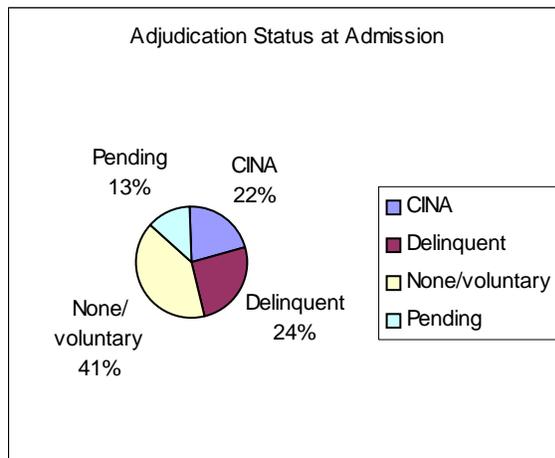
**Table 9. Last Grade Completed**

Grade	N	%
7th	4	9%
8th	13	29%
9th	13	29%
10th	13	29%
11th	2	4%
<b>Total</b>	<b>45</b>	<b>100%</b>

The remaining figures are proportional representations of the Discoveries sample on several selected variables of interest: whether the youth was referred primarily by DHS or Juvenile Court (Fig. 1), adjudication status at intake (Fig. 2), whether or not the youth was living in a single parent home at admission (Fig. 3), whether or not the youth reported attending church or other religious activity on a regular basis during the three months prior to the baseline interview (Fig. 4), and whether or not the girl lived with someone other than her parents or extended family members during the three months prior to the baseline interview (Fig. 5). The number of girls who answered each question or for whom that data was available is provided below each figure.



**Figure 1. Referral Source, N=45**



**Figure 2. Adjudication Status, N=45**

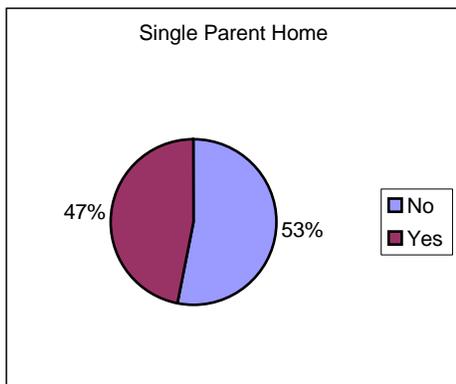


Figure 3. Single Parent Home, N=45

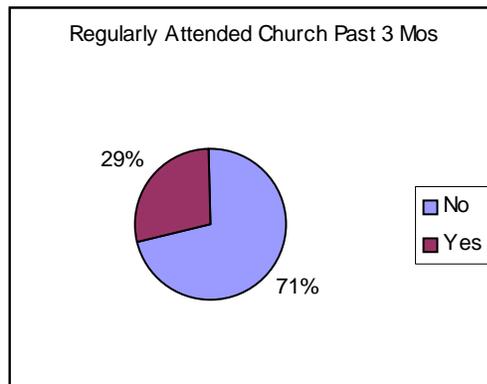


Figure 4. Attended Church, N=34

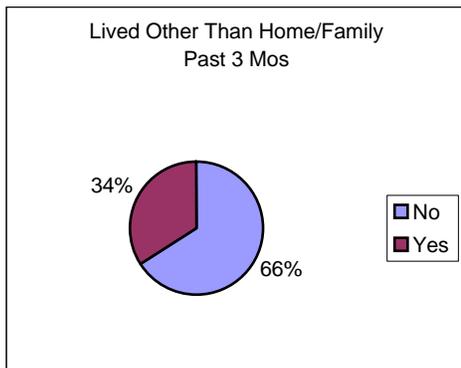


Figure 5. Lived out of Home, N=35

In examining other baseline data for all of the samples, it became clear that there are important differences between them that will have a bearing on the outcomes analyses. Because the differences are most clear and relevant pertaining to the MACI, those results are reported here. However, selected baseline scores for the MACI and for other measures will typically be reported again and further along in the report as part of documenting the results of pre-post analyses within and between the groups. For now, it is important to note that a review of all of the MACI baseline data shows that the comparison sample is starting in a different place clinically than the treatment sample (which is what one would expect).

First, Table 10 lists the results of a comparison of mean scores between the treatment and comparison groups on each of the twelve MACI Personality Pattern scales at T1, baseline. These comparison (t test) results were statistically significant at  $p \leq .05$ <sup>1</sup> on 7 of the 12 scales (shown in bold in the table) and a review of the Means for those scales reveals the direction of the differences: the treatment girls are scoring higher than the comparison girls on the Doleful, Unruly, Forceful, Oppositional, and Borderline Tendency subscales, but much lower on Conforming and Submissive. There are other scales on which the samples may be different overall, but any observed differences were not statistically significant.

<sup>1</sup> Levene's test for homogeneity of variance was used to assess the assumption of equality of variance, and the appropriate degrees of freedom are reflected in the t test results. This can be assumed throughout the report.

**Table 10. Baseline Group Differences on MACI Personality Patterns**

MACI Subscale	Group Statistics				Independent samples t test		
	Sample	N	Mean	Std. Dev.	t	Sig. (2-tailed)	Mean Diff.
T1 Introversive	Treatment	34	44.1	16.1	-0.59	0.555	-2.5
	Comparison	45	46.6	20.5			
T1 Inhibited	Treatment	34	48.1	19.3	-0.94	0.351	-4.2
	Comparison	45	52.3	19.7			
<b>T1 Doleful</b>	Treatment	34	<b>56.6</b>	22.2	2.09	<b>0.040</b>	11.4
	Comparison	45	<b>45.1</b>	25.4			
<b>T1 Submissive</b>	Treatment	34	<b>49.2</b>	18.0	-3.55	<b>0.001</b>	-12.8
	Comparison	45	<b>62.0</b>	12.4			
T1 Dramatizing	Treatment	34	61.2	18.5	-0.72	0.474	-3.5
	Comparison	45	64.6	23.0			
T1 Egotistic	Treatment	34	48.7	21.5	-0.47	0.643	-2.4
	Comparison	45	51.2	23.8			
<b>T1 Unruly</b>	Treatment	34	<b>76.4</b>	17.7	3.95	<b>0.000</b>	17.5
	Comparison	45	<b>58.9</b>	20.8			
<b>T1 Forceful</b>	Treatment	34	<b>61.6</b>	24.3	4.94	<b>0.000</b>	25.6
	Comparison	45	<b>36.0</b>	21.6			
<b>T1 Conforming</b>	Treatment	34	<b>39.1</b>	17.5	-4.34	<b>0.000</b>	-17.7
	Comparison	45	<b>56.7</b>	18.2			
<b>T1 Oppositional</b>	Treatment	34	<b>68.6</b>	20.2	2.51	<b>0.014</b>	11.4
	Comparison	45	<b>57.2</b>	19.9			
T1 Self-Demeaning	Treatment	34	55.4	19.3	0.28	0.784	1.3
	Comparison	45	54.2	21.5			
<b>T1 Borderline Tendency</b>	Treatment	34	<b>53.6</b>	23.2	2.18	<b>0.032</b>	10.8
	Comparison	45	<b>42.8</b>	20.8			

Similar to Table 10, Table 11 lists the t test results for each of the MACI's eight Expressed Concerns and seven Clinical Syndromes scales. In this case, the treatment sample was scoring higher than the comparisons on Social Insensitivity, Family Discord, Childhood Abuse, Substance Abuse Proneness, Delinquent Predisposition, and Impulsive Propensity, and lower on Sexual Discomfort and Anxious Feelings. It is worth noting that they appeared to score higher on Identity Diffusion and Suicidal Tendency at baseline also, although the differences were not significant at  $p \leq .05$ .

**Table 11. Baseline Group Differences on MACI Expressed Concerns and Clinical Syndromes**

MACI Subscale	Group Statistics				Independent samples t test		
	Sample	N	Mean	Std. Dev.	t	Sig. (2-tailed)	Mean Diff.
T1 Identity Diffusion	Treatment	34	53.7	21.1	1.77	0.081	8.7
	Comparison	45	45.0	22.2			
T1 Self-Devaluation	Treatment	34	56.5	21.9	1.09	0.280	6.0
	Comparison	45	50.4	26.1			
T1 Body Disapproval	Treatment	34	44.7	31.5	0.14	0.891	1.0
	Comparison	45	43.8	29.7			
<b>T1 Sexual Discomfort</b>	Treatment	34	<b>42.9</b>	15.5	-2.97	<b>0.004</b>	-10.2
	Comparison	45	<b>53.1</b>	14.9			
T1 Peer Insecurity	Treatment	34	49.1	25.9	-0.14	0.890	-0.7
	Comparison	45	49.9	22.0			
<b>T1 Social Insensitivity</b>	Treatment	34	<b>69.5</b>	17.6	4.09	<b>0.000</b>	15.5
	Comparison	45	<b>54.0</b>	16.0			
<b>T1 Family Discord</b>	Treatment	34	<b>85.2</b>	16.1	3.23	<b>0.002</b>	13.0
	Comparison	45	<b>72.2</b>	18.8			
<b>T1 Childhood Abuse</b>	Treatment	34	<b>51.3</b>	21.1	2.60	<b>0.011</b>	14.1
	Comparison	45	<b>37.2</b>	25.7			
T1 Eating Dysfunctions	Treatment	34	40.3	27.9	0.34	0.733	2.0
	Comparison	45	38.3	24.6			
<b>T1 Substance-Abuse Proneness</b>	Treatment	34	<b>72.4</b>	30.6	4.74	<b>0.000</b>	31.2
	Comparison	45	<b>41.2</b>	27.7			
<b>T1 Delinquent Disposition</b>	Treatment	34	<b>73.4</b>	15.6	5.10	<b>0.000</b>	19.7
	Comparison	45	<b>53.7</b>	17.9			
<b>T1 Impulsive Propensity</b>	Treatment	34	<b>72.4</b>	23.4	4.48	<b>0.000</b>	24.9
	Comparison	45	<b>47.4</b>	25.2			
<b>T1 Anxious Feelings</b>	Treatment	34	<b>47.6</b>	17.5	-5.00	<b>0.000</b>	-18.2
	Comparison	45	<b>65.8</b>	14.9			
T1 Depressive Affect	Treatment	34	69.6	25.8	0.61	0.542	3.5
	Comparison	45	66.1	24.0			
T1 Suicidal Tendency	Treatment	34	47.6	26.4	1.96	0.054	12.1
	Comparison	45	35.6	27.6			

Means testing comparing the two main samples were conducted on all MACI subscales at all four time points, and a few brief observations can be made in regard to those results. First, the pattern of the comparison girls overall being quite a bit more “timid” than the treatment girls can be seen in all of the data, not surprisingly. The primary outcome question to be answered will be to what extent do the treatment girls come to resemble the apparently less troubled girls over time. Second, “apparently” less troubled is noted because in fact, the comparison girls tended to score *higher* than the treatment girls at most time points on the Depression scale, and these differences became statistically significant at two of the time points when the t tests were conducted comparing the treatment girls to the girls from Mount Vernon alone (which was an early type of analysis done, not often repeated later). Third, and even more

strikingly, the comparisons scored much higher on Anxiety at all time points than the treatment girls did; and in this case, the differences were statistically significant at every point when comparing the treatment girls to the comparison-no services girls (which is how the analyses were done in this case). The differences between the treatment and comparison-services girls were significant only at T1 (when the latter scored on average 61.3, with a st.dev. of 16.7). The results comparing the treatment and no service sample are shown in Table 12. Neither group appears to change much over time, and it would appear that the Discoveries group doesn't really need to, but something is going on with the *comparison* girls that needs attention.

**Table 12. Differences Between Groups on Anxiety Subscale Measures**

MACI Subscale	Group Statistics				Independent samples t test		
	Sample	N	Mean	Std. Dev.	t	Sig. (2-tailed)	Mean Diff.
T1 Anxious Feelings	Treatment	34	47.6	17.5	-6.3	0.000	-23.8
	Comparison-No services	20	71.4	10.1			
T2 Anxious Feelings	Treatment	28	50.9	18.3	-4.5	0.000	-19.7
	Comparison-No services	16	70.6	10.5			
T3 Anxious Feelings	Treatment	19	49.2	24.3	-3.4	0.002	-22.4
	Comparison-No services	17	71.6	13.6			
T4 Anxious Feelings	Treatment	13	51.8	17.2	-3.0	0.009	-27.2
	Comparison-No services	4	79.0	9.8			

#### IV. QUANTITATIVE RESULTS

In this section, results from the quantitative statistical analyses are reported, organized according to the domains addressed by the original research questions for the evaluation: school, empowerment, prosocial behavior, relationships, and health. In addition to some descriptive outcomes, the statistical results reported are primarily from paired samples and independent samples t tests. The primary focus of the analysis for this report has been on identifying and testing the significance of any observed changes on key measures for the Discoveries girls between T1 (intake) and T3 (end of aftercare, essentially the end of treatment contact), and occasionally, when the results were interesting or significant for those, T1 and T4 (follow-up). A second focus has been on between-groups means testing. In some cases, independent samples tests were conducted simply comparing the Discoveries and comparison girls on their scores at T3. More often, change scores for each subject were calculated, and then the independent samples tests were conducted comparing the differences in change score means (average amount of change). When these results were significant, it typically meant the Discoveries girls were improving on a given measure whereas the comparison girls were not; it's important to remember in these instances that the treatment girls and possibly the comparison-services girls as well probably had more need for improvement than the other girls did, which in itself can affect the outcome.

To compare changes over time between groups, as noted, change scores were calculated by subtracting each subject's score on a given measure at T3 from her score at T1. The means of these change scores for each group were then compared in those t tests. However, a more vital point is that understanding the

results requires careful attention to the direction of both the measure/scale used as well as the results. T1 minus T3 frequently resulted in a negative mean change score, which may in fact reflect a group improvement on that particular measure. In other cases, the reverse is true. Therefore, the scale and/or meaning of each measure is discussed in turn, along with the direction of observed results.

To sum up the analyses, and make it easier to skim the results if desired, the tables that follow are typically either:

- a) paired samples t tests comparing T1 and T3 means for the treatment girls (or T1 and T4);
- b) independent samples t tests comparing treatment and comparison means at T3; or
- c) independent samples t tests comparing treatment and comparison change score means T1-T3;

It's important to keep in mind throughout the report the relatively small sample sizes for some analyses. Sample sizes can affect the results in many ways, encouraging erroneous conclusions as well as missing possible important ones. Generally, the smaller the sample size, the greater the effect needs to be for statistical significance to be achieved. This is a common issue in these results, where in some cases, the treatment girls' scores appear to have improved more than the comparisons but significance has not been achieved – in which case, the observed difference may or may not be worth noting. The opposite situation occurs frequently also, when because well-validated measures were used (such as the MACI scales), significance is detected even when the sample size is relatively small – in which case, caution is still called for, but more confidence can be placed in the results than not.

One other note on samples sizes: Different analyses of course make use of different subsamples in terms of how they are performed. A certain sample size may be reported for a given measure at T1, which might then be reduced when a *paired* sample t test is performed because the test calls for matching scores for each individual (say, between T1 and T3) and therefore only those girls who had a score at both time points can be included in the analysis. This may be an obvious point to make for some readers, but is intended to simply help avoid confusion when comparing sample sizes from one table to the next.

## **A. School Domain**

Although several data sources included items pertaining to grade in school, hours spent in school, and the like, three items in the evaluation attempted to measure outcomes related to school. Two were from the Youth Questionnaire, and one was from the CAFAS. The Youth Questionnaire asked the girls to rate on a scale of 1 (Almost Always) to 5 (Never) how often they skip school or cut classes. They were also asked to rate how often they feel the school work they are assigned is meaningful and important, using the same scale (which means a lower score is better in this case).

The pre-post results for the Discoveries sample on those two questions are shown in Table 13. Overall, the girls are reporting improved school attendance at the end of treatment (T3) and at follow-up (T4) with a mean at both time points of 4.5 (between Seldom and Never), as compared to intake when they reported only around 3.7. The change from T1 to T3 was statistically significant, although T1 to T4 was not, most likely due to the smaller sample size. On the question of meaningful assignments, the scores appeared to deteriorate on average a little from T1 to T3, but improve from T1 to T4. Neither of the latter results was significant however.

**Table 13. Self-Report School Outcomes for Discoveries Clients**

Discoveries sample only	Group Statistics				Paired Samples Test			
					Paired Differences		t test	
	Measure	N	Mean	Std. Dev.	Mean Diff.	Std. Dev.	t	Sig. (2-tailed)
How Often Skips School*	T1	19	3.7	1.3	-0.8	1.6	-2.2	0.043
	T3	19	4.5	1.0				
How Often Skips School	T1	11	3.6	1.4	-0.9	1.9	-1.61	0.138
	T4	11	4.5	0.7				
School Meaningful & Important*	T1	17	2.8	1.0	-0.2	1.3	-0.78	0.450
	T3	17	3.1	1.2				
School Meaningful & Important	T1	10	2.7	0.9	0.2	0.9	0.69	0.509
	T4	10	2.5	1.1				

\*Scale for both is 1=Almost Always, 5=Never

In Table 14, the results of comparing the change score means for the treatment and comparison groups on how often they report cutting school are shown. As noted above, the treatment girls reported a significant improvement between T1 and T3 in school attendance. Comparing their average change with the very small average change reported by the comparison girls, however, did not reach statistical significance. Remember, a negative change score number on *this* question is an improvement, because change scores

**Table 14. Between Groups Comparison of Change in Self-Report School Truancy**

	Group Statistics				Independent samples t test		
	Sample	N	Mean Change	Std. Dev.	t	Sig. (2-tailed)	Mean Diff.
T1-T3 How Often Skips School	Treatment	19	-0.8	1.6	-1.7	0.106	-0.7
	Comparison	36	-0.1	1.1			

for each girl were calculated as T1 minus T3, and the scale shows that 5 means *Never* cutting school: thus the Discoveries girls T1 mean of 3.7 minus their T3 mean of 4.5 resulted in a mean

change score of -.8. In contrast, the comparison girls only improved by 1/10 of a percent (-.1).

The third outcome measure in the school domain was taken from the CAFAS. The School/Work Role Performance domain subscale was scored by staff for each client as follows: 0=no/minimal impairment, 10=mild impairment, 20=moderate impairment, 30=severe impairment. As shown in Table 15, the difference between the sample’s mean score at intake of 21.2 and their mean at discharge (T3) of 14.8 was highly significant at p<.001.

**Table 15. CAFAS School/Work Domain, Pre-Post Difference in Means**

	Discoveries	
	Intake	Discharge
N	42	42
Mean	<b>21.2</b>	<b>14.8</b>
Std. Dev.	10.2	11.3
	Paired Differences	
Mean	6.4	
Std. Dev.	8.8	
t	4.7	
Sig. (2-tailed)	<b>0.000</b>	

### **B. Empowerment & Self-Efficacy**

This section is devoted to discussion of the analysis and results from those measures seen by the Research Director as gender-related issues involving power and self-image. The entire list of measures in this and the next section is fraught with assumptions and judgment calls that others may or may not agree with, but that needed to be made in order to at least organize the reporting of the results, not to mention spark critical thinking about both the measures used and the interpretation of the results. For a number of reasons that will become clear, it is in this area more than any other that it must be noted that no person who worked on this project has received formal training in the MACI. Although not explicitly required, specific training leading to a broader understanding of the theory or research base for the MACI may (or may not) have led to different interpretations.

First, it's important to note which measures pertaining to power were *not* included in this section, but might have been if one were to take a closer look at what the goals are for the juvenile justice system and what we might like them to be. There are a number of MACI subscales that can be seen as pertaining in some way to *compliance*, including Conforming, Submissive, Unruly, Forceful, and Oppositional. These results were postponed for discussion until the next section entitled Prosocial Behaviors and Attitudes, because it was presumed that they do in fact reflect the traditional objectives of the juvenile justice system – bringing the youth's behavior into compliance with society's rules and norms. It was hypothesized that the program would increase the girls' behavioral compliance, that is, increase their "prosocial behavior and attitudes". However, in the context of developing and applying a gender-specific approach to treating girls, many of whom have histories of abuse and neglect, who live in a society that at some level devalues them and teaches them to suppress themselves in order to meet the expectations of male-dominated institutions and systems, perhaps these MACI measures belong in this section: perhaps we should have hoped for the girls to feel empowered to be less conforming, less submissive, more forceful, more unruly. Perhaps anger and oppositional behavior is called for, given their circumstances! Of course this means they would be even more at odds with their parents, their teachers, and the law, which wouldn't serve them either, or society in the end perhaps. Therefore, given the current state of affairs, some level of conformity may need to be required, as long as the reasons for the lack of conformity in the first place are understood. In any case, one thing we may all be able to agree on is that the systems need to be changed, in order to address and channel what could be seen by some as appropriate responses to the girls' experiences. This section may address part of that question – while the next section's measures may look at how well the program helped the girls to become better behaved, the subscale measures below address

the extent how the girls view themselves and the extent to which they are able to interact with and impact the world around them.

That said, the MACI subscales that most pertain to the concepts of empowerment and self-efficacy are shown in Table 16, along with the results from the initial set of t tests, those testing differences between the Discoveries girls’ mean scores at T1 (intake) and T3 (end of aftercare). The judgment calls made on these measures include which direction of change constitutes an improvement. A decrease in scores from pre- to post- on Introversive and Inhibited is probably an improvement and certainly is on Doleful, Self-Demeaning, Identity Diffusion, and Self-Devaluation. The Dramatizing and Egotistic measures were more challenging. These two personality patterns were described in the MACI manual in less than positive ways, suggesting that like the others, a decrease on those subscales might be indicated. However, in this gender-specific context, a positive outcome of the program would clearly be an increase in the Egotistic subscale, not a decrease. While increasing scores might suggest that the program was encouraging girls to be more “narcissistic” and “unconcerned for the needs of others”, more likely it would mean that it was facilitating their “self-assuredness” and “confidence in their abilities”. The meaning of the Dramatizing scale for this project was less clear, but treatment improvement would still be more likely to be reflected by an increase in the subscale score than a decrease. An increase in this subscale might mean the girls are becoming more “manipulative” or “attention-seeking”, or just as likely that they are “becoming bored with routine”, or “more talkative, charming, exhibitionistic, or emotionally expressive”. The association of some of the terms used to describe this personality pattern, including “manipulative” and “seductive”, with traditional views of women and their ‘feminine wiles’, made this judgment call more of an instinctual one than a scientific one; others are welcome to delve into the question further and draw their own conclusions.

**Table 16. T1-T3 Comparisons on Empowerment Subscales**

MACI Subscales Re: Empowerment	Group Statistics				Paired Samples Test			
					Paired Differences		t test	
	Measure	N	Mean	Std. Dev.	Mean	Std. Dev.	t	Sig. (2-tailed)
Introversive	T1	18	40.9	10.0	0.7	15.9	0.18	0.861
	T3	18	40.2	18.0				
Inhibited	T1	18	42.6	17.8	2.6	14.1	0.79	0.441
	T3	18	39.9	13.7				
Doleful	T1	18	<b>54.7</b>	21.6	13.2	25.4	2.21	<b>0.041</b>
	T3	18	<b>41.5</b>	30.0				
Dramatizing	T1	18	<b>67.1</b>	12.6	-8.5	12.8	-2.81	<b>0.012</b>
	T3	18	<b>75.6</b>	21.2				
Egotistic	T1	18	<b>53.2</b>	22.4	-16.2	15.0	-4.59	<b>0.000</b>
	T3	18	<b>69.4</b>	24.0				
Self-Demeaning	T1	18	54.3	20.7	8.6	22.1	1.65	0.118
	T3	18	45.7	22.9				
Identity Diffusion	T1	18	<b>50.9</b>	23.3	9.4	13.3	3.00	<b>0.008</b>
	T3	18	<b>41.5</b>	22.9				
Self-Devaluation	T1	18	<b>51.8</b>	22.6	10.7	15.3	2.96	<b>0.009</b>
	T3	18	<b>41.1</b>	26.4				

In reviewing the results, it can be seen that the Discoveries girls improved while in treatment on most of the Empowerment subscale scores, and most were statistically significant at  $p \leq .05$  (bolded). Three of them were in fact significant at  $p < .01$ , and one (Egotistic) was highly significant at  $p < .001$ . The girls also appeared to have improved quite a bit on the Self-Demeaning scale, although the result was not significant. There appeared to be

essentially no change on the Introversive scale. Although the results are not shown here, these same tests were conducted for the comparison group. In their case, only one of the eight tests showed a statistically significant change over time. That group's mean Self-Demeaning score went down from 55.4 at T1 to 50.8 at T3.

The same t tests were also done comparing the Discoveries girls' scores at T1 with T4, to see if any of the improvements over time were apparent at follow-up. In fact, the same pattern of results appeared: there was essentially no change on Introversive, a small apparent improvement on Inhibited, and all of the other scores increased at about the same magnitude as at T3 (7-10 points). However, only one difference was statistically significant, shown in Table 17. The Egotistic scale increase was apparently strong enough to overcome even the small sample size of 12.

**Table 17. T1-T4 Significant Differences on Empowerment Subscales**

MACI Subscales Re: Empowerment	Group Statistics				Paired Samples Test			
	Measure	N	Mean	Std. Dev.	Paired Differences		t test	
					Mean	Std. Dev.	t	Sig. (2-tailed)
Egotistic	T1	12	41.8	23.2	-17.3	19.8	-3.02	<b>0.012</b>
	T4	12	59.0	24.7				

Now the question is - how well did the treatment girls do in relation to the comparison girls? Perhaps their apparent improvement was a coincidence or in fact, all girls this age improved on these things during that time frame last year. Although we'll never know for sure, the results of the between groups comparison suggest that it does appear the Discoveries girls improved more in some areas than the comparison girls did. When t tests were conducted on the five subscale measures that were statistically significant above (those that the treatment girls clearly showed improvement on from T1 to T3), three were significant (in bold) and one was very close (in italics).

**Table 18. Between Groups Comparison of Change on Empowerment Subscales**

MACI Subscales Re: Empowerment	Group Statistics				Independent Samples t test		
	Sample	N	Mean Change	Std. Dev.	t	Sig. (2-tailed)	Mean Diff.
Doleful	Treatment	18	<b>13.2</b>	25.4	2.55	<b>0.014</b>	15.5
	Comparison	32	<b>-2.3</b>	17.4			
Dramatizing	Treatment	18	-8.5	12.8	-1.52	0.134	-7.3
	Comparison	32	-1.3	17.7			
Egotistic	Treatment	18	<b>-16.2</b>	15.0	-3.2	<b>0.002</b>	-12.8
	Comparison	32	<b>-3.4</b>	12.8			
Identity Diffusion	Treatment	18	<b>9.4</b>	13.3	2.7	<b>0.010</b>	10.5
	Comparison	32	<b>-1.1</b>	13.2			
Self-Devaluation	Treatment	18	10.7	15.3	2.0	<i>0.057</i>	8.6
	Comparison	32	2.1	14.7			

When reading this chart, note that it depicts the mean change scores for each group (Mean Change column) and that these were calculated by subtracting T1 minus T3; thus a positive number indicates the girls scores went down over time. For example, on Doleful, the treatment girls scores went down by about 13 points, whereas the comparison girls increased by 2. Since we want the girls to be less Doleful, this is a positive outcome. The same direction of change is

desired and can be observed for Identity Diffusion and Self-Devaluation. The treatment girls scores went down by about 9 points on Identity Diffusion; the comparison girls went up by 1. On Self-Devaluation, they both went down, but the treatment girls by 10 and the comparison girls only 2. By contrast, when looking at the two dimensions on this chart for which increases in scores over time were desired, Dramatizing and Egotistic, negative numbers would indicate that the T3 scores were higher than the T1 scores (again, T1 minus T3). And in fact, this is the case for both groups on both scales. However, on both scales, the treatment girls changed more, and in the case of Egotistic, the differences in the mean change scores was again significant at  $p < .01$ . *It is very likely that the program is having a positive effect on the girls' self-image while they are in treatment.*

One other instrument, the Youth Questionnaire, included two Empowerment-related measures. When the girls were asked to rate at each time point the extent to which they had experienced changes in any of the listed areas “compared to three months ago”, the Discoveries girls’ self-report ratings improved somewhat from T1 to T3 on the Empowerment-related items and they scored a little higher than the comparison girls at T3. However, for the most part, t tests revealed no significance to the differences observed. Table 19 below lists the mean ratings for the three groups at T1 and T3 on those items.

**Table 19. Self-Reported Changes in Empowerment-Related Areas**

Item	Measure	Sample	N	Mean	Std Dev
Change in past 3 mos: Self efficacy as a person	T1	Discoveries	35	3.7	1.2
	T3	Discoveries	20	4.1	1.0
	T1	Comparison-services	27	3.5	1.4
	T3	Comparison-services	20	3.9	1.0
	T1*	Comparison-no services	22	3.5	0.9
	T3*	Comparison-no services	16	3.8	0.9
Change in past 3 mos: Self efficacy as a girl	T1	Discoveries	35	3.9	1.0
	T3	Discoveries	20	4.3	0.9
	T1	Comparison-services	28	3.9	1.0
	T3	Comparison-services	20	4.0	0.9
	T1	Comparison-no services	22	3.9	0.8
	T3	Comparison-no services	16	3.8	0.9

Self-efficacy as a person refers to the item “feeling good about myself as a person, that I am worthwhile and capable”, and self-efficacy as a girl refers to “feeling good about myself as a girl, that I have the right to speak for myself and pursue my own goals”. The ratings were 1) much worse, 2) a little worse, 3) about the same, 4) somewhat better, and 5) much better. The results therefore tend to move in the desired direction but are inconclusive, since the girls tended to see somewhat positive

\* This is the only pair that showed significance *within* groups over time at  $p \leq .05$ .

None of the differences *between* samples at T1 or at T3 on these items was significant.

changes occurring before treatment and the comparison girls tended to answer somewhat positively also. It is interesting to note the pattern between the *three* samples, that the average scores at T1 were nearly the same on both items, and then at T3 were highest for the Discoveries, next highest for the comparison-services, and lowest for the no-services girls. This pattern is not uncommon in the results throughout the report, suggesting the Discoveries girls gained the most (and often had the most to gain).

### **C. Prosocial Attitudes & Behavior**

Having already noted above some of the issues related to the goal of increasing prosocial behavior among youth involved in the system(s), this section presents the various measures included in the evaluation to assess the extent to which the program may have achieved this objective. From a more traditional perspective, therefore, the working hypotheses for the MACI measures were that the Discoveries girls

mean scores would increase on the Submissive and Conforming subscales, and decrease on the Unruly, Forceful, Oppositional, Social Insensitivity, Delinquent Predisposition, and Impulsive Propensity subscales. The analyses and results for the MACI Compliance-related subscales are presented in the same format as the Empowerment data above. These are then followed by results from other measures in the evaluation of relevance to prosocial behavior, including items from the CAFAS, criminal activity known by Juvenile Court, self-reported problems from the Youth Questionnaire, and self-reported criminal activity. While substance abuse is certainly a prosocial/compliance issue in some ways, it is included in the Health section later in the report.

The results of the t tests for the Compliance-related MACI subscales were mixed. As shown in Table 20, very small (if any) changes can be observed in treatment sample means from T1 to T3 on Submissive, Unruly, Forceful, and Delinquent Predisposition, although the scores on the latter two subscales are reasonably high (in the 70's), suggesting changes were needed. One other item that was high and did not show significant results was Impulsive Propensity, but the mean in this case did drop from 74.1 to 67.1. The remaining three subscales showed statistically significant results (bolded), but only two were in the hypothesized direction. The girls apparently became more Conforming and less Oppositional while in treatment as presumably intended, but their scores on Social Insensitivity actually *increased*, suggesting they may have become more indifferent to the rights of others and the negative impact of their misbehavior on themselves and society.

**Table 20. T1-T3 Comparisons on Compliance Subscales**

MACI Subscales Re: Compliance	Group Statistics				Paired Samples Test			
					Paired Differences		t test	
	Measure	N	Mean	Std. Dev.	Mean	Std. Dev.	t	Sig. (2-tailed)
Submissive	T1	18	46.8	19.1	-2.1	10.8	-0.81	0.430
	T3	18	48.9	21.7				
Unruly	T1	18	77.6	19.0	0.7	11.3	0.25	0.806
	T3	18	76.9	19.5				
Forceful	T1	18	60.9	27.7	2.6	19.4	0.56	0.584
	T3	18	58.3	27.8				
<b>Conforming</b>	T1	18	<b>39.9</b>	20.3	-8.6	13.0	-2.81	<b>0.012</b>
	T3	18	<b>48.6</b>	25.9				
<b>Oppositional</b>	T1	18	<b>68.4</b>	22.6	7.9	12.7	2.65	<b>0.017</b>
	T3	18	<b>60.5</b>	24.9				
<b>Social Insensitivity</b>	T1	18	<b>71.1</b>	22.7	-6.2	12.1	-2.16	<b>0.045</b>
	T3	18	<b>77.2</b>	18.0				
Delinquent Predisposition	T1	18	74.2	19.3	-3.2	11.9	-1.15	0.268
	T3	18	77.4	15.2				
Impulsive Propensity	T1	18	74.1	27.2	6.9	19.7	1.50	0.153
	T3	18	67.1	29.6				

Like for the Empowerment subscales, t tests for the Compliance subscales comparing means at T1 and T4 were also conducted. This time the Social Insensitivity test was not statistically significant, although the

means increased from 66.3 to 73.0; however, the N was only 12, which may explain why the result was not significant. As shown in Table 21, the Oppositional subscale was again significant, even with only 12 cases, and the Conforming scale was nearly so.

**Table 21. T1-T4 Significant Differences on Compliance Subscales**

MACI Subscales Re: Compliance	Group Statistics				Paired Samples Test			
					Paired Differences		t test	
	Measure	N	Mean	Std. Dev.	Mean	Std. Dev.	t	Sig. (2-tailed)
<i>Conforming</i>	T1	12	33.8	17.3	-14.5	23.9	-2.10	0.060
	T4	12	48.3	25.9				
<b>Oppositional</b>	T1	12	<b>75.3</b>	18.8	13.2	18.8	2.42	<b>0.034</b>
	T4	12	<b>62.1</b>	22.3				

Now, looking at the between groups comparisons, we see additional support for an improvement in some aspects of compliance for the girls who attended Discoveries. As shown in Table 22, the negative mean change score of -8.6 on Conforming for the treatment sample compared to a positive 1.3 for the comparison suggests that the Discoveries girls improved considerably whereas the comparison became slightly less conforming. (T1 minus T3 resulting in a negative score means the score went up over time, and conforming is supposed to go up.) The improvement in Discoveries girls scores on Oppositional compared to the comparison sample was likewise significant (although in this case T1-T3 is positive, indicating the scores went down, as intended.) For Social Insensitivity, a measure we hypothesized would decrease over time, we instead have both groups increasing - the Discoveries girls the most, but the difference is not significant.

**Table 22. Between Groups Comparison of Change on Compliance Subscales**

MACI Subscales Re: Compliance	Group Statistics				Independent Samples t test		
	Sample	N	Mean Change	Std. Dev.	t	Sig. (2-tailed)	Mean Diff.
<b>Conforming</b>	Treatment	18	<b>-8.6</b>	13.0	-2.40	<b>0.021</b>	-9.9
	Comparison	32	<b>1.3</b>	14.5			
<b>Oppositional</b>	Treatment	18	<b>7.9</b>	12.7	2.29	<b>0.026</b>	7.6
	Comparison	32	<b>0.3</b>	10.4			
<i>Social Insensitivity</i>	Treatment	18	-6.2	12.1	-1.59	0.118	-5.5
	Comparison	32	-0.7	11.5			

Turning to the other instruments that included measures of prosocial behavior, the Child and Adolescent Functional Assessment Scale (CAFAS) offers two subscales relevant to this area. First, however, it's important to look at the overall CAFAS scores for the Discoveries girls at intake and discharge. Table 23 shows that there was a dramatic decrease in the total CAFAS scores for the clients, with their average total dropping from 108.1 at intake to 81.4 at discharge (end of aftercare) on a scale of 0-200. This t test result was highly significant statistically.

**Table 23. CAFAS TOTAL Score, Pre-Post Difference in Means**

	Discoveries	
	Intake	Discharge
N	42	42
Mean	<b>108.1</b>	<b>81.4</b>
Std. Dev.	29.7	36.9
	Paired Differences	
Mean	26.7	
Std. Dev.	28.3	
t	6.1	
Sig. (2-tailed)	<b>0.000</b>	

Recall that the sample size of 42 for the t tests on the CAFAS scores includes all of the girls who entered treatment during the study period for whom both intake and discharge CAFAS data were available. Also it's important to note the relatively high standard deviation, meaning there was a wide range in how the girls were scoring (and thus presumably functioning). A review of the raw data reveals that a few of the girls actually did worse from admission to discharge (although most obviously did not) and that there was in fact a wide variation in how functional clients appeared to be when they entered treatment and how much they improved while there. This is to be expected since the sample includes the entire client population during that time

period, not just those who successfully completed treatment, which would appear to make the results even more impressive. On the other hand, there are some reliability issues with the CAFAS, primarily because scores are based entirely on staff ratings which can change with training and experience, vary from one staff to the next, and rely on limited information, especially at intake.

One of the primary domains of the CAFAS is Role Performance, which is broken down into three subscale areas, School/Work, Home, and Community. Table 24 shows the results of the t test comparing means on the Community subscale at intake versus discharge, accompanied by the t test results for the Behavior Toward Others domain scale in Table 25. Both improvements were statistically significant.

**Table 24. CAFAS - Community Domain, Pre-Post Difference in Means**

	Discoveries	
	Intake	Discharge
N	42	42
Mean	<b>16.4</b>	<b>13.6</b>
Std. Dev.	10.1	10.1
	Paired Differences	
Mean	2.9	
Std. Dev.	8.3	
t	2.2	
Sig. (2-tailed)	<b>0.032</b>	

**Table 25. CAFAS - Behavior Toward Others, Pre-Post Difference in Means**

	Discoveries	
	Intake	Discharge
N	42	42
Mean	<b>19.5</b>	<b>15.0</b>
Std. Dev.	7.0	7.4
	Paired Differences	
Mean	4.5	
Std. Dev.	5.5	
t	5.3	
Sig. (2-tailed)	<b>0.000</b>	

The data for the CAFAS domain of Thinking are reported here also (although it could perhaps better be categorized with mental health concerns). The mean for Thinking was only .5 at both intake and discharge, indicating very few youth had any problems at all with disordered thinking or hallucinations.

**Table 26. CAFAS Thinking Scores**

	Mean	N	Std. Dev.
<b>Thinking - Intake</b>	<b>0.5</b>	<b>42</b>	<b>2.2</b>
<b>Thinking - Discharge</b>	<b>0.5</b>	<b>42</b>	<b>2.2</b>

In addition to data provided by youth and program staff, data from the Juvenile Court database system was examined to assess the extent to which the subjects in the study were engaging in criminal activity as known and recorded by the justice system. The data available included a list of the arrests/allegations made against the youth, the charges actually filed, and the dispositions, including placements, going back from the data of the query (February, 2003) to the youth's first contact with the system. About half of the subjects in the comparison-services sample (16 of 33 names queried) had a record in the system while none of those in the comparison-no services sample did (recall that that was one criterion for being placed in the comparison-services sample). Of the 42 total youth in the treatment sample, 33 had been discharged from Discoveries as of the date of the query and had a record in the JCO database.

There were a number of ways to examine and present this data. Table 27 attempts to present the most data that is the most pertinent in the simplest format, focusing on the Discoveries samples only. It concentrates on the number of allegations made against a youth when she was apprehended for some reason, which of course may or may not have eventually turned into actual charges filed. The allegations may have been made on one or more occasions. In the first set of three columns (in green), are the girls who were Successfully Discharged from Discoveries (N=16). Prior to intake, 1 of these girls had had no allegations at all, 6 of them had had one allegation, 7 had had two, and 1 each had had four and five respectively. When the numbers were tallied using the girls' discharge dates instead of intake dates, the number of allegations changes, showing that one of the girls moved from the 0 allegations row to the 1 allegation row – meaning that according to the records, one of the girls gained an allegation while in treatment. No other changes occurred while those girls were in treatment. In the third column of that set, it shows that only 2 of the 16 who were successfully discharged from treatment had received any allegations by the police between the time they were discharged and the date of the records check.

Similarly, the girls whose treatment was deemed Partially Successful by program staff at discharge are shown in the next set of three columns in blue (N=5). Two of these girls had no allegations during the “follow-up” period, whereas 3 did. As for the clients whose treatment was not successful or did not attend treatment at all (N=12), it appears at least 3 of them picked up additional allegations between intake and discharge. Three of the 12 also received allegations after they were discharged to date.

**Table 27. Criminal Allegations Recorded by the Juvenile Court Pre-, During, and Post-Treatment**

N of allegations	Successful Discharges			Partially Successful			Unsuccessful or no treatment*		
	Prior to intake	Prior to discharge	Post discharge	Prior to intake	Prior to discharge	Post discharge	Prior to intake	Prior to discharge	Post discharge
0	1	0	14	0	0	2	1	0	9
1	6	7	2	3	3	1	4	5	3
2	7	7	0	1	1	2	4	3	0
3	0	0	0	0	0	0	1	2	0
4	1	1	0	1	1	0	2	1	0
5	1	1	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	1	0
<b>Total Youth**</b>	<b>16</b>	<b>16</b>	<b>16</b>	<b>5</b>	<b>5</b>	<b>5</b>	<b>12</b>	<b>12</b>	<b>12</b>

\* 2 youth in this sample did not attend treatment because their parents did not consent

\*\* Total Youth includes all girls from the Discoveries sample who have been discharged and have a record in the Juvenile Court database (N=33)

**Table 27a. Months since Discharge**

	N	%
0-3 mos	7	17%
4-6 mos	13	31%
7-9 mos	13	31%
10-12 mos	9	21%
Total	42	

Since the likelihood of being “arrested” depends to some extent on the length of the follow-up period, Table 27a to the left shows the number of months since discharge for the Discoveries girls at the time of the JCO database query.

Other measures included in the general compliance-related area were some of the self-reported problems experienced by subjects in the recent past. On the Youth Questionnaire, subjects were asked to report whether or not they had experienced problems in a number of areas, including a few related to prosocial behavior. Means testing could not be done on these items because the girls simply checked the item if they experienced it during the prior 3 months. However, it appears that some of the Discoveries clients may have made some gains on the two problem areas shown below. In Table 28, the percentage of Discoveries girls reporting problems with fighting with parents, at school, or with others during the past three months dropped from 71.4% at intake to 55% at discharge. Similarly, only 5% reporting having problems with stealing at discharge, compared to 34% at intake. The sample size for the discharge percentages, however, is almost half the intake sample, and it is likely that many of the remaining respondents for these items were those whose treatment was most successful.

**Table 28. Self-Reported Problems in Past Three Months at T1 and T3, Delinquency Issues**

Discoveries sample only		T1		T3	
		N	%	N	%
Had problem in past 3 mos: Fighting w/parents, at school, others	Yes	25	71.4	11	55.0
	No	10	28.6	9	45.0
Had problem in past 3 mos: Stealing	Yes	12	34.3	1	5.0
	No	23	65.7	19	95.0

Yes simply means that item on the list was checked by that subject; no means it was not checked.

A final set of measures related to Prosocial Behavior were included on the Youth Questionnaire, where subjects were asked to report the number of times they were arrested for and actually committed a variety of crimes. Although a great deal of effort was put into making the data collection independent and confidential, and letting the subjects know that, it’s hard to say of course how comfortable they felt about being honest on these questions. In any case, the results for their self-reported criminal activity are reported in Table 29.

The data was aggregated into three categories: crimes against persons, crimes against property, and alcohol and other drug-related (AOD) crimes. The number of crimes each person reported committing in each category were totaled at each time point, and then a change score was calculated for each subject by subtracting the number of crimes they reported at T1 minus the number they reported at T3. The mean change score for the treatment group was then compared in a t test with that of the comparison group. In each of the three crime categories, the treatment sample mean decreased more than the comparison

sample's did. However, the differences were not statistically significant. It should also be noted again that there were some problems with reliability on these questions, particularly the AOD related crimes, in which cigarette smoking for example was counted differently by different subjects.

**Table 29. Change in Self-Reported Criminal Activity**

Self-Reported Crimes Past 3 Months	Group Statistics				Independent Samples t test		
	Sample	N	Mean Change	Std. Dev.	t	Sig. (2-tailed)	Mean Diff.
T1-T3 Crimes Against Persons	Treatment	19	1.6	4.9	1.00	0.331	1.1
	Comparison	37	0.5	1.7			
T1-T3 Crimes Against Property	Treatment	19	4.5	16.2	1.12	0.277	4.2
	Comparison	37	0.4	2.5			
T1-T3 AOD Related Crimes	Treatment	19	2.4	23.8	0.32	0.747	1.8
	Comparison	37	0.6	17.7			

#### D. Relationships

In this section, results are reported for measures from the MACI, the CAFAS, and a few special items from the Youth Questionnaire on family and personal relationships.

The MACI subscales selected as most relevant to Relationships were Family Discord and Peer Insecurity. (The subscale on Child Abuse would be relevant also, but it was not used as an outcome measure in the same way as the other scales since it was not expected to change as a result of treatment participation, and if the scores did change, the interpretation would have been too complex an undertaking in this context anyway.) The t test comparing the Discoveries sample means between T1 and T3 actually show an increase in Peer Insecurity, which is worth noting, although it was not significant statistically. The mean decrease in Family Discord was, however, significant, and in the intended direction, as shown in Table 30.

**Table 30. T1-T3 Comparisons on Relationship Subscales**

MACI Subscales Re: Relationships	Group Statistics				Paired Samples Test			
					Paired Differences		t test	
	Measure	N	Mean	Std. Dev.	Mean	Std. Dev.	t	Sig. (2-tailed)
Peer Insecurity	T1	18	40.4	22.5	-4.3	20.9	-0.87	0.396
	T3	18	44.7	19.8				
Family Discord	T1	18	<b>85.7</b>	18.6	7.8	13.9	2.40	<b>0.028</b>
	T3	18	<b>77.9</b>	17.9				

The improvement on Family Discord was also significant between the T1 and T4 means, as shown in Table 31.

**Table 31. T1-T4 Significant Differences on Relationship Subscales**

MACI Subscales Re: Relationships	Group Statistics				Paired Samples Test			
					Paired Differences		t test	
	Measure	N	Mean	Std. Dev.	Mean	Std. Dev.	t	Sig. (2-tailed)
Family Discord	T1	12	<b>89.0</b>	11.0	13.5	16.0	2.93	<b>0.014</b>
	T4	12	<b>75.5</b>	17.1				

Table 32 shows that the comparison group mean also increased by 1.2 points between T1 and T3, as compared to the 7.8 point improvement made by the treatment sample. Although this may be a real difference, it's hard to say because it was not statistically significant. Chances are the program is having some effect on family conflict.

**Table 32. Between Groups Comparison of Change on Relationship Subscales**

MACI Subscales Re: Relationships	Group Statistics				Independent Samples t test		
	Sample	N	Mean Change	Std. Dev.	t	Sig. (2-tailed)	Mean Diff.
Family Discord	Treatment	18	7.8	13.9	1.56	0.125	6.6
	Comparison	32	1.2	14.7			

Results from the t test comparing intake and discharge means for the (Role Performance at) Home subscale on the CAFAS are shown in Table 33. As on other CAFAS subscales, the Discoveries girls on average appeared to improve in functioning in this area, the difference of 6.2 in pre- and post-test means being statistically significant.

**Table 33. CAFAS Home Domain, Pre-Post Difference in Means**

	Discoveries	
	Intake	Discharge
N	42	42
Mean	<b>23.3</b>	<b>17.1</b>
Std. Dev.	11.1	10.1
	Paired Differences	
Mean	6.2	
Std. Dev.	9.1	
t	4.4	
Sig. (2-tailed)	<b>0.000</b>	

The Youth Questionnaire included five questions measuring the subject’s perceptions of family togetherness and family conflict on a scale of 1 to 5, with 5 indicating greater cohesion and/or lesser conflict. The rating chosen for each of the five questions was averaged to serve as a composite score for each subject. These composites were then used to calculate mean scores for the samples. As shown in Table 34, the Discoveries girls scored on average the same as the comparison-services group and lower than the no-services group. Interestingly, neither of the comparison group means changed at all between T1 and T3, whereas the Discoveries did improve to a small degree. The difference between the two scores was not significant in a paired samples t test, however, (not shown here) nor was the change score significantly different than that of the comparison groups (which was nearly 0 in both cases of course). It’s important to keep in mind that the small improvement shown by the Discoveries may or may not be worth noting - although the five questions were drawn from validated instruments measuring these concepts, this does not assure that these five themselves thereby constitute a valid composite scale statistically. With a larger, better instrument devoted to this question alone and/or a larger sample, the results might have been significant.

**Table 34. Family Cohesion scores**

	Measure	Sample	N	Mean	Std Dev
<b>Family Cohesion composite score</b>	T1	Discoveries	19	3.3	1.1
	T3	Discoveries	19	3.5	0.9
	T1	Comparison-services	20	3.5	0.9
	T3	Comparison-services	20	3.5	0.9
	T1	Comparison-no services	17	4.0	0.5
	T3	Comparison-no services	17	4.0	0.6

All three of the groups showed modest improvement between T1 and T3 on most of the Youth Questionnaire self-report improvement items pertaining to relationships, with a few exceptions, although none were statistically significant. The Discoveries girls did report at T3 significantly more improvement over the past three months than the other groups in getting along with teachers and peers, and expressing anger (which in this case means simply that they scored higher at T3 than the other groups did, not that they necessarily scored different statistically than themselves at T1). Table 35 charts out the descriptive statistics; the t test results are not shown, although the data are highlighted for the two variables where significance was achieved between groups.

**Table 35. Perceived Past 3 Month Improvement in Relationships**

Item	Measure	Sample	N	Mean	Std Dev
Changes in past 3 mos: Getting along w/parents & teachers	T1	Discoveries	35	3.7	1.3
	T3	Discoveries	20	4.3	1.0
	T1	Comparison-services	28	3.7	0.9
	T3	Comparison-services	20	3.7	1.1
	T1	Comparison-no services	22	3.3	0.7
	T3	Comparison-no services	16	3.4	0.9
Changes in past 3 mos: Expressing anger	T1	Discoveries	35	3.7	1.2
	T3	Discoveries	20	4.2	0.8
	T1	Comparison-services	28	3.7	1.1
	T3	Comparison-services	20	3.9	0.9

	T1	Comparison-no services	22	3.4	0.7
	T3	Comparison-no services	<b>16</b>	<b>3.6</b>	<b>0.7</b>
Changes in past 3 mos: Friendships	T1	Discoveries	35	4.0	1.1
	T3	Discoveries	20	4.1	0.9
	T1	Comparison-services	28	4.1	1.1
	T3	Comparison-services	20	3.9	1.1
	T1	Comparison-no services	22	3.7	0.8
	T3	Comparison-no services	16	3.8	0.9
Changes in past 3 mos: Romantic relationships	T1	Discoveries	35	3.7	1.4
	T3	Discoveries	20	3.8	1.0
	T1	Comparison-services	27	3.7	1.1
	T3	Comparison-services	20	3.4	1.1
	T1	Comparison-no services	22	3.7	0.8
	T3	Comparison-no services	22	3.7	0.8

### E. Health Domain

In the Health area, which was broadly defined for the evaluation to include mental health, substance abuse, and spirituality as well as physical health, the MACI had seven subscales. The Discoveries sample was expected to improve on each of the scales between T1 and T3, as represented on these items by decreases in the mean scores (resulting in a positive mean *difference*). As shown in Table 36, this was the case on the Body Disapproval subscale, which was the only difference that was statistically significant (bolded). The Eating Dysfunctions and Suicidal Tendency subscale means also decreased, and the differences were nearly significant (italicized). None of the means for these subscales were

**Table 36. T1-T3 Comparisons on Health Subscales**

MACI Subscales Re: Health	Group Statistics				Paired Samples Test			
					Paired Differences		t test	
	Measure	N	Mean	Std. Dev.	Mean	Std. Dev.	t	Sig. (2-tailed)
Body Disapproval	T1	18	<b>40.1</b>	32.8	11.4	18.1	2.68	<b>0.016</b>
	T3	18	<b>28.7</b>	25.6				
Sexual Discomfort	T1	18	43.4	17.2	-4.7	12.2	-1.62	0.124
	T3	18	48.1	17.4				
<i>Eating Dysfunctions</i>	T1	18	36.6	29.1	9.1	18.8	2.05	<i>0.056</i>
	T3	18	27.4	21.4				
Substance Abuse Proneness	T1	18	69.9	34.9	0.6	15.8	0.15	0.883
	T3	18	69.3	36.5				
Anxious Feelings	T1	18	46.3	20.8	-3.3	15.7	-0.89	0.387
	T3	18	49.6	24.9				
<i>Depressive Affect</i>	T1	18	62.9	25.9	9.7	25.6	1.61	0.125
	T3	18	53.2	30.0				
<i>Suicidal Tendency</i>	T1	18	46.1	26.9	9.8	20.2	2.07	<i>0.054</i>
	T3	18	36.3	27.0				

particularly high to begin with. By contrast, the Depressive Affect mean was high enough to warrant a decrease, which did occur, although the t test was not significant. Likewise, the Substance Abuse Proneness mean was reasonably high at intake; however, it did not decrease by discharge. The

remaining two subscales, Sexual Discomfort and Anxious Feelings, actually got a little worse, in that the means increased slightly over time. The results were not significant, however, so the differences may not mean much.

Recall the finding reported early in the narrative that the Discoveries and comparison girls (the comparison-no services girls in particular) scored very differently on the Anxious Feeling scales; the results shown here (as well as those not shown) demonstrate that neither group changed significantly over time. Overall, the Discoveries girls continued to score much lower on Anxious Feelings than the other groups at all time points.

The decrease in suicidal tendency score is somewhat intriguing, and comparing the mean changes between the treatment and comparison groups on that scale did produce a significant result, as shown in Table 37. The decrease in Depressive Affect for the treatment sample compared to the small increase for the comparison group of .9 remains notable as well, although not statistically significant. Both groups appeared to get better over time on Body Disapproval and Eating Dysfunctions, making it a little more difficult to say if the changes observed on those items above have real meaning.

**Table 37. Between Groups Comparison of Change on Health Subscales**

MACI Subscales Re: Health	Group Statistics				Independent Samples t test		
	Sample	N	Mean Change	Std. Dev.	t	Sig. (2- tailed)	Mean Diff.
Body Disapproval	Treatment	18	11.4	18.1	1.15	0.257	5.5
	Comparison	32	5.9	15.0			
Eating Dysfunctions	Treatment	18	9.1	18.8	0.43	0.668	1.9
	Comparison	32	7.2	12.2			
<i>Depressive Affect</i>	Treatment	18	9.7	25.6	1.65	0.112	10.7
	Comparison	32	-0.9	12.8			
<b>Suicidal Tendency</b>	Treatment	18	<b>9.8</b>	20.2	2.15	<b>0.037</b>	10.6
	Comparison	32	<b>-0.8</b>	14.5			

Breaking down the mean scores for the Suicidal Tendency subscale for all three subsamples further reveals that both the Discoveries group and the comparison-services group average scores decreased over time (with one anomalous change at T3 for the service group), suggesting it is the no-services comparison group responsible for the difference on that measure.

**Table 38. Changes in Suicidal Tendency by Sample**

Measure	Treatment		Comparison-services		Comparison-no services	
	N	Mean	N	Mean	N	Mean
T1	34	47.6	25	47.8	20	20.3
T2	28	40.8	20	41.3	16	27.3
T3	19	36.0	19	45.8	17	27.0
T4	13	35.2	13	34.0	4	32.3

Going on to review whether or not any of the apparent changes that may have occurred for the Discoveries girls during treatment might have been maintained following treatment, Table 39 shows the t test results comparing T4 with T1 means on those measures that had significant or interesting T1-T3 results. The pattern tends to be pretty much the same as at T3, but interestingly, the means for the follow-up sample on these measures at both T1 and T4 tended to be quite a bit higher than those of the larger sample available at discharge. This pattern arose on some other measures, but this is one of the more marked occurrences.

**Table 39. T1-T4 Significant Differences on Health Subscales**

MACI Subscales Re: Health	Group Statistics				Paired Samples Test			
					Paired Differences		t test	
	Measure	N	Mean	Std. Dev.	Mean	Std. Dev.	t	Sig. (2-tailed)
<i>Body Disapproval</i>	T1	12	50.2	32.7	11.1	19.1	2.01	0.070
	T4	12	39.1	28.6				
Eating Dysfunctions	T1	12	45.4	28.2	7.1	20.1	1.22	0.249
	T4	12	38.3	23.8				
Depressive Affect	T1	12	72.8	27.8	10.9	32.3	1.17	0.266
	T4	12	61.8	31.3				
<b>Suicidal Tendency</b>	T1	12	<b>52.7</b>	26.6	16.5	21.9	2.61	<b>0.024</b>
	T4	12	<b>36.2</b>	24.7				

In addition to the MACI subscales above, most of the key measures in the evaluation pertaining to health were drawn from the Testwell. The Testwell Wellness Inventory for Adolescents measures health and wellness on multiple dimensions, including areas not traditionally considered by many of us as wellness-related domains such as “occupational” and “social and environmental” wellness. More importantly, it covers domains of particular interest to the Discoveries and other gender-specific programs such as sexuality and spirituality, substance abuse and emotional expression, as well as nutrition and physical self-care. Results presented below from this instrument are provided in a little more detail than has typically been the case for previous measures. Also, please note that sample sizes in these analyses were affected by the version administered, whether or not the respondent completed all of the questions for a given subscale, had all valid subscales scores from which to calculate a total score, and then had both a pre- and a post-test for the given t test in question.

To begin, Table 40 on the next page provides the descriptive data and t test results comparing the T1 and T3 mean total Testwell scores for both the Discoveries group and the comparison group. And in fact, the Discoveries group did show a statistically significant increase in total Testwell scores between intake and discharge. The comparison group also improved, but the difference was not significant. Not shown here, there was an increase in the mean score for the Discoveries girls between T1 and T4 also; that difference was not significant, but the sample size was very small which has a bearing on the result (N=10).

**Table 40. T1-T3 Comparisons on Total Testwell Scores by Sample**

	Discoveries		Comparisons	
	Pretest	Post	Pretest	Post
N	21	21	18	18
Mean*	<b>164.4</b>	<b>177.9</b>	174.9	181.0
Std. Dev.	25.8	27.3	20.2	21.7
	Paired Differences		Paired Differences	
Mean	-13.5		-6.2	
Std. Dev.	19.5		19.0	
t	-2.9		-1.5	
Sig. (2-tailed)	<b>0.009</b>		0.151	

\*Individual items on the Testwell and thus the mean scores for those items can range from 1 to 5. On the subscales, scores can range from 5 to 25, and total Testwell scores from 50 to 250.

Table 41 provides descriptive data and t test results comparing mean Testwell total scores between the treatment and comparison groups at each of the four test points. It shows that the comparison groups scored higher than the Discoveries girls at all four points, and that the differences were significant at T1 and T2. It also shows that the gain made by the Discoveries girls during treatment *may* have been partially lost by follow-up (or it reflects a change in sample characteristics).

**Table 41. Between Groups Comparison of Testwell Total Scores at all Time Points**

Testwell Total Scores by Time Point	Group Statistics			
	Sample	N	Mean	Std. Dev.
T1	Treatment	26	<b>159.5</b>	26.7
	Comparison	37	<b>179.1</b>	28.0
T2	Treatment	18	<b>157.2</b>	30.1
	Comparison	35	<b>181.5</b>	27.4
T3	Treatment	19	177.0	26.9
	Comparison	25	180.8	22.6
T4	Treatment	12	165.2	60.1
	Comparison	13	172.3	55.6
Independent Samples t test	t	Sig. (2-tailed)	Mean Diff.	
T1	-2.8	<b>0.007</b>	-19.6	
T2	-3.0	<b>0.005</b>	-24.3	
T3	-0.5	0.609	-3.8	
T4	-0.3	0.760	-7.1	

When broken down by the three samples, the same results hold true, in that all subgroups improved on the Testwell over time, with the exception of changes between T3 and T4 (aftercare and follow-up for the Discoveries girls). As shown in Tables 42 and 43, the Discoveries girls' scores in fact typically appeared

to improve more than the other two samples; however, none of the differences were shown to be significant in t tests. Significance might have been obtained with larger sample sizes, since the changes appear to move in the intended pattern and direction (again, with the exception of T3-T4). As in previous sections, when interpreting results, careful attention must be paid to the intended direction of change scores. On all of the Testwell scales, program improvements would be suggested by increasing scores over time and thus negative Mean Change scores.

**Table 42. Discoveries vs. Comparison-services sample on Testwell Change Scores**

Group Statistics	Sample	N	Mean Change	Std. Dev.
T1-T2	Discoveries	14	-4.7	18.3
	Comparison-services	12	-1.0	22.5
T2-T3	Discoveries	11	-2.7	21.2
	Comparison-services	9	-8.6	18.8
T3-T4	Discoveries	9	17.3	52.0
	Comparison-services	6	31.5	65.5
T1-T3	Discoveries	18	-13.5	19.5
	Comparison-services	9	-8.3	17.1
Independent Samples Test	t-test for Equality of Means			
		t	Sig. (2-tailed)	Mean Diff.
T1-T2		-0.5	0.647	-3.7
T2-T3		0.6	0.528	5.8
T3-T4		-0.5	0.648	-14.2
T1-T3		-0.7	0.506	-5.2

Taking note of the T1-T3 results in particular (intake minus discharge from aftercare for the Discoveries girls), we see that the Discoveries group improved on average by 13.5 points on the Testwell, the comparison-services group by 8.3 points (Table 42), and the comparison-no services group by 4.6 (Table 43).

**Table 43. Discoveries vs. Comparison-No services sample on Testwell Change Scores**

Group Statistics	Sample	N	Mean Change	Std. Dev.
T1-T2	Discoveries	14	-4.7	18.3
	Comparison-NO services	19	-2.6	12.4
T2-T3	Discoveries	11	-2.7	21.2
	Comparison-NO services	13	-1.0	15.4
T3-T4	Discoveries	9	17.3	52.0
	Comparison-NO services	2	-6.5	4.9
T1-T3	Discoveries	18	-13.5	19.5
	Comparison-NO services	12	-4.6	20.9
Independent Samples Test	t-test for Equality of Means			
		t	Sig. (2-tailed)	Mean Diff.
T1-T2		-0.39	0.699	-2.1
T2-T3		-0.23	0.820	-1.7
T3-T4		0.62	0.549	23.8
T1-T3		-1.19	0.243	-8.9

Because the program targeted certain key health issues, t test comparisons were conducted on two specific items on the Testwell: a) avoiding fad diets and b) avoiding sex, or if having sex, using condoms. Analyses were undertaken for the means between T1 and T3, and T1 and T4, for both the treatment group and the comparison group. The descriptive data only are shown in Table 44, since the t test results were not very interesting. The comparison group means increased each time, but only one test was significant (between T1 and T3, shown in bold), and the treatment group means were completely mixed. The latter had one pair that showed a notable improvement, an increase in avoiding fad diets between T1 and T4, but the t test was not significant. Perhaps it would have been with a larger sample size.

**Table 44. Changes in Dieting and Safe Sex Measures by Sample**

Sample		Measure	Mean	N	Std. Dev.	
comparison	<b>Pair 1*</b>	T1 Dieting	<b>2.9</b>	29	1.3	
		T3 Dieting	<b>3.6</b>	29	1.1	
	Pair 2	T1 No sex/safe sex	3.9	28	1.7	
		T3 No sex/safe sex	4.1	28	1.4	
	Pair 3	T1 Dieting	3.9	11	1.0	
		T4 Dieting	3.9	11	1.3	
	Pair 4	T1 No sex/safe sex	3.2	12	1.9	
		T4 No sex/safe sex	3.5	12	1.9	
	treatment	Pair 1	T1 Dieting	3.1	19	1.3
			T3 Dieting	2.8	19	1.5
		Pair 2	T1 No sex/safe sex	3.8	19	1.4
			T3 No sex/safe sex	4.1	19	0.9
Pair 3		T1 Dieting	2.8	11	1.5	
		T4 Dieting	3.7	11	1.1	
Pair 4		T1 No sex/safe sex	3.7	11	1.5	
		T4 No sex/safe sex	3.6	11	1.4	

\*(t=-2.7, p=.01)

There are ten formal subscales in the Testwell. Table 45 shows the means for each subscale for the Discoveries sample at T1 and T3, as well as the t tests results for each pair. As shown, the Discoveries girls improved in every domain to some degree, except Nutrition, and in four areas (bolded), the differences were statistically significant. Not shown here, the comparison group also improved in all areas to some extent, but only one was significant – interestingly, in the area of Nutrition (which for them increased from 13.6 to 15.8, N=28).

**Table 45. T1-T3 Comparisons on Testwell Subscales**

Testwell Subscales	Measure	N	Mean	Std. Dev.	Mean Diff.	Std. Dev.	t	Sig. (2-tailed)
Physical Activity	T1	19	13.7	4.6				
	T3	19	14.8	4.9	-1.1	5.0	-0.96	0.350
Nutrition	T1	19	14.5	3.7				
	T3	19	13.8	4.4	0.7	4.4	0.74	0.472
Self Care	T1	19	15.1	4.8				
	T3	19	15.5	4.2	-0.4	4.3	-0.37	0.714
Safety	T1	19	18.3	3.4				
	T3	19	18.8	3.8	-0.6	2.9	-0.86	0.404
Social & Environmental Wellness	T1	19	14.4	4.4				
	T3	19	15.4	4.6	-1.0	3.3	-1.31	0.207
<i>Emotional Awareness &amp; Sexuality</i>	T1	19	19.8	4.5				
	T3	19	21.1	3.0	-1.3	3.1	-1.78	0.091
<b>Emotional Management</b>	T1	19	<b>17.4</b>	4.8				
	T3	19	<b>20.7</b>	4.0	<b>-3.3</b>	<b>3.8</b>	<b>-3.80</b>	<b>0.001</b>
<b>Intellectual Wellness</b>	T1	19	<b>13.6</b>	4.1				
	T3	19	<b>15.8</b>	3.9	<b>-2.2</b>	<b>4.2</b>	<b>-2.30</b>	<b>0.033</b>
<b>Occupational Wellness</b>	T1	18	<b>17.5</b>	4.0				
	T3	18	<b>20.9</b>	3.6	<b>-3.4</b>	<b>4.2</b>	<b>-3.45</b>	<b>0.003</b>
<b>Spirituality &amp; Values</b>	T1	19	<b>19.2</b>	3.7				
	T3	19	<b>21.6</b>	3.8	<b>-2.4</b>	<b>3.8</b>	<b>-2.76</b>	<b>0.013</b>

Despite the fact that the differences *between* the groups on improvements in *total* Testwell scores were not statistically significant (Tables 42 and 43 previously), some of the improvements on *subscale* scores were, as shown in Table 46 on the next page. Three of the four subscales that the Discoveries girls showed significant improvement on over time above (bolded) were significantly different between groups as well. The mean differences over time on the Nutrition scores for the two groups above, where the comparisons improved while the Discoveries girls appeared to lose ground between T1 and T3, was also a statistically significant difference between the groups.

**Table 46. Between Groups Comparison of Change on Testwell Subscales**

Testwell Subscales T1-T3 Changes	Group Statistics				Independent Samples Test		
					t-test for Equality of Means		
	Sample	N	Mean Change	Std. Dev.	t	Sig. (2-tailed)	Mean Diff.
Physical Activity	Treatment	19	-1.1	5.0			
	Comparison	27	-0.9	4.5	-0.13	0.899	-0.2
<i>Nutrition</i>	Treatment	19	<b>0.7</b>	4.4			
	Comparison	28	<b>-2.2</b>	4.7	2.16	<b>0.036</b>	3.0
Self Care	Treatment	19	-0.4	4.3			
	Comparison	28	-1.0	3.4	0.53	0.600	0.6
Safety	Treatment	19	-0.6	2.9			
	Comparison	30	0.7	3.7	-1.30	0.201	-1.3
Social & Environmental Wellness	Treatment	19	-1.0	3.3			
	Comparison	27	-0.7	3.6	-0.32	0.749	-0.3
Emotional Awareness & Sexuality	Treatment	19	-1.3	3.1			
	Comparison	28	-0.6	3.3	-0.68	0.500	-0.7
<b>Emotional Management</b>	Treatment	19	<b>-3.3</b>	3.8			
	Comparison	28	<b>0.0</b>	2.4	-3.37	<b>0.002</b>	-3.3
Intellectual Wellness	Treatment	19	-2.2	4.2			
	Comparison	28	-1.0	3.2	-1.12	0.268	-1.2
<b>Occupational Wellness</b>	Treatment	18	<b>-3.4</b>	4.2			
	Comparison	26	<b>0.0</b>	3.2	-3.11	<b>0.003</b>	-3.5
<b>Spirituality &amp; Values</b>	Treatment	19	<b>-2.4</b>	3.8			
	Comparison	28	<b>0.4</b>	2.8	-2.72	<b>0.011</b>	-2.8

(Remember that negative change scores on the Twell signify improvements over time.)

Given the range of reasonably positive results from the MACI and the Testwell, one would expect that girls who made improvements in one area would be likely to have made improvements in another. To investigate this possibility, Pearson correlations were calculated to assess the extent of any association between and among the T1-T3 change scores on the MACI subscales and the four total Testwell change scores available from analyses on the Testwell to this point.

Not surprisingly, changes in several of the MACI subscales were very highly correlated with one another (although not charted out, some examples are: Forceful and Submissive,  $r = -.63$ , and Egotistic and Self-Devaluation,  $r = -.71$ , Eating Dysfunctions and Body Disapproval,  $r = .91$ , and interestingly, Social Insensitivity and Self-Devaluation,  $r = -.67$ ). T1-T3 changes in a few of the MACI dimensions correlated with a few of the changes over time in total Testwell scores also, as bolded in Table 47. Although some of these results are probably anomalous, those in the far right column are most likely to be of real significance, where the change in total Testwell scores from T1 to T3 are correlated with four key MACI dimensions – Egotistic, Conforming, Identity Diffusion, and Self-Devaluation.

**Table 47. Correlations in MACI and Testwell Changes over Time**

MACI Subscales	Statistic	T1-T2 Wellness change	T2-T3 Wellness change	T3-T4 Wellness change	T1-T3 Wellness change
Submissive Change T1-T3	Pearson Correlation Sig. (2-tailed) N	-0.26 0.151 33	0.44 <b>0.017</b> 29	-0.15 0.608 14	0.28 0.096 36
Unruly Change T1-T3	Pearson Correlation Sig. (2-tailed) N	-0.17 0.354 33	-0.18 0.364 29	-0.43 0.121 14	-0.19 0.259 36
Forceful Change T1-T3	Pearson Correlation Sig. (2-tailed) N	-0.08 0.673 33	-0.24 0.205 29	-0.56 <b>0.037</b> 14	-0.25 0.139 36
Egotistic Change T1-T3	Pearson Correlation Sig. (2-tailed) N	0.03 0.867 33	0.30 0.118 29	-0.53 0.053 14	0.41 <b>0.014</b> 36
Conforming Change T1-T3	Pearson Correlation Sig. (2-tailed) N	-0.02 0.898 33	0.32 0.095 29	-0.46 0.101 14	0.34 <b>0.044</b> 36
Oppositional T1-T3	Pearson Correlation Sig. (2-tailed) N	0.04 0.814 33	-0.15 0.437 29	-0.11 0.718 14	-0.18 0.299 36
Identity Diffusion T1-T3	Pearson Correlation Sig. (2-tailed) N	-0.29 0.106 33	-0.22 0.256 29	-0.12 0.675 14	-0.51 <b>0.001</b> 36
Self-Devaluation T1-T3	Pearson Correlation Sig. (2-tailed) N	-0.15 0.407 33	-0.27 0.163 29	0.32 0.271 14	-0.36 <b>0.029</b> 36
Body Disapproval T1-T3	Pearson Correlation Sig. (2-tailed) N	0.11 0.524 33	-0.26 0.165 29	0.60 <b>0.023</b> 14	-0.09 0.598 36
Social Insensitivity T1-T3	Pearson Correlation Sig. (2-tailed) N	0.04 0.810 33	-0.16 0.411 29	-0.13 0.660 14	-0.01 0.957 36
Family Discord T1-T3	Pearson Correlation Sig. (2-tailed) N	-0.35 <b>0.046</b> 33	-0.04 0.844 29	0.16 0.587 14	-0.26 0.119 36
Childhood Abuse T1-T3	Pearson Correlation Sig. (2-tailed) N	-0.20 0.273 33	-0.24 0.203 29	0.11 0.719 14	-0.27 0.115 36
Substance Abuse T1-T3	Pearson Correlation Sig. (2-tailed) N	-0.11 0.552 33	-0.25 0.193 29	0.00 0.987 14	-0.21 0.228 36
Eating Dysfunctions T1-T3	Pearson Correlation Sig. (2-tailed) N	0.061 0.737 33	-0.207 0.281 29	0.464 0.094 14	-0.01 0.965 36
Delinquent Disposition T1-T3	Pearson Correlation Sig. (2-tailed) N	0.05 0.781 33	-0.46 <b>0.013</b> 29	-0.35 0.224 14	-0.19 0.265 36

When the correlations were calculated again for the four key MACI subscales above with the Testwell total change scores, this time for each of the three samples separately, one of the associations continued to be statistically significant. Changes from T1 to T3 in the Identify Diffusion subscale continued to be negatively correlated with T1-T3 changes in total Testwell scores for both the Discoveries group and the

comparison-no services group, as shown in Table 48. The other associations appeared again as well at some level, but the correlations were not significant at  $p \leq .05$ .

**Table 48. Correlations of Changes in Identity Diffusion and Testwell Scores by Sample**

	Statistic	T1-T3 Change in Wellness & Identity Diffusion
Discoveries T1-T3	Pearson Correlation	-0.55
	Sig. (2-tailed)	<b>0.028</b>
	N	16
Comparison-services T1-T3	Pearson Correlation	-0.03
	Sig. (2-tailed)	0.947
	N	9
Comparison - no-services T1-T3	Pearson Correlation	-0.63
	Sig. (2-tailed)	<b>0.037</b>
	N	11

Moving on to results pertaining to Health from the Youth Questionnaire, such as the Discoveries girls' self-reports at T1 versus T3 regarding which problems they had experienced in the past 3 months, modest reductions over time can be seen in the percentages who reported having problems with unsafe sex and with sexual abuse. As shown in Table 49, no appreciable change occurred in substance abuse, with 40% still reporting at discharge they had experienced problems during the past three months. Similarly, no change appeared to occur in regard to self-harm, although only about 15% reported having problems in that area (representing 3 girls at T3).

**Table 49. Self-Reported Problems in Past Three Months at T1 and T3, Health Issues**

Discoveries sample only		T1		T3	
		N	%	N	%
Had problem in the past 3 mos: Substance Abuse	Yes	15	42.9	8	40.0
	No	20	57.1	12	60.0
Had problem in the past 3 mos: Eating Problems	Yes	6	17.1	3	15.0
	No	29	82.9	17	85.0
Had problem in the past 3 mos: Self-Harm	Yes	6	17.1	3	15.0
	No	29	82.9	17	85.0
Had problem in the past 3 mos: Unsafe Sex	Yes	18	51.4	9	45.0
	No	17	48.6	11	55.0
Had problem in the past 3 mos: Sexual Abuse	Yes	10	28.6	2	10.0
	No	25	71.4	18	90.0

As to be expected, the Discoveries sample scored significantly differently on the substance abuse measures at baseline than did the comparison-no services subjects, as shown in bold in Table 50. On a scale of 1 to 5, the comparison-no services girls had the highest mean score on the attitude composite measure (How Wrong it is to Use .... various substances). Likewise, the number of days they reportedly used tobacco in the past 30 was only 2.7 on average (while the Discoveries girls used tobacco an average of 12.2 days in the past month), and alcohol or other drugs less than one day (.8). Their score for the

number of friends who use substances was the highest, meaning they had closer to few or none. The comparison-services group tended to score roughly the same as the Discoveries girls on these measures.

**Table 50. Substance Abuse Measures at Baseline**

T1 Substance Abuse Measures	Sample	N	Mean	Std. Dev.	t	Sig. (2-tailed)
How Wrong to Use score	Treatment	35	1.8	0.7	0.57	0.571
	Comparison-services	28	1.7	0.9		
	Comparison-no services	22	2.5	0.5		
N of Days in Past 30 Used Tobacco	Treatment	34	12.2	13.7	-0.17	0.865
	Comparison-services	26	12.8	13.4		
	Comparison-no services	22	2.7	8.8		
N of Days in Past 30 Used AOD	Treatment	33	4.1	8.4	-0.19	0.848
	Comparison-services	25	4.5	9.8		
	Comparison-no services	22	0.8	2.3		
Friends Who Use Substances score	Treatment	34	1.8	0.5	-0.44	0.665
	Comparison-services	28	1.8	0.6		
	Comparison-no services	22	2.3	0.4		

Although not shown here, the results above continued to hold true at T3, in that the Discoveries girls still scored significantly differently on these four AOD composite measures than the no-services group. Testing for differences in the means at T1 and T3 for the Discoveries girls reveals why. While the treatment group did improve on the number of days they used alcohol or other drugs while in Discoveries, which is a very good result, the difference was not statistically significant, as shown in Table 51. Moreover, the number of days they smoked cigarettes actually increased on average while they were in treatment from 9.2 to 12.4 (at least for the smaller sample of those who responded to these questions at both T1 and T3). There was basically no change in their attitudes toward use, or the number of friends they had who used.

**Table 51. Discoveries sample, Changes in Substance Use Measures, T1 - T3**

	Measure	N	Mean	Std. Dev.	Mean Diff.	Std. Dev. of diff.	t	Sig. (2-tailed)
How Wrong to Use score	T1	20	1.9	0.7	-0.01	0.7	-0.08	0.935
	T3	20	1.9	0.8				
N Days in Past 30 Used Tobacco	T1	19	9.2	12.8	-3.3	15.3	-0.93	0.364
	T3	19	12.4	14.0				
N Days in Past 30 Used AOD	T1	18	5.2	9.7	3.8	10.4	1.57	0.135
	T3	18	1.4	2.9				
Friends Who Use Substances score	T1	19	1.8	0.6	0.1	0.4	1.12	0.277
	T3	19	1.7	0.5				

The results from the CAFAS Substance Use subscale support the notion that what substance use problems were occurring in the Discoveries group were reduced by the time clients were discharged from the program (at least, other than tobacco). Their mean score dropped from 9.3 at intake to 4.3 at discharge, a

statistically significant difference, although both were assessed on average as being below the level of even minimal impairment even at intake.

**Table 52. CAFAS Substance Use Domain, Pre-Post Difference in Means**

	Discoveries	
	Intake	Discharge
N	42	42
Mean	<b>9.3</b>	<b>4.3</b>
Std. Dev.	10.0	7.0
	Paired Differences	
Mean	5.0	
Std. Dev.	8.3	
t	3.9	
Sig. (2-tailed)	<b>0.000</b>	

Two other CAFAS domain scores are of relevance to the Health area. First, as shown in Table 53, the results for the Self-Harmful Behavior subscale show that as a group, the Discoveries girls had few problems at intake in the area of self-harmful behavior as defined by the CAFAS, and little change occurred between intake and discharge.

**Table 53. CAFAS Self-Harmful Behavior Scores**

	Mean	N	Std. Dev.
<b>Self-Harm - Intake</b>	<b>5.2</b>	<b>42</b>	<b>8.9</b>
<b>Self-Harm - Discharge</b>	<b>5.7</b>	<b>42</b>	<b>9.4</b>

Second, the group’s mean score on Moods/Emotions subscale (reflecting impairment in that area) dropped on average by a small amount from intake to discharge, and the difference in means was almost statistically significant.

**Table 54. CAFAS Moods/Emotions Domain, Pre-Post Difference in Means**

	Discoveries	
	Intake	Discharge
N	42	42
Mean	12.6	11.0
Std. Dev.	7.3	7.6
	Paired Differences	
Mean	1.7	
Std. Dev.	5.4	
t	2.0	
Sig. (2-tailed)	0.051	

As is not uncommon in studies of this type, even when a careful attempt has been made to limit the data collection to specific research questions, there is much more data available for analyses than there is time and space to examine and/or report. Much more can be done with the data for those who might have an

interest, and it should be clear at this point that the data is in fact worth further examination, given the number of positive results that came out of the initial work. Therefore, with appropriate precautions taken to preserve client confidentiality, inquiries on data sharing and collaboration on future research activities will be welcome and encouraged.

However, before moving on to a discussion of qualitative results and then conclusions and recommendations, there are a few remaining results that were drawn from analyses of the primary data collection tools. These have been gathered together into the next and final quantitative results section, Program & Services Data.

## **F. Program & Services Data**

One of the important types of data collection when undertaking program evaluation is to document the nature and types of services provided, so that if and when certain outcomes are observed, one may have some basis upon which to associate positive change with the actual treatment received. In this case, the goal was to measure the volume of gender-specific services provided, which led to the development of a special data collection tool that documented the number of minutes/hours clients spent in individual and group sessions while at Discoveries. The instrument also allowed staff to select from a list of program topics that were addressed at each session (and/or add to the list if needed). Unfortunately, analyzing by topic by time by session by client by outcome proved too complex for this report. However, the total amount of time each client spent in each type of session (group versus individual) was aggregated, and correlations were calculated for each of those with changes in some of the key outcome measures. The results are shown below.

First, in Table 55, Pearson correlations are shown for the total Testwell change scores (T1 minus T3, intake minus discharge from aftercare) with each of the variables of group services hours, individual counseling hours, and then total hours combined. None were correlated. This may mean that: a) the number of hours the clients spend in group or individual sessions does not affect their health outcomes (or theoretically vice versa); b) the cause of health outcomes is program-related but occurs in the relatively small amount of time spent outside of sessions or is caused by some other program factor unrelated to time (such as the nature of the relationships established, or exposure to certain materials); and/or c) the Sessions instrument was an unreliable measure of service hours.

**Table 55. Correlations between Change in Testwell scores and Service Hours**

		Group Service Hours	Individual Counseling Hours	Total Counseling Hours
T1-T3 Testwell change scores	Pearson Correlation	0.13	0.09	0.12
	Sig. (2-tailed)	0.611	0.724	0.622
	N	18	18	18

Correlations were also calculated for each of the service hour variables with each of the MACI subscale change scores (T1-T3 only). As shown in Table 56, there was a reasonably high correlation between each of the service hours variables with the Submissive subscale change scores, in the 40% range, although none quite reached significance. The only pair that was significant was Substance Abuse with Group Service hours, which was a negative correlation of 47%, suggesting that as substance abuse goes up, the number of service hours goes down, and vice versa. There were some other correlations that are worth

noting, given the pattern, although the correlations are not all that high. Unruly, Oppositional, and Delinquent Predisposition each had correlations with Group Service hours in the 20% range, and Body Disapproval was over 30%, all negative, suggesting the program *might* be a factor in any changes observed in those areas. Individual counseling hours as a variable did not correlate with any of the MACI subscales except Submissive. While both Group and Individual hours were highly correlated with Total Hours, of course, they did not contribute equally to the total, and more importantly, they were only correlated at  $r=.57$  with one another. This suggests that the number of individual hours received by clients varied quite a bit, and that this was only partly associated with how long they were in the group/program.

**Table 56. Correlations between Change in MACI subscale scores and Service Hours**

MACI Subscale	Statistic	Group Service Hours	Individual Counseling Hours	Total Counseling Hours
Introversive Change T1-T3	Pearson Correlation	-0.08	0.10	-0.05
	Sig. (2-tailed)	0.749	0.706	0.848
	N	18	18	18
Submissive Change T1-T3	Pearson Correlation	0.40	0.45	0.43
	Sig. (2-tailed)	0.102	0.064	0.079
	N	18	18	18
Unruly Change T1-T3	Pearson Correlation	-0.26	0.03	-0.21
	Sig. (2-tailed)	0.299	0.913	0.398
	N	18	18	18
Forceful Change T1-T3	Pearson Correlation	0.07	0.13	0.09
	Sig. (2-tailed)	0.769	0.610	0.726
	N	18	18	18
Egotistic Change T1-T3	Pearson Correlation	-0.02	-0.04	-0.02
	Sig. (2-tailed)	0.937	0.876	0.922
	N	18	18	18
Conforming Change T1-T3	Pearson Correlation	0.08	-0.06	0.05
	Sig. (2-tailed)	0.757	0.805	0.835
	N	18	18	18
Oppositional T1-T3	Pearson Correlation	-0.25	-0.17	-0.24
	Sig. (2-tailed)	0.327	0.506	0.338
	N	18	18	18
Identity Diffusion T1-T3	Pearson Correlation	-0.05	-0.04	-0.05
	Sig. (2-tailed)	0.834	0.885	0.837
	N	18	18	18
Self-Devaluation T1-T3	Pearson Correlation	0.19	0.17	0.19
	Sig. (2-tailed)	0.453	0.505	0.444
	N	18	18	18
Body Disapproval T1-T3	Pearson Correlation	-0.33	-0.11	-0.30
	Sig. (2-tailed)	0.185	0.665	0.232
	N	18	18	18
Social Insensitivity T1-T3	Pearson Correlation	-0.12	-0.10	-0.12
	Sig. (2-tailed)	0.645	0.690	0.640
	N	18	18	18
Family Discord T1-T3	Pearson Correlation	-0.01	0.07	0.01
	Sig. (2-tailed)	0.978	0.780	0.973
	N	18	18	18
Childhood Abuse T1-T3	Pearson Correlation	-0.31	0.01	-0.26
	Sig. (2-tailed)	0.204	0.977	0.294

	N	18	18	18
<b>Substance Abuse T1-T3</b>	Pearson Correlation	-0.47	-0.05	-0.40
	Sig. (2-tailed)	<b>0.051</b>	0.831	0.098
	N	18	18	18
Eating Dysfunctions T1-T3	Pearson Correlation	-0.14	0.04	-0.11
	Sig. (2-tailed)	0.569	0.874	0.657
	N	18	18	18
Delinquent Disposition T1-T3	Pearson Correlation	-0.26	-0.10	-0.24
	Sig. (2-tailed)	0.296	0.689	0.339
	N	18	18	18
Group Service Hours	Pearson Correlation	1.00	0.57	0.98
	Sig. (2-tailed)	.	0.000	0.000
	N	45	45	45
Individual Counseling Hours	Pearson Correlation	0.57	1.00	0.72
	Sig. (2-tailed)	0.000	.	0.000
	N	45	45	45
Total Counseling Hours	Pearson Correlation	0.98	0.72	1.00
	Sig. (2-tailed)	0.000	0.000	.
	N	45	45	45

One of the items on the Coalition Discharge forms was that staff were asked to assess how much progress each client had made toward their program goals as of the time of discharge. When the client was being discharged when less than 50% progress had been made towards goals, the reasons for discharge were noted. The data available for the Discoveries group during the study period are listed below in Tables 57 and 58, followed by a breakdown in Table 59 of where the girls were placed following discharge (which included remaining at home).

**Table 57. Progress Made During Treatment**

Discharge Assessment	N	%
Substantial progress towards 80 percent of goals	15	31.9
Substantial progress towards 50 percent of goals	10	21.3
Less than 50 percent progress towards goals	17	36.2
<b>Total</b>	<b>42</b>	<b>100.0</b>

**Table 58. Reasons for Less than 50% Progress**

Reason for Discharge	N	%
Higher level of care indicated	9	52.9
Lack of official approval	1	5.9
Runaway	4	23.5
Other	2	11.8
Other-client refusal	1	5.9
<b>Total</b>	<b>17</b>	<b>100.0</b>

**Table 59. Placement At Discharge**

Placement situation	N	%
Homeless	1	2.4
Runaway	4	9.5
Parent home	25	59.5
Relatives home	2	4.8
Foster care	3	7.1
Shelter care	1	2.4
Residential trmt community	1	2.4
Juvenile detention	4	9.5
State mental hospital	1	2.4
<b>Total</b>	<b>42</b>	<b>100</b>

## V. QUALITATIVE RESULTS

Although the bulk of the evaluation was focused on the collection of quantitative data to document outcomes, the study design did call for the research team to gather information from staff and clients regarding their perspectives and experiences with the Discoveries program. Although program materials are very good sources of qualitative information (i.e., narrative descriptions, client newsletters), they can be somewhat limited in nature and availability. In the absence of an experimental design, current qualitative data from clients can also help link observed changes in clients (outcomes) to the program or treatment itself (process). However, caution must be exercised in using data from qualitative interviews. It becomes tempting to draw firm conclusions from statements made by individuals despite the anecdotal/testimonial nature of the information when it's presented in a summary format and the information appears to have substance and color. However, one individual's experience does not necessarily reflect the group's, and certainly the interview samples in this case were small and nonrepresentative. Yet if the results are seen for what they are, a small window into the real experiences of the interviewees, an opportunity to gain some insight into the actual treatment process, then they can be very useful.

In this section, the results from the two sets of qualitative interviews are provided. Standard precautions were taken to protect the confidentiality of clients participating in the process, and careful editing and review was undertaken in an attempt to ensure that sensitive client comments, suggestions, and quotes could not be attributed to a given individual, particularly by anyone other than other clients. Care was also taken to obscure the sources of sensitive comments and quotes from the staff interviews, except where express permission was obtained in advance. Results are organized in general according to questions or question areas discussed in the interviews.

### A. Summary of Youth Interviews

#### Goals that the Girls had for themselves in Attending Discoveries

A common response was: "at first, to get out of there as soon as possible", with other issues becoming more important for some as treatment progressed.

Other responses made by more than one girl:

- Relationships/communication/getting along with parents
- "Male" relationships, "male dependency"
- Drug use

Single responses:

- Anger management
- Getting along with females
- My attitude and my communication

#### Goals that Others Had for the Girls (such as Counselor, Probation Officer, DHS Worker, Parents, etc.)

Common responses:

- Family relationships

Attend school & keep up grades  
“Male” relationships  
Drug abuse

Single responses:

Independent living skills  
Pay off criminal fines  
Stay out of trouble

### **Were the Goals Met? Why/Why Not?**

Most of the girls gave mixed responses, in that some of the goals were met in some ways, and some were not. Examples:

*“Some have and some haven’t. I get along with my mom a lot better since I’ve moved out on my own, but I relapse when it comes to some stuff. I had another bad bad relationship with a man. Some of my goals were met with my parents, we get along a lot better. When I was in Discoveries, yea, because they kept a really close eye on you. You basically live there so you really couldn’t do anything in your spare time. So that did help with that, but as far as when I got out, I don’t think they ‘installed’ enough to keep you from doing it.”*

*“My family relationship, no, and my male relationship, yes, and my attitude, in ways. Because me and my family, we are still fighting, my mom mostly. My male relationship, I’m just not having them. And my attitude, it’s getting better, but I’m ... still cussing at people.”*

A couple explained how “staff have helped as much as they can”, but pointed out the responsibility they felt for their own success as well. For example,

*“We are in here, and yea, they are going to be here to help you, but it’s more you need to do what you need to do on your own, because they aren’t going to be there holding your hand when you are out of the program.”*

None of the girls responded that no goals had been achieved, although one girl avoided answering the question.

### **Most Important or Valuable Issues Addressed**

Common responses:

Healthy “male” relationships  
School  
Drugs and alcohol

Single responses:

“Holding your peers accountable”  
“the violence”  
“the family”  
pregnancy

### **Issues that were Missed that Should Have been Addressed**

A couple of respondents replied there weren’t really any issues they could think of that were missed. All others essentially brought up the same item – communication/getting along with/dealing with feelings about parents.

*“...more of parent/child groups, you know, about communication about getting along better. About having good conversations. I know I’m not the only girl ... who is having problems with my parents. .... But I think if we had at least one or two groups on that, me and other girls would have a better understanding of why parents act that way or why teenagers act this way. Things like that. How not to get upset when parents say this or that.”*

*“I think they should have a lot more discussions about how important family is because all of us that were in there were having problems with our parents or our guardians.”*

*“I think we need to work on communication. ... They told us we had to get along with our parents, but really didn’t help us learn how..... because they didn’t have a lot of groups around feelings, it was more issues.”*

*“Probably the way I feel about my dad and stuff.”*

*“Probably like family stuff. They don’t have a lot of groups on family. How to build healthy relationships.”*

### **Extent to Which You Learned to Express Feelings Appropriately**

Subject responses in this area were very mixed, sometimes off topic, and in general, difficult to summarize. There was some suggestion by a couple of respondents that their feelings were not always of primary importance when a conflict with staff arose (i.e. “ they really didn’t talk about how people felt, they just told you this, told you that”). There was also discussion about how and when staff should intervene when the girls are showing “attitude” toward one another. The most intriguing comment was:

*“I think a lot of staff there have issues. They want you to do something like it’s a personal thing, not just their job. My case manager always has these moral ethical conversations with me. I don’t care, I’m in here because I have to be, I don’t want to talk to you any more.”*

### **Extent to Which You Learned to Develop and Maintain a Positive Support System**

Despite concerns noted above about staff not taking their feelings into account, the same girls answered this question by emphasizing how they can continue to use their case managers for support after leaving the program. Other single responses included a complaint about staff siding with parents, in one case, and in another, about how the treatment period was an opportunity for her to learn who her “real” friends were (those who were willing to make time for her during her restricted free hours and respect her early curfew).

### **Program Addressed Body Image**

The responses were very mixed on this question. Some examples are:

*“Yes. They told people to look at themselves and respect themselves.”*

*“I’m not sure about that one. They just make you fat (laughing). They feed you all kinds of stuff, pasta, etc.”*

*“There have been times when girls have said, ‘Oh, I’m fat and this guy doesn’t like me because I’m fat’ and the staff will say something like, ‘You shouldn’t put yourself down like that,’ and after that, about 2 minutes, that’s it, it’s done. That’s a real self-confidence issue there. I think a lot of girls in this world have this problem. I know I do.”*

*“They did help me a lot when it came to that because I had a really bad eating disorder.... They helped me to realize that I am the way I am. .... You are around \_\_\_ other girls all the time. You become sisters with those girls and you develop a bond with most of them.”*

*“I didn’t really feel they addressed it at all.”*

*“I think aftercare was a lot better because aftercare was focused right on it.”*

### Program Addressed Hygiene, Exercise, Sleep Habits, and Nutrition

Hygiene – most girls mentioned toothbrushes and deodorant, one mentioned ear cleaning

Sleep – was discussed by only one girl

Exercise – a few girls mentioned it, and noted that physical exercise was infrequent

Nutrition – the “etiquette meal” was noted, and that meals were planned out in advance

The overall impression received was that these issues are addressed, but to a lesser degree than other more primary topics.

### Program Addressed Mental Health

Some of the respondents thought mental health was not addressed at all, including by one that mentioned having depression. However, a few indicated that they believed mental health was addressed, but only on an individual level, when someone had a specific mental health problem or diagnosis. There was one exception, however – one girl mentioned a group that had talked about suicide.

### Preferential Treatment for Some Girls

Interestingly, those few cases of preferential treatment that respondents discussed were in reference to themselves receiving it and the reasons for it being either because they were older or because they had been in the program a long time. One girl was discussed as receiving preferentially poor treatment (being “smashed” on) and it was agreed by that respondent group that it was because that girl “kept screwing (up) repeatedly”.

### Program Addressed Differences in Backgrounds

Only one respondent made mention of racial issues in regard to this question, and in that case, to note that race is not usually a concern (although a couple of girls noted that a group activity on race/ethnicity had been offered). The issue that arose more commonly was economic background, in which case, those who were perceived as coming from a higher economic level were more likely to be the target of comment.

### Safe and Supportive Environment

The most common response was that Discoveries offered an emotionally supportive environment. Staff and other clients particularly offered emotional support when a specific family crisis arose for some girls.

*“They give you support when you really need it.”*

*“When \_\_\_ was there with us, if she saw somebody not getting along with somebody else, she would sit down with you and talk to you.”*

*“They are always there to talk to. If you wanted to call a little bit earlier, you were allowed to call. They were always willing to give it their best to what the certain situation was you might be in. They kept everything confidential. Like if you told another girl and you told them not to tell, they wouldn’t tell. They were really good about that.”*

*“When my \_\_\_ died, .... They were just really understanding of my emotional needs and I always could talk to somebody. ... The whole group was like are you okay; do you want to talk about it?”*

Although some subjects reported feeling physically safe, others offered comments on sometimes feeling less than safe in regard to other clients and/or the house.

*“...Because, not that we weren’t afraid of anybody in there, but it was the fact that they would leave us outside by ourselves. They would leave us by ourselves outside or in the living room where at any point in time somebody could jump somebody.”*

*“It’s safe, but they don’t have locks on the doors or anything. Or they do, but they aren’t locked. People walk in and out of here. People ... know this is a girls’ home and girls are here. ... I don’t think it’s too safe.”*

*“It’s not that they don’t address it. There is this one girl there who says, ‘I don’t like you so get the hell out of my face, go away.’ They just don’t do anything really unless it’s something with staff, like if the staff tells them to do something and they say they won’t then they get upset because they are saying no to them. But when something happens between us, they are like, ‘oh’.”*

### **Program Addressed Sexuality**

Girls reported talking about sexual topics on a regular basis, both in groups and “at lunch”, particularly that they shouldn’t have sex and/or that if they do, they should use protection. Program restrictions around sexuality were noted, such as “appropriate” clothing and for some girls, not being alone with boyfriends. Other topics included STDs, HIV/AIDS, and sexual orientation and sexual identity. Most respondents seemed to have no reluctance to discuss this question, and had a lot to say about it. Some disagreed with the encouragement they received to avoid sexual activity, but none complained about being urged to use protection or birth control.

*“I think they are kind of mad because we do have sex. I think they are happy that we use condoms.”*

### **Program Addressed Abuse Issues**

All of the respondents reported frequent and regular discussion of all sorts of abuse issues, with the exception of one girl who said, “I don’t know, maybe half the time”. Physical and sexual abuse were mentioned most frequently, as was abuse from males/boyfriends. Abuse was reportedly discussed in both groups and in individual sessions. One girl also commented on one of the part-time therapists being “a real advocate for spousal abuse and child abuse and things like that”. Two respondents noted that parental abuse was not discussed and/or not discussed as much as boyfriend abuse. On the other hand, a different girl said family abuse was discussed, but not rape or date rape. One respondent indicated that the type and nature of the topics depended on who was in the group and their individual needs.

*“What they normally do is put things together with what people need to work on, then make groups out of that.”*

### **Suggestions for Staff on How to Improve the Program**

*“I think the individual rules would have been better. Like have a level system. The better you are doing in your program, the higher up you are, the more privileges you get. I know \_\_\_ is \_\_\_ (age), her curfew was midnight, and mine was midnight, and I was almost \_\_\_ (age). There is a lot of different maturity levels between \_\_\_ and \_\_\_ and me. Know what I mean? They should have individualized rules. Stupid rules like, you have to leave with your parents, well, I never got along with my parents so I wouldn’t want to leave with them. There should have been an individualized program, but it was all everybody had the same rules and everybody had to do this and that. For some people, it’s easier to not do something than for others. You set a bunch of rules that aren’t hard for a person to follow, then they aren’t going to be challenged. You should have rules according to that person.”*

*“I do think they should have taken into account the way people learn.”*

*“Just enforce what they say they are going to enforce in the beginning and not let up.”*

Two girls expressed some anger and hurt feelings over staff “whispering” about them to other staff during “quiet time” where other clients could hear. They were upset that the other staff were being informed about something before it was discussed with them personally, and because other girls could hear what was being said.

Interestingly, although nearly everyone commented during the interviews about the clothing and make-up restrictions at the house, the only real complaint made was about having to do so on family night, which was seen as “embarrassing”.

Lastly, a suggestion was made to have a current advisory group, different from a monthly group of former clients. This group would be composed of current clients who would meet and make suggestions on the program during their own time in treatment.

## **B. Summary of Staff Interviews**

### **Staff Qualifications**

All staff interviewed have at least a Bachelor’s degree in the social sciences, primarily social work. The Program Coordinator has an MSW. All had some prior experience working with adolescents (one worked with the program part-time before becoming a regular staff member). Only one individual had prior formal education or training focusing on gender issues (a minor in women’s studies), although another person noted that such issues are often embedded in various college courses. One staff member has been with the program since it opened.

### **In Service Training**

The agency as a whole requires and provides for a certain number of hours of continuing education of all of its employees. Discoveries staff take advantage of a range of training opportunities that arise in the community and in the state. Training in gender-specific programming did not seem to be a priority over other adolescent treatment topics, although there was some implication that it might have been earlier in the evolution of the program. Two (older) staff mentioned the unfortunate loss of a “binder” that seemed to serve as the program’s library of original gender-specific program materials.

### **Program Goals**

Typical initial response: to keep the kids in the community and out of residential treatment.

Also emphasized was meeting the individual goals for each girl

Other responses:

- To rehabilitate her so she can abide by the rules and the law
- To develop the whole girl (healthy body, healthy attitude, healthy spirit)
- To help her make better and different choices
- To build positive relationships with women
- To avoid/break dependency on males
- To develop self-image, self-esteem
- To teach them to be independent
- To help them go to college and get jobs

## What Makes a Program Gender-Specific for Girls?

Varied responses:

- Knowing that girls learn differently than boys do, less linear processing of concerns
- All female staff, and close enough in age so the girls can relate
- Creating an environment that is safe for girls to be heard
- Speakers on topics of relevance to girl (i.e., sexual responsibility, domestic violence)
- Sometimes using consequences that are less authoritarian and more meaningful (i.e., talking, writing assignments)
- Girls can be more themselves because boys aren't around to try to impress
- Recognizing how girls are different from boys (i.e. how girls can hold a grudge if you don't process concerns with them, girls need more 'self-assurance', girls express feelings more, etc.)
- Using the 11 principles of gender-specific programming, with "Relationship-Based" being the most important – staff developing a strong one-to-one bond with each girl that comes through the program

*"That is a lot of what gender-specific programs need, especially when working with females, is relationship-building. They have to feel that connectiveness and they have to feel safe, otherwise it's not going to work. For example, one girl that we've had ... she was pregnant, and we were the first people she told she was pregnant. I talked to her about it later. She said that she knew even if we were disappointed in her, she knew we wouldn't leave her. She didn't feel that same concept with her mom, friends, or probation officer. That's why she told us first, because she knew she could trust us and we would help her make the right decision and we would help her tell people that she needed some help. So to me that speaks a lot of the relationship building we have here."*

Having a physical atmosphere that sets the right mood and encourages relationship building (a clean, house-like setting)

## How "gender-specific" is Discoveries?

Each staff person went on from this point to be very frank about what they saw as the very strong aspects of their program as well as those issues they would like to see improve, including how the program does in addressing "gender-specific" issues. When asked to begin by offering a numerical rating of 1 to 10 on how well the program currently attends to the gender-specific needs of females, the highest was an 8 and one person offered a 6.5. The following sentiment seemed representative of the comments made in support of the ratings:

*"...as far as our intentions to create a strong gender-specific program, we are definitely at a 10. But it is so easy to get mired down in the daily details of keeping the program going and dealing with girls that have problem behavior. It's easy to focus on the behaviors instead of the major goals we have for the program. ...."*

## Program Strengths

Everyone mentioned something about having a very devoted staff – professionals who are committed to the goals, willing to be on call every weekend and work during off hours, who really care about the girls

*"It takes special people who are willing to do that, to give up their time on the weekends ... not everything is about money".*

Other responses:

Small staff makes relationship building easier  
 Case management structure allows relationship building, each girl has one staff assigned to her for individual attention  
 Girls can call case managers on the weekends  
 Cohort structure builds relationships within the groups  
 Graduation celebration  
 Youth advisory board  
 Adjusting program to needs of individual girls coming through – 2 people said this  
 Thoughtful groups, not just make-work  
 Work hard at creating a safe emotional atmosphere  
 Girls having to “dress out” (change into casual clothes and remove make-up)  
 Not giving up easily on the kids, always looking at what can we do now or differently with youth who has continuing problems  
 Positive peer culture aspects, relationship-building among the girls  
 School liaison position

*“Kids always remark that once they get used to the program and get used to the staff, it always goes fast and it’s not nearly as bad as they thought it would be. So I think that it’s a program that we teach them skills; I don’t think it’s boring; I think it’s somewhat interactive, where they can learn something even if they don’t think they’re learning something. They end up thinking it’s not as bad as they thought it would be.”*

### Program Changes Needed

More unstructured time for relaxed relationship building  
 More spirituality activities (mentioned more than once)  
 More space, for meetings with outside professionals and for graduation ceremonies  
 More/better training for back-up staff to replace regular staff who need a night off  
 More fun field trips  
 Lengthen the program, and perhaps even make it a residential program  
 Finding time to meet more consistently one-to-one with individual girls  
 Better able to address certain mental health problems (PTSD we do well; depression, anxiety disorders, etc., not so well)  
 Be more creative in working with kids when we don’t have much family support, especially when the youth is a DHS referral instead of JCO (i.e., can’t order detention as a consequence).  
 Be more consistent in assessing consequences (i.e. JCO youth might go to detention while DHS might not)  
 Recognize strengths more – be less problem-focused (in spite of all the pressures to do so)  
 More time for individual counseling with girls; more time for girls to spend with fewer professionals (too many professionals coming in sometimes for some girls)  
 Put more emphasis on the girls investing in each other, being more responsible and accountable to one another, and invested in the community; this requires more trust and thus the ability to do more in-depth groups

*“... Of course you have to have time to explore that depth. Because sometimes when you say, if they were to say something personal, it would be really hard for me to just leave it hanging. You don’t ever want someone to share something with you really personal and then your response is, “Ok, what do you have to say?” to the next person or “What can you share?” and you just leave it out there. They need some sort of response back about, “let’s explore that, that must have been a really tough experience, how did you end up handling that?” I want the girls to be thinkers; I think that’s the key. If they are thinkers, they are better able to process things, make good decisions, better choices. Able to*

*recognize that there are a million different opportunities, thoughts, actions, responses that they could make and they have the ability to think through them and be innovative, creative, and unique. Thinkers.”*

### **Addressing Sexism in Society**

Although not everyone was invited to speak to this issue, no one identified any program activities beyond the individual level. When given the opportunity, one person did discuss definitions of feminism, and gave multiple examples of situations in which the program helps individual girls deal with gender stereotyping and oppression in their lives. However, no examples of community, institutional, or societal level program activities were mentioned in the interviews.

### **Self-Disclosure with Clients**

All agreed it was an important tool when used properly, defined variously as:

- “mindfully”, with a purpose in mind
- not just taking up the client’s time
- on topic
- not overly sensational for teens (such as *perhaps* substance use or sexual activity, depending on the content, the group, the background, etc.)
- to invite the identification of feelings
- to share problem-solving strategies
- to build trust based on personal disclosure and shared experiences

### **All Female Staff**

All identified a concern that when a male enters the house, the girls all focus on him, and that in general girls are less likely to attend to their priorities when males are present. Two mentioned that girls will spend their time trying to impress the males, regardless of age. Past victimization and unhealthy relationships with men were also noted as a problem that gets in the way of treatment when some of the staff are men. One person, however, did discuss the possible advantages in having part-time male staff teach the girls about healthy relationships with men.

### **Staff Relationships**

All reported that staff don’t spend time together outside of work hours. Unlike some youth programs where most of the staff members are younger singles who socialize in bars, this group is mostly married or engaged and starting to have children. Most personal socializing among the staff currently appears to occur on-site over lunch, for those who are available at that time of the day.

Otherwise, the bulk of the relationship-building among the staff tends to occur either through goal-setting and activity development for each new client cohort, or through day to day shifts working with the girls. Until recently, however, most shifts with the girls (afternoon/evenings) were split into two teams that worked the same days each week, and activities tended to be developed separately by those individuals or teams. Therefore, despite this being a small staff of only 5 full-timers, the structure encouraged subunit bonds and boundaries that eventually seems to have caused rifts when disagreements arose on program issues.

Thus, there was a sense of quiet conflict among the staff. Concerns seem to center around issues regarding *how* to resolve disagreements, with some comments focused on how it's hard to get people to "agree to disagree", that everyone seems to want full consensus or for everyone to feel good about the decision before it gets made. Other comments, however, suggested that the decision-making and program development are not democratic enough, that the program is not growing as much as it might because ideas for change are not implemented, and that too much emphasis has at times been placed on management hierarchy. More than one person felt that efforts to improve the program and in particular to maintain its gender-specific content were inconsistent, at times getting lost in the day to day pressure of the job.

On the other hand, all noted a major positive change in scheduling had occurred that rotated staff during the week, allowing each person to work with each of the others. One person also noted that an integrated activity planning session had been held for the first time (as of the time of that interview) and that communication and consistency among staff had improved. Perhaps in part due to the small size of the staff, everyone consistently spoke of their concerns and one another very respectfully and professionally. One is left with an impression of their strong commitment to the program and what they are collectively trying to accomplish.

*"I would say this is a pretty close staff, that we are comfortable with (each other), which is why we were able to solve the scheduling thing because we were able to say, 'this isn't working'."*

From the Program Coordinator:

*"... It's really hard for a new person just starting. To be honest, when I hire people, beyond all criteria do I think they are going to be good at this job, the next thing I look for is will they fit in? In another setting, that might not be a consideration, but here it is a huge consideration."*

## VI. CONCLUSIONS AND RECOMMENDATIONS

One of the first conclusions that can be drawn in reviewing the project overall is that this has been a worthwhile endeavor that has actually yielded useful information. Thanks to helpful clients, cooperative program staff, good research tools, and detailed protocols, the project was able to produce affirmative results, in spite of a relatively short time frame, recruitment challenges, and small sample sizes. It is not uncommon for evaluation studies such as this to end with inconclusive results because no particular patterns of significance appeared in the data analyses, making it impossible to even speculate with any confidence about whether or not the program was having any effect on its clients. Reasonable funding support and a few key components made it possible for this one to reveal what some of the program's benefits and weak areas may truly be.

In the area of **school and/or job performance**, the results showed that the Discoveries girls' self-reported school attendance increased while they were in treatment. This result probably continued through into the follow-up period on average, and was an improvement that did not seem to appear (or need to occur) for the comparison group. The Discoveries girls also appeared to show an increase in their perceptions of school as meaningful and important, but the opposite seemed to be true at follow-up, in that their scores at follow-up were actually lower than their intake scores on that scale. Statistically significant improvements were made in CAFAS school-work domain scores.

Study Outcomes: Improved school attendance during treatment  
Somewhat improved school functioning  
Temporary focus on school as important/meaningful

In the area of **empowerment and self-efficacy**, statistically significant improvements were seen on several MACI subscales. The Discoveries girls' scores decreased on average on the Doleful, Identity Diffusion, and Self-Devaluation scales and increased on the Dramatizing and Egotistic scales. The improvements on the Doleful, Egotistic, and Identity Diffusion scales were all significantly better than what was observed for the comparison group sample. Three of the scale improvements were still observable at follow-up: Egotistic, Dramatizing, and Self-Devaluation. Egotistic continued to be significant, and Self-Devaluation was nearly so. No apparent change was observed for Discoveries girls on the Inhibited and Introversive scales. Possible improvements were noted on the Discoveries girls' self-report ratings for change in the past three months in regard to their feelings of self-efficacy as a girl, and self-efficacy as a person (increases were noted, but were not statistically significant). The improvements appeared to be greater than those shown by the comparison sample.

Study Outcomes: Decrease in doleful personality tendencies  
Increase in emotional self-expression and sense of entitlement  
*During treatment and continuing into follow-up:*  
Increase in understanding of personal identity, goals and values  
Improvement in self-confidence and self-image

In the area of **prosocial behavior and attitudes**, results were mixed. The results showing improvement in the MACI subscales of Conformity and Oppositional were statistically significant across time and between groups, and the Submissive, Forceful, and Impulsive Propensity scales appeared to show improvement over time but were not significant. However, the Unruly scale showed no apparent improvement, and the girls appeared to actually lose ground over time on the Delinquent Predisposition and Social Insensitivity scales. On Social Insensitivity, the increase during treatment was statistically significant. On most other measures in the evaluation, the girls appeared to do well (i.e., fewer criminal allegations during and after treatment, self-reported problems related to fighting with parents or at school or with stealing; self-reported crimes against people, property, or related to drugs), but only the CAFAS improvements were statistically significant. Changes during treatment in the CAFAS overall scores, the Community subscale scores, and the Behavior Toward Others subscale scores were all significant.

Study Outcomes: Increase in overt compliance with rules and the law  
Decrease in impulsivity  
*During treatment and continuing into follow-up:*  
Less rebellious or oppositional behavior  
Little or no improvement in traditional "prosocial" attitudes

In the area of **relationships**, those measures that were not covered in other areas also showed mixed results. A key positive outcome was a lasting statistically significant reduction in Family Discord, although the mean score remained moderately high. The between groups test on this measure was also somewhat positive in that it showed that the Discoveries girls' improvement was greater than the comparison group's, and the result was nearly significant. There was also a significant improvement in the CAFAS Home domain subscale scores, and on the differences in ratings between groups on the

changes the girls experienced in the past three months in getting along with parents and teachers, and in expressing anger. There was no apparent change in the past 3 months, as rated at discharge, in friendships or romantic relationships. The Discoveries girls did show a small improvement on a Family Cohesion composite scale over time, and the improvement was greater than that shown by the comparison group, although these results were not significant. Finally, the Discoveries sample's scores on the MACI subscale of Peer Insecurity *increased* on average during treatment; this result was not significant either, however.

Study Outcomes:

Little change in peer relationships, possibly more insecurity  
Little change in romantic relationships

*During treatment and continuing into follow-up:*

Somewhat less conflict with parents and other family members

In the area of **health and mental health**, a complex set of results were found. Statistically significant results for the Discoveries girls both over time and in contrast to the comparison sample included improvements on the Testwell subscales of Emotional Management, Occupational Wellness, and Spirituality & Values. The MACI Body Disapproval scale, the Testwell total score, and the Testwell Intellectual Wellness subscale were all significant over time for the Discoveries group as well and the Eating Dysfunctions scale was nearly so, but the Discoveries group's greater improvement over the comparison group on the Testwell total score and Intellectual Wellness subscale were not significant and the comparison group showed a similar level of improvement on the two MACI items. The MACI Suicidal Tendency subscale was very likely an important improvement to be noted, since it was significant over time on the T1-T4 within groups test and between groups on the T1-T3, and showed improvement on the T1-T3 within groups as well even though it wasn't significant. The results were promising on the MACI Depressive Affect as well, in that the Discoveries group showed improvements during treatment and into follow-up, and the t test between groups was not quite significant but nearly so.

Less promising but still positive results were those from the Testwell subscales of Emotional Awareness & Sexuality, Physical Activity, and Social & Environmental Wellness, which showed small improvements over time for the Discoveries group but no differences between the Discoveries and comparison groups. No change at all was apparent for the Discoveries girls in the MACI scales of Substance Abuse Proneness and Anxious Feelings, and the Testwell scales of Self-Care and Safety, despite a relatively high score on the Substance Abuse scale. On other substance abuse measures, the Discoveries girls self-reported number of days using alcohol decreased during treatment and their CAFAS scores in the Substance Abuse domain showed a significant improvement, but the number of days they reported using tobacco *increased*, and the number of their friends who use substances, their ratings on how wrong it is to use substances, and their self-reported frequency as a group of having problems with substance abuse did not change. Similarly, the results showed no change at all on the Testwell Nutrition scale, compared to an increase in scores by the comparison group (a statistically significant difference between groups), and the Discoveries girls average score on the Testwell dieting question decreased from intake to discharge and then increased at follow-up. As a group, they reported no change at discharge in eating problems during the past 3 months compared to intake. In one other area, the Discoveries girls appeared to partly lose ground. On the MACI Sexual Discomfort scale, the Discoveries girls' scores *increased* on average, a negative result, although the difference was not significant. On the Testwell safe sex question, their average ratings increased from intake to discharge, suggesting they were avoiding sex or using safe sex methods, but their scores at follow-up were in fact lower than their intake scores, suggesting that they were either saying what they were expected to say at discharge or lost what they

learned by follow-up. On the other hand, a lower percentage at discharge than at intake reported having had problems during the past three months with unsafe sex and with sexual abuse. Finally, one last set of measures showed that as a group, the Discoveries girls' reported problems related to self-harm, and their CAFAS scores on the Self-Harm and Thinking domain scales (which were pretty low to begin with), did not change between intake and discharge. Their CAFAS Moods/Emotions subscale scores, however, did improve over time; the difference was not significant but nearly so.

Study Outcomes:

Little actual change in sexual behavior  
Temporary restriction on alcohol use, increase in tobacco use  
No change in attitudes toward substance abuse or sexuality  
No change in physical activity, safety, or self-care activities  
Improved goal-setting & decision-making  
Increased knowledge of strengths, skills, career choices

*During treatment and continuing into follow-up:*

Enhanced body image & reduced eating dysfunctions, possibly as a function of improved self-image overall, not connected to a better understanding of Nutrition or Physical Health  
Reduced depression & hopelessness  
Enhanced self-confidence & personal belief system

A few analyses were also conducted across measures. One interesting result showed a significant correlation between the MACI Identity Diffusion subscale and the Testwell total scores for both the Discoveries sample and the no-service comparison sample. In fact, there were four MACI subscales that had significant correlations overall with the Testwell total scores: Egotistic, Conforming, Identity Diffusion, and Self-Devaluation. While the two instruments did show correlations on these key measures, few significant correlations were found between them and the number of service hours documented for the Discoveries girls on the group and individual session instruments. The correlations between hours and the Testwell scores were low, and there were only two tests that approached significance between hours and the MACI scales; those were both over .40 for the group session hours with both the Submissive and the Substance Abuse scales (in the latter case, the result being a negative correlation, as would be expected).

Taking the above quantitative results into account, along with the qualitative information provided by the staff and the clients, the following conclusions and recommendations are offered for the Discoveries program, other programs interested in developing and/or modifying gender-specific approaches for treating adolescent girls, and/or planners and policy-makers taking a look at how best to stimulate and fund best practices in the field:

**1. The girls who are completing the Discoveries program generally leave feeling a lot better about themselves than when they came in.** On the whole, they are stronger, less depressed, better able to express their feelings, more self-confident, have an enhanced body image and fewer eating dysfunctions, are better able to identify their skills and set goals for themselves, can speak for themselves, and are more likely to see themselves as capable and worthwhile young women.

**2. While in Discoveries, the clients are able to bring their behavior into line with societal and parental expectations, although their underlying attitudes may not be changing a great deal.** While in Discoveries and in some cases after they leave, the girls tended to reduce their alcohol use, stop

committing crimes, stop skipping school, report avoiding unsafe sex, argue less with their parents and teachers, and in general, show a decreased tendency toward impulsiveness and oppositional behavior. However, their tobacco use actually increases while in treatment, they may be paying less attention to the rights of others, and their *attitude* toward delinquency does not appear to change. Nor does there appear to have been any appreciable change in their relationships with peers and romantic partners, and possibly their sexual behavior, despite extensive discussion of such issues.

**3. Recognize that this is a pretty good mix of outcomes for a group of troubled adolescent girls today.** It could be argued that as soon as these girls have some restrictions lifted and they come of age, you'll have a good group of capable, self-confident criminals. On the other hand, when one considers the backgrounds and victimization experiences many of the girls have had, having an improved self-confidence and sense of self combined with an opportunity to make progress by virtue of simply staying in school and out of further trouble, is a very real step forward. The fact that they may be retaining some rebellious attitudes, and engaging in some unhealthy and even unsafe activities, says as much about the world they live in today and their attempt to cope with it as it does anything else.

**4. More attention needs to be given to mental health issues for girls before and after they enter the juvenile justice system.** Studies have shown that women who become involved in the justice system are very likely to have a mental health disorder (Teplin, et.al., 1996), including anxiety & panic disorders, PTSD, depression, obsessive-compulsive disorder, and substance abuse. The evaluation lends support to this research, with the Discoveries girls scoring *an average* of 69.6 at intake on the MACI Depression scale, not far from the 75 point cut off point warranting serious clinical attention. Although the t test comparing T1 with T3 showed a drop in means from 62.9 to 53.2, suggesting their involvement in Discoveries led to some improvement overall in that scale, the numbers clearly suggest mental health should be an area of concern for the program and the systems that are referring them. Less of a concern but also noteworthy was their *average* score of 47.6 for Anxiety at intake, which actually increased a little while they were in the program to 49.6 at T3, although their T1 mean of 47.6 on Suicidal Tendency dropped considerably to approximately 36. Many professionals involved with adult systems would now agree that attention to mental health is absolutely critical for helping women who have substance abuse problems and who are involved in criminal justice. It would only make sense that the earlier these issues are recognized in young people, the more likely they will be treated properly. The National GAINS Center for People with Co-Occurring Disorders in the Justice System is a recommended resource for information on and approaches to addressing the mental health needs of women and girls in the justice system.

Understanding that mental health problems frequently contribute to initial and continued involvement in the justice system is one thing. What is more disconcerting is the high scores on these scales by the comparison girls, particularly the supposedly "normal" no-services comparison group. While appearing stable and healthy on most other measures in the study, these girls actually scored *higher* on Anxiety on average (65.8 at T1) than both the Discoveries and the comparison-services groups, and stayed that way at every measurement point. While the Discoveries girls' Depression scores tended to decrease while in treatment, the "normal" girls did not, despite having an *average* of 66.1 at T1 themselves. What is it about these girls and/or their environment that leads to such a high incidence of anxiety and depression? Is high anxiety the price of successful functioning for adolescent girls currently? It would be very interesting to repeat the administration of the MACI with a larger sample of Iowa girls to see whether this finding would endure, and explore a bit more what the causes of this might be.

**5. Innovative approaches are needed to address sexuality and romantic relationships with teens in our society.** Unlike some areas, the study results did not show demonstrable change in attitudes or behavior among the Discoveries girls in regard to sexuality or romantic relationships, apparently despite extensive attention to such issues during the course of treatment, particularly ways to avoid it and ways to avoid pregnancy and disease. One of the Discoveries clients seemed to sum it up by saying, “they seem to get mad when we have sex but they’re happy when we use condoms”. Unfortunately, the results likewise did not provide any clear direction for program recommendations for Discoveries specifically.

A few comments on the subject in general are offered, however, for professionals and parents interested in the issues. Everyone agrees the pressure on kids to be involved in sex is tremendous, and youth that have been sexually abused are even more likely to become sexually active early. Yet, because of our ambivalence on the subject, our conflicting values, even our shame, the support we’re offering as adults is so meager compared to that pressure, it seems that kids still pretty much have to find their own way. To girls that have been deeply, profoundly, and perhaps irrevocably affected by having been sexually molested at a young age by a parent’s paramour, for example, we may offer a few months of a positive treatment environment at best, much much better than nothing, but much less than the long-term healing experience that is needed. To girls that are under pressure to learn proper oral sex techniques or must make a decision about whether or not to allow anal sex with a boyfriend, some adults advising them are still picturing a heated exchange after the dance in which the girl must resist his “natural” advances long enough to maintain her good reputation and self-respect. While more “progressive” professionals are often beyond that thinking, sometimes we too get stuck in a comfort zone where the discussion tends to revolve around the mechanics of prevention, whether it’s pregnancy, disease, (or simply sex) prevention. In any case, it’s still unfortunately common for many of us as adults to think we’re talking about the same thing as teens when we talk about sex, even when we’re worlds apart. Programs that use even current standard sex education approaches cannot hope to compete with the culture (and sometimes the families) the kids live in. Whether it’s the new abstinence movements, in-depth long-term sexual abuse treatment as a standard component of certain treatment programs for girls (and boys?), a new feminist revolution in our society, or at the very least sex education that starts where the kids are, something new needs to happen in this area.

**6. This evaluation overall lends enough support to the concept of “gender-specific” programming to keep moving and build on what’s been learned so far.** Discoveries includes many of the components recommended for gender-specific programs, and the evaluation shows that the program is producing a number of good outcomes. It is clearly a relationship-based program with a strong, committed staff; abuse/victimization, emotional/physical safety, female role modeling, sexuality/relationships, and career development are all among its primary concerns. This *suggests* that the outcomes observed are due at least in part to the gender-specific features of the program. While the evaluation design allowed only preliminary exploration of the question of whether or not gender-specific programs in general are better than those that are not, the results clearly show the program is performing better than no services at all and in fact, there is a little evidence to suggest a pattern whereby the Discoveries girls improved the most, the comparison-services girls some, and comparison-no services girls the least or not at all. This pattern appeared, for example, on the MACI Anxiety, Depression, & Suicidality scales, on the Youth Questionnaire Self-Efficacy, Family Cohesion, and self-reported improvement measures, and on the Testwell total change scores. More t test analyses of the subsamples are needed to determine whether or not the pattern is observable on other measures. Analysis of variance would be the next step also in examining questions related to this issue.

In the interim, professionals and policy-makers interested in developing and/or enhancing gender-specific programs for adolescent girls may wish to take note that the following components of the Discoveries program appear to be its greatest strengths, based on the outcomes reported above and the qualitative comments provided by staff and clients:

- \* Individual clients assigned to specific full-time staff member with 24 hr-7 day access by cell phone
- \* Intensive group sessions with daily/weekly topics planned for each cohort of girls
- \* Use of outside speakers and organizations from the community
- \* Staff is predominately comprised of strong independent women educated in helping professions, with some racial/ethnic diversity (part-time staff)
- \* Family therapy offered as an adjunct support
- \* Topics/ideas promoted intensively include:
  - Personal boundaries, self-respect
  - Self-image and self-confidence
  - Freedom from traditional gender role expectations & oppressive media stereotyping
  - Healthy emotional responses and communication, including anger and grief
  - Trauma and abuse/victimization avoidance and healing
  - Sexual restraint, protection, and birth control
  - Goal-setting, career development and college planning
  - Healthy romantic relationships, including how to relate with boys and men
  - Behavior management at home, school, and in the community
  - Family relationships, getting along with parents
  - Healthy eating habits, avoiding eating dysfunctions
- \* Relationship-building within the group, among/between girls and staff
- \* Full-time school liaison staff position maintains communications with the schools, ensures successful experience for each girl in her home school while in treatment
- \* Enforcement of personal space, appropriate touch among everyone, including clients with each other
- \* Policy of casual clothing for clients provided by the program, and no make-up or jewelry while on site
- \* Policy of no communication between clients outside of the program while in treatment
- \* Policy of no casual communication (gossip) among and about peers in or outside the program
- \* Transportation provided to and from the program, as well as meals on site

- \* Monthly family social night on site
- \* Behavioral monitoring through case managers on call in communication with families/parents;
- \* Consequences for behavioral infractions are assigned for both on- and off-site (depending on cooperation from families and systems), including writing assignments, weekend curfews, detention stays, etc.
- \* Physical environment that is homelike & comfortable, yet includes positive materials on gender, race, and behavior

Other less emphasized components:

Racial/ethnic diversity unit  
 Menu planning & meal preparation  
 Journaling & client newsletter  
 Staff training and policies on physical restraint, crisis intervention, drug free workplace, child abuse/criminal background checks, etc.

**7. The following recommendations are offered by the Research Director for enhancing the Discoveries program, based on her interpretations of the qualitative and quantitative study results:**

- \* reinstitute regular spirituality components
- \* add a fine arts component as an aspect of spirituality enhancement
- \* require training in mental health diagnoses and treatment for staff
- \* purchase on-site psychiatric nurse visits for staff consultation and diagnostic purposes
- \* however, try to minimize the number of different professionals each youth is involved with
- \* allow the girls to dress and wear make-up at whatever level the staff does (staff wear sweats also or allow the girls to wear jeans if that is how the staff dresses, etc.) and in any case, allow the girls to dress and make up as they choose on family night
- \* offer a physical self-defense class once in each cohort
- \* experiment with a new component that involves girls in gender- or race-based community action (i.e., Women for Peace, or a project to change a school policy on athletics) that gives them experience in community activism (appropriate “rebelliousness”), volunteerism, group belonging, etc., while working to end discrimination or some other just cause
- \* develop an *ongoing* (long-term/two years) weekly aftercare support group for all former clients, with incentives to attend (i.e., paid stipends and/or court orders) to encourage retention of program outcomes

- \* lock the doors to the house and require visitors to ring the bell to get in
- \* be careful not to assume the girls feel 100% physically safe with one another, even with staff in the next room (they may or may not be, and even if they are, they may or may not feel that way)
- \* incorporation of *carefully*-constructed positive male relationship-building, possibly into long-term aftercare component (i.e., co-ed facilitators)
- \* develop mechanisms for encouraging unified program vision and growth, such as: 1) conduct integrated all-staff planning sessions for each cohort's groups and activities; and 2) shut down the program for one or two days annually and have a staff retreat, the first one to be facilitated by an outside person, agenda to be "visioning", program planning, staff renewal, etc. (Multi-purpose: reestablish shared gender-specific vision, strengthen staff relationships, revitalize staff and the program components).

**8. The extent to which a program is "gender-specific" is not just an issue at start-up.** Although the evaluation has shown some positive client outcomes and it's clear Discoveries is doing good work, there remains an elusive sense of possible erosion in the "gender-specific" nature of the program design. The impression is an elusive one because it would appear that many of the components recommended for gender-specific programs for girls are in place, positive outcomes for girls are occurring, and staff seem very committed to the concept. However, staff do seem to have varying ideas on what it is that makes boys and girls different, and a number of the components recommended for gender-specific programs that were once in place seem to be less emphasized currently, most notably spirituality, although nutrition education and physical activity are missing as well. As noted by one staff member, it can be very difficult to maintain a strengths-based approach to treatment, and given the program's outcomes pertaining to self-confidence, that would seem less of an issue than some others. It's interesting that the lost resource notebook with the compendium of gender-specific training materials has never been replaced, and that training for new staff does not seem to be as grounded in the state's gender-specific "movement" as it once was. There were few indications that gender oppression as a social problem was to be addressed by the program, and clearly no expectations that the program would attempt to have an impact at any level other than the individual girl (or possibly her immediate family). It would have been interesting to learn if any attempt, recently or in past years, was made to get a school policy changed even on behalf of an individual girl, let alone the entire population of females in a given school (the school liaison was not interviewed). Of course others might not agree this is an appropriate level of intervention at all, and that certain details have been missed in the review, and that some details have been emphasized over others in support of this concern and recommendation. All of that may be true.

The conclusion and recommendation is simply this: Organizations and programs all evolve over time. They have an initial growth period, a maintenance period, and then they begin to stagnate or a period of new growth begins. Let's say enthusiastic program developers get excited about this new gender-specific concept (especially a charismatic director as in the case of Discoveries), start up some new components or approaches that are embedded with other more traditional-looking approaches ("good gender-specific services begin with good services"), and then time passes, some staff turnover occurs, and the press of everyday work occurs, kids having day to day problem behaviors that must be dealt with one way or another. Perhaps some aspects of the new approach begin to lose their urgency and some of the traditional ways of doing things begin to slip back in. Perhaps some angst is expressed about that among the older staff while at the same time, the newer staff begin feeling somewhat stifled, both because not

enough new planning and revitalization is taking place to refresh the program concept. Perhaps over time the program begins to look a little more traditional than anyone wants it to, but no one feels like they have the time or the power to do what it takes to change things. The good news is there are ways to deal with organizational change, once the need is identified. For Discoveries, the recommendation therefore is to take a look at this, and begin by following the specific program planning suggestions offered above. For other agencies and programs, the recommendation is to build in mechanisms pertaining to this issue in order to manage what is a very natural process (sort of a reversion to the mean), such as ongoing strong gender-specific training requirements, annual staff visioning retreats, and perhaps even external evaluations like this.

**9. Continually challenge yourselves not to see “compliance” or “conformity” as the only or even primary goal of treatment or juvenile justice, especially for girls.** As parents, social workers, juvenile court officers, and policy-makers, often our first objective with kids is to design a way to make them follow the rules. We know how important it is to instill in them a sense of right and wrong, and to teach them to respect others, and we also know how important it is for them that we be consistent in our enforcement of the rules, and so we try to use the systems we’ve set up to advance those principles, forgetting that those systems were not necessarily designed with girls’ needs in mind, or even how best to address the root causes of delinquency. We assume that conformity or compliance is a good choice, when in fact rebelliousness may be the very thing we want girls to learn in a family or a school or a society that abuses and devalues them. Noncompliance in some situations is an appropriate response! The fact that girls don’t always know when and how to express their feelings of rebelliousness appropriately is to be expected in such a world, yet the longer they are in school, the better relationships they have with others, and of course, the longer we can keep them out of lock-up, the better off they are. Therefore, it would appear that the primary goal is not compliance, but in fact, *discernment*. Teaching them to be able to discern when and how to rebel, and when and how to conform, is the real issue. And of course, in order to do that, we have to know when and how ourselves.

## VII. REFERENCES

Bates, M., 2001. “The Child and Adolescent Functional Assessment Scale: Review and Current Status,” Clinical Child and Family Psychology Review, 4(1), 63-84.

Hodges, K., 1997. Child & Adolescent Functional Assessment Scale (CAFAS). Functional Assessment Systems: Ann Arbor, MI.

Hodges, K., 1995. CAFAS Self-Training Manual. Author: Ann Arbor, MI.

Hodges, K. & Wong, M., 1996. “Psychometric Characteristics of a Multidimensional Measure to Assess Impairment: the Child and Adolescent Functional Assessment Scale (CAFAS),” Journal of Child and Family Studies, 5(4) 445-467.

Iowa Commission on the Status of Women, 1999. Providing Gender-Specific Services for Adolescent Female Offenders: Guidelines and Resources. Iowa Commission on the Status of Women, Department of Human Rights: Des Moines, IA.

Iowa Commission on the Status of Women, 1997. Female Juvenile Justice Report. Iowa Commission on the Status of Women, Department of Human Rights: Des Moines, IA.

Millon, T. with Millon, C. & Davis R., 1993. Millon Adolescent Clinical Inventory & (MACI) Manual. NCS Assessments & NCS Pearson, Inc.: Minneapolis, MN.

Millon, T., 1991. "Classification in Psychopathology: Rationale, Alternatives, and Standards," Journal of Abnormal Psychology, 100, 245-261.

Millon, T., 1969. Modern Psychopathology: A Biopsychosocial Approach to Maladaptive Learning and Functioning. W.B.Saunders: Philadelphia, PA.

National Wellness Institute, 1999. Testwell: Wellness Inventory for Adolescents. National Wellness Institute, Inc.: Stevens Point, WI.

Office of Juvenile Justice and Delinquency Prevention (OJJDP), 1998. Guiding Principles for Promising Female Programming: An Inventory of Best Practices. Office of Juvenile Justice and Delinquency Prevention, US Department of Justice: Washington, D.C.

Quist, R., 2000. "The Child and Adolescent Functional Assessment Scale (CAFAS): A Dynamic Predictor of Juvenile Recidivism," Adolescence, Spring, 2000.

Teplin L, Abram K, and McClelland, 1996. "Prevalence of Psychiatric Disorders Among Incarcerated Women," Arch Gen Psychiatry, 53, 505-512.