University of Rhode Island

DigitalCommons@URI

Open Access Master's Theses

2000

Understanding the Efficacy of Residential Treatment for Female and Alaskan Native Adolescents: Outcome Study and Meta-Analysis

Kathleen Cady-Webster
University of Rhode Island

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

Recommended Citation

Cady-Webster, Kathleen, "Understanding the Efficacy of Residential Treatment for Female and Alaskan Native Adolescents: Outcome Study and Meta-Analysis" (2000). *Open Access Master's Theses*. Paper 1677.

<https://digitalcommons.uri.edu/theses/1677>

This Thesis is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.

BF724.3
S36
C339
2000

UNDERSTANDING THE EFFICACY OF RESIDENTIAL TREATMENT FOR
FEMALE AND ALASKAN NATIVE ADOLESCENTS: OUTCOME STUDY AND
META-ANALYSIS

BY
KATHLEEN CADY-WEBSTER

A THESIS
COMPLETED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS IN PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND

2000

#45824294

Abstract

Today adolescent violent crime and juvenile delinquency is an increasing problem in our society. Along with the high number of minority offenders, there is also evidence from state and local agencies that the number of female offenders is also increasing. Once youth enter the family court system, they are often adjudicated for treatment into a residential facility. There is little information available regarding the efficacy of these types of placement especially for female and minority problem youths. Phase I of this project involved the examination of such a facility utilizing a longitudinal study spanning four years of data collection. Year four included data on 21 (66%) of the original 32 youths. Forty-four percent (14) of participants were female (nine Caucasians and five Alaskan Natives) and 56% were male (nine Caucasian and nine Alaskan Native) resulting in 44% of participants reporting minority status.

Scores on the Achenbach Youth Self-Report (Achenbach YSR: 1991), Rosenberg Self-Esteem (Rosenberg, 1965), and FACES III (Olson, Porter & Lavee, 1985) were subjected to repeated measure ANCOVAs and Multiple Regression. As hypothesized, it was found that Alaskan respondents reported significantly higher levels of problem behaviors and lower levels of self-esteem. When overall efficacy was examined by group, results indicated that females, and especially Alaskan females, reported no significant increases in levels of self esteem from intake to year four. Post-hoc tests indicated that Alaskan females reported significantly lower levels of self-esteem than Caucasian females at the 12 and 48-month interviews. Alaskan males experienced lower levels of self-esteem than Caucasian males until year four at which time there was virtually no difference between these two groups. Alaskan females remained within the clinical range on total scores of the YSR at all follow-ups with the exception of the 36-

month when they were slightly below. Alaskan males also reported significantly higher levels of problem behaviors than Caucasian males at the 24 and 36-month interviews and just missed significance at the 48-month. Overall Multiple regression revealed that being White and being Male was a significant predictor of higher levels of self-esteem. Overall effect size within the outcome study was quite large for the field with eta-squared = .18, which coincides with a large effect of .80 as specified by Cohen's d (1988). Average power within the outcome study was also somewhat high in the field .60, but well below the recommended level of .80 (Cohen, 1988). Discussion includes addressing cultural and gender development as important indices of outcome, and methodological issues were explored.

Phase II involved the completion of a meta-analysis where 16 psychological journals associated with the treatment of problem youth containing 3600 citations published between 1994 - 1999 were searched for key words related to juvenile delinquency and residential treatment. This effort resulted in the examination of 700 abstracts leading to the selection of 86 (12%) articles including information on residential treatment specifically. Of those eighty-six articles included, 26 (30%) discussed outcome, while only 12 of these (46%) included the relevant information to be included in the statistical analysis (pre-post and or comparison groups and the necessary summary statistics). The 12 articles included 306 statistical tests with an average power of .25, .68, and .87 for detecting effect sizes for Cohen's d of small, medium, and large effects respectively. Only 3 articles offered enough information to determine effect size resulting in an average effect of .26. Corrections for bias due to sample size (Hedge's g^u , 1985) lowered this to .25. The overall meta-analysis revealed a significant Z of 8.83 indicating positive outcome between pre-post tests and/or comparison groups.

Rosenthal's fail-safe N indicated that 333 null findings completed between 1994 - 1999 would be necessary to make the overall Z score non-significant. Unfortunately, only six of the studies included outcome associated with females and only 2 for minority youths. While the number of articles including information regarding outcome for females was small limiting reliability, the overall Z score associated with these articles was also significant $Z = 3.56$. This finding is suspect as it does not meet the tolerance level for null effects (Rosenthal, 1979) as only 22 would be necessary to threaten this significant finding. Methodological issues, treatment modalities, limitations, and future directions are discussed.

Phase III of this project includes an Epilogue that explores the philosophical debates existing in the field of scientific inquiry. It highlights the importance of expanding our existing methods and ontological positions so that a more comprehensive "truth" may be evident: one that may not be possible while relying on a mechanistic and objective model of reality and dichotomize testing. It directs our attention to what science was, what it is now, and speculates as to what it could be in the future. These expansions in thought and method are then applied to the findings within this paper.

Acknowledgements

Getting here requires not only dedication and commitment, but it also requires the help of many. I would first like to thank Dr. Barbara Minton from the University of Alaska. As a returning student, it was through her encouragement and inspiration that I was able to navigate my way through my undergraduate work and came to realize my potential. The lesson she taught me was to lift as you climb. This has been the inspiration behind my dedication in working with undergraduates. I too may help someone find their way. I would also like to thank the youths that participated in this study. This work would not have been possible without their efforts.

I would also like to acknowledge the support I received as a non-matriculating student here at the University of Rhode Island. I returned to Rhode Island to help my mom with homecare for my terminally ill brother. I held onto my dream of going to graduate school and took classes here. Dr. Bernice Lott, Dr. Al Lott, and Dr. John Stevenson taught those classes. My acceptance into this Graduate Program is a direct result of their support and belief in my ability beyond that represented in GRE scores. Their efforts, along with Dr. Valentino's support, gave me the opportunity to prove that I was capable of excelling in such a challenging environment.

Additionally, I would like to thank all of my professors whose dedication and commitment to teaching continues to astound me. I would especially like to thank Dr. Rossi for his continued dialogue and guidance, thought provoking questions, somewhat quirky sense of humor, the use of his power programs, and his great Sushi! While I'm at the CPRC, I would also like to thank Dr. Jay Maddock and Trish Jorden for sharing both their knowledge and products of meta-analytic procedures with me. Your work served as wonderful examples that helped clarify many of the procedures necessary to complete

this paper. It is important to acknowledge my gratitude to my friend and undergraduate assistant, Rachel Kaplan. Rachel's flexibility, dedication, and swiftness in completing tasks assisted in keeping this project manageable and on schedule. Her continued patience and laughter were always a joy, even at the most trying of times. Thank you Rachel.

This process is also made easier as a result of the friendships that are made along the way. I would like all of you to know that I truly value our friendship and look forward to a lifetime of connections. Yes, that means you too Bill.

This brings me closer to heart of my training. I would like to first and foremost thank my major professor, Dr. Lisa Harlow. Her incredible enthusiasm for the complex world of statistics and teaching is contagious. It was through her tutoring and example that I have come to love methodology. She has taught me that statistics are more than just numbers. They are potentials that are capable of revealing complex phenomenon. I am honored that I am within her direction and scope. In addition, Dr. Harlow has provided me with more than just knowledge. She has been a great source of emotional and spiritual support as well. This has meant more to me than she may ever know. I walk lighter as a result of her efforts.

I would also like to thank my committee members for their time and effort in guiding my work. Dr. Grebstein's pragmatic approach to problem solving will help to keep me grounded in the applicability of my work. In contrast, I would like to thank Dr. Burbank for her ability to allow one to explore realms of reality that are not visited by many in the field of scientific inquiry. I have a purer understanding of myself and my work as a result of these journeys in understanding one's philosophical position. I would

also like to thank Dr. Adams for committing his time and energy to a student that he has yet to meet!

Finally I would like to thank my family for their incredible support. Without them, my life would be hollow. You are my foundation. I would also like to extend my love and appreciation to Michael. As I have pursued my dreams, he has endured years of lonely nights and summer visits. He has allowed me to fly without guilt and without fear of losing what we have. All of you have embraced me, and I am honored.

Table of Contents

Abstract	ii
Acknowledgements	v
Table of Contents	viii
List of Tables	ix
List of Figures	x
Chapter 1: Introduction	1
Chapter 2: Phase I	12
Methods	13
Results	19
Descriptive	22
ANCOVAs and Self-Esteem	23
ANCOVAs and YSR	29
Multiple Regression	35
Chapter 3: Phase II	40
Methods	48
Results	50
Effect Size	53
Power Analysis	56
Meta-Analysis	60
Chapter 4: Discussion Phase I	63
Discussion Phase II	73
Contributions and	
Future Directions	77
Chapter 5: Phase III: Epilogue	80
References	105
Appendices	
A: YSR	113
B: FACES	118
C: Rosenberg Self-Esteem	120
D: References of Retrieved	
Articles for Meta-Analysis	122
Bibliography	134

List of Tables

Phase I

Table 1:	Expected Group Differences in Phase I	13
Table 2:	Original and Adjusted Mean Scores and Standard Deviations for the Rosenberg Self-Esteem Total Scores	20
Table 3:	Original and Adjusted Mean Scores and Standard Deviations for Youth Self-Report Total Scores	20
Table 4a:	Standard Multiple Regression of Problem Behaviors, Early Sexual Contact, Family Cohesion Gender, and Ethnicity on Self-Esteem at Initial Interview	36
Table 4b:	Standard Multiple Regression of Problem Behaviors, Early Sexual Contact, Family Cohesion, Gender, and Ethnicity on Self-Esteem at 12 Months	37
Table 4c:	Standard Multiple Regression of Problem Behaviors, Early Sexual Contact, Gender, and Ethnicity on Self-Esteem at 24 Months	38
Table 4d:	Standard Multiple Regression of Problem Behaviors, Early Sexual Contact, Gender, and Ethnicity on Self-Esteem at 36 Months	39
Table 4e:	Standard Multiple Regression of Problem Behaviors, Early Sexual Contact, Gender, and Ethnicity on Self-Esteem at 48 Months	40

Phase II

Table 5:	Article Demographics, Gender, and Ethnicity	53
Table 6:	Summary of Effect Size Analysis	55
Table 7:	Articles Included in Power Survey of Meta-Analysis	58
Table 8:	Results of Meta-Analysis	61

List of Figures

Figure 1:	Logic Model of Program Components and Expected Outcome	17
Figure 2:	Mean Scores Associated With Levels of Self-Esteem Across Time by Group	24
Figure 3:	Mean Scores Associated With Levels of Self-Esteem by Ethnicity Across Time	25
Figure 4:	Mean Scores Associated With Levels of Self-Esteem Across Time for Females	26
Figure 5:	Mean Scores Associated With Levels of Self-Esteem Across Time for Males	27
Figure 6:	Mean Scores Associated With Levels of Self-Esteem by Gender Across Time	28
Figure 7:	Mean Scores Associated With Reports of Problem Behaviors by Ethnicity Across Time	29
Figure 8:	Mean Scores Associated With Problem Behaviors by Group Across Time	30
Figure 9:	Mean Scores Associated With Levels of Reported Problem Behaviors Across Time for Females	31
Figure 10:	Mean Scores Associated With Levels of Reported Problem Behaviors Across Time for Males	33
Figure 11:	Mean Scores Associated With Reported Problem Behaviors Across Time for Gender	34

Understanding the Efficacy of Residential Treatment for Female and Alaskan Native Adolescents: Outcome Study and Meta-Analysis

Statement of the Problem

Today adolescent crime and juvenile delinquency is an increasing problem in our society. Along with the high number of minority offenders, there is also evidence from state and local agencies that the number of female offenders is also increasing. Once youth enter the family court system, they are often adjudicated for treatment into a residential facility. O'Mally (1996) informs us that "Residential treatment is being offered as a treatment modality all over the country by a greater number of providers than ever before" (p. 62). Until recently, there has been little investigation into the effectiveness of treatment programs designed to address this particular placement. There is even less understanding regarding the effectiveness of residential treatment with regard to the female and/or minority adolescent. Additional research suggests that overall, residential facilities are not necessarily predictive of positive outcome (Wells, 1991; Asarnow, Aoki, & Elson, 1996). Researchers in the field posit that before residential treatment can be predictive of positive outcome, a comprehensive examination of "practice, policy, treatment components,...and research methods" must be forthcoming (Wells, 1991 p, 326). Most of the outcome studies that are available report findings primarily associated with male outcome with little or no attention being given to gender or ethnicity. However, if female and minority adolescent crime is an increasing problem in our society, and the

courts are placing these youth into residential treatment as a means of referral, then efforts must be made to identify the treatment modalities that are effective and meaningful for these populations as well.

Support for the Problem

Through the collaborative efforts of the University of Alaska and a non-profit adolescent residential treatment facility, I held the position of senior research assistant and program demographer on a longitudinal treatment outcome study between 1992 - 1997. The purpose of this project was to determine the efficacy of this facility by examining outcome for youth that were residents. In the course of writing annual reports for this project, descriptive statistics revealed that many females in the study were reporting lower levels of self-esteem after treatment as compared to intake on a number of items on the Rosenberg Self-Esteem Test (1965). Females also reported experiencing more problem behaviors on the YSR, Achenbach Youth Self-Report (1991) than did the males.

The first step in the current project is to examine the significance of these differences within this sample. Specifically, analyses will seek to investigate how residential treatment may impact female and Alaskan Native adolescents' outcome in comparison to that of White males. It will be determined if differences in self-esteem and reported problem behaviors are significantly different across groups, as well as examine treatment practices that may be more predictive of higher levels of self-esteem and lower levels of problem behaviors indicative of more positive outcome.

Due to the fact that this population is an intake sample into residential care, the number of participants is quite small (n=32). However, the literature review completed

for this project indicated that this number is respectable in the field. Nonetheless, it does limit statistical options. Therefore, this thesis will include three phases. Phase I is the outcome study introduced above. Phase II includes the completion of a meta-analysis regarding the efficacy of treatment outcome for females and minority adolescents receiving residential care. The purpose of the meta-analysis is to allow for a comprehensive understanding of treatment outcome with respect to treatment modalities and outcome measures utilized by facilities. In addition, it helps to establish statistical norms with regard to effect size and power of such research. The meta-analysis also allows for the identification of important outcome and mediating variables that may be interacting with treatment modalities affecting outcome for female and minority adolescents. Discussion includes the theoretical issues associated with outcome for females and minority adolescents as presented in the introduction. To continue illustrating the need for multiple assessments and perspectives regarding the "theoretical, empirical, and practical" meanings associated with our findings (Harlow, PSY 533 discussion, 2000), this paper then concludes with an Epilogue (Phase III) that examines the ways in which our ontological and philosophical scientific positions, either as a scientific community or as an individual scientist, can, many times, dictate the questions we ask. Consequently, this orientation determines the methods we use resulting in a truth or a reality that may not be inclusive for all. Methodological, philosophical, and ontological expansions are discussed based on the ideas inherent in the "New Physics." To keep it practical, these ideas are then examined in the context of this paper.

Justification for the Study

For nearly two decades, research on adolescent residential treatment facilities have not been necessarily predictive of positive outcome after discharge (Wells, 1991; Asarnow, Aoki, & Elson, 1996). There is much debate in the field about whether this lack-luster treatment outcome is due to the treatment modalities adolescents receive while in the residential facility, or if it's associated with individual characteristics of the residents themselves. Others suggest that perhaps it is a result of inadequate or absent aftercare support mechanisms (Curry, 1991; Blackman, Eustace & Chowdhury, 1991). While still others conclude that inconsistencies in efficacy are the result of methodological problems in the field (Zimmerman, 1996; Pratt & Moreland, 1996). These researchers advocate the use of outcome research, but also find fault with some of the rigorous methods inherent in traditional scientific inquiry such as the sole reliance on significance testing over qualitative analysis, utilizing statistical methods that demand large samples, and adherence to randomization as evidence of a sound design. They go on to contend that the goal of outcome research "extends beyond evaluating the general question of whether psychotherapy [or any other treatment] is effective. . . [and] should be to improve the quality and efficacy of treatment in order to reduce suffering in children and adolescents" (Pratt & Moreland, 1996, p. 3). Zimmerman (1996) and Pratt & Moreland (1996) go on to suggest that the field has been remiss in providing a detailed description of those treatments that may be most effective. While this may be true, the current project seeks to take the ideas of Zimmerman (1996) and Pratt & Moreland (1996) one step further and ask, if these treatments are effective, then define for whom they are effective.

Searches have indicated that there is evidence regarding the efficacy of treatment outcome for male residents in the literature (Asarnow, Aoki, & Elson, 1996; Wright, 1995; Valliant, 1993; Small, Kennedy, & Bender, 1991; Epps, 1991; Burnett & Rathbun, 1993). However, closer examination reveals that many of the outcome studies and treatment modalities utilized with juvenile offenders do not take into consideration treatment issues relevant to the female or minority resident. To illustrate this point, twenty-four articles associated with residential treatment were initially examined to show support for this project. Of those articles, seventeen specifically discussed the efficacy of residential treatment between 1990 and 1996. Only two (11%) of these articles (Hiller, 1996; Mennen & Meadow, 1993) offer outcome and treatment modalities specific to the female resident. Three articles (18%) offered a cursory comparison of male and female outcomes (Gutierrez, Felipe-Russo, & Urbanski, 1994; Gutierrez & Todd, 1997; Scheidlinger, 1992). Seven of the 17 articles (41%) were specific to the male resident (Asarnow, Aoki, & Elson, 1996; Wright, 1995; Valliant, 1993; Small, Kennedy, & Bender, 1991; Epps, 1991; Burnett & Rathbun, 1993; Wells, 1991). The remaining 5 articles (29%) lumped participants together and did not distinguish outcome between genders (Blackman, Eustace, & Chowdhury, 1991; Burks, 1995; Pfeiffer & Strzelecki 1990; Curry, 1991; Duchnowski, Johnson, Hall, & Kutash, 1993). No article examined the efficacy of their program with regard to ethnicity.

This lack of available information on treatment outcomes for female and minority youth in residential care is of great concern. Researchers involved in this area suggest that treatment modalities for adolescent females need to address issues specifically associated with the female resident. For instance, Scheidlinger (1992) investigated the relationship

of gender identity with the structure of child and adolescent treatment groups. He found that "When groups of any age contain girls, a female therapist is desirable" (p. 107). There is also an indication that young female adolescents may do better in same-sex groups with the same-sex therapist. Miller's et al. (1995) examination of the literature revealed that adolescent female offenders need "unique consideration" when implementing treatment. They contend that female offenders have different developmental needs than males; they experience greater physical and sexual abuse than male offenders, are more vulnerable to social disapproval, depression and anxiety, experience lower levels of self-worth, may receive harsher sentencing from juvenile judicial courts, and internalize their feelings to a greater degree which may be related to addictive behaviors such as drinking and or drug use. Gutierres & Todd (1997) go on to stress the importance of understanding the prevalence of physical, emotional, and sexual abuse for female adolescents. They found that those who were abused as children had higher incidence of depression and lower self-esteem scores both at entry and at treatment completion. Mennen & Meadow (1993) also supported this finding when they contend that a history of sexual abuse in emotionally disturbed girls correlated with higher levels of depression and anxiety and lower levels of overall self-esteem. Yandow (1990), director of a center for female substance abusers in Battelboro Vermont, states that in her facility, "85 percent of all women admitted into treatment for addiction have a family history of sexual abuse, and the numbers are 100 percent for all adolescent girls admitted to treatment" (Video). Therefore, when considering the issue of abuse and its ability to be addressed in a co-ed environment, Yandow contends that many women may not feel safe expressing themselves. Yandow (1990) believes that females, especially in early

treatment, may do better in an all-female environment with treatment being directed toward those gender-specific issues relevant to the female. Therefore Yandow declares, that in order for treatment to be effective, the issue of abuse must be dealt with, otherwise, a great many relapses will be evident. In fact, Yandow claims that relapses are indicative of a treatment's negligence in addressing the issue of abuse.

Feminist theorists (Miller, 1986; Lott, 1994; Kaplan, Gleason & Klein, 1991; Surrey, 1991; Stern, 1991) discuss a bias in our society that suggests successful development involves separation, autonomy, and individualization. Whereas this model may be indicative of health in White male Western middle-class society, there is increasing evidence that this may not be true for most females and many non-Western societies including subcultures within the US. Kaplan, Gleason & Klein (1991) suggest a different paradigm for females as well as for other minority cultures in and out of Western society. They suggest a model of development based on the Self in Relation: "Within this model, women's core self-structure emerges out of experience of a relational process" (p. 123). They go on to contend that for females and people from other cultures, "Being in relationships, empathetically sharing with another, and maintaining the well-being of relationships, function as important motivators for action and sources for self-esteem and self-affirmation" (p. 123).

This theory has important implications for residential treatment. Youths in residential treatment are removed from their homes an average of 8 months to a year (Curry, 1991). Specifically understanding how female and other minority residents experience this break in the continuity of primary relationships may be paramount to understanding the behavior and symptomology that will result. This information can then

guide future treatment methodologies. For instance, if self-esteem is bolstered by the maintenance of relationships as suggested by "Self in Relation" theory, then perhaps self-esteem can be increased if treatment modalities focus on providing social support and substituting secondary relationships for the disruptions in the primary ones.

Issues of self-esteem are important indices of psychosocial well being and resiliency (Hiller, 1996; Brooks, 1994). Brooks (1994) suggests that self-esteem "is a key variable in determining resiliency" (p. 547). Brook describes self-esteem as a feeling of worth about oneself that emanates from a place of caring and respect for self and others. He goes on to describe self-esteem as a reciprocal process that "guides and motivates actions and the outcome of those actions in turn affects self-esteem" (p. 547). He concluded that youth will continue to report low levels of self-esteem when they are in situations where they do not feel successful or connected to others. Therefore, understanding how an adolescent's self-esteem may be eroded by society through their socialization processes or current situations are important considerations when determining the types of treatment/s that would be effective. Based on the idea of resiliency being related to self-esteem and a relational model of development (Kaplan, Gleason & Klein, 1991), it may be that follow-up care is more important to the female and minority adolescent as their desire to reconnect with previously interrupted relationships may be stronger than that of White males. That is, disengagement may be easier for boys as they are socialized and encouraged to separate from parents and others. In contrast, girls and many youth from minority cultures are encouraged and expected to reestablish and continually adapt themselves in such a way that maintaining relationships is of primary importance (Surrey, 1991). Burt (1993) discusses this reconnect process when

she describes her work with displaced American Indian Youth in residential treatment. She discusses the process of a 14 year-old girl who failed therapy based on treatments that emphasized separation-individuation. Burt (1993) states that we must stop imposing a "culturally and gender biased model that negates the importance of interdependency and relationships. . ." (p. 150). Burt tells us that only after this fourteen year old girl was allowed to experience and value the importance of interdependence was she able to stop being aggressive, suicidal, and self-mutilating, and eventually re-connect with her culture.

With regard to culture, mainstream ideas about identity formation in Western society focus on the idea of individuation and separation or "individualism." Hsu (1995) suggests that these Western mainstream models of identity formation do not work cross culturally. He concludes, like Burt (1993), that they are inherently culture bound because they focus on seeing and defining individuals as being separate and autonomous from others. Enns (1994) states that "American psychology has historically emphasized the centrality of the individual, independent self in identity formation" (p. 205). It is contended here that treatment modalities directly reflect this Western worldview and are therefore not necessarily effective for females and youth from other cultures. Hsu (1995) speaks about ethnic identity as something that transcends the Western concept of personality and states "the time has come to replace this concept [individualism], with something more serviceable . . . An understanding of how self is defined or operates in different cultures is the basic means of unlocking the secrets of social and cultural stability and change" (p. 24 - 25).

This discussion is based on the idea that there are critical differences associated with belonging to a collectivist or individualistic worldview. Tafarodi & Swann (1996)

describe collectivist cultures, of which Native Americans are considered to be representative, as emphasizing "social interdependence, connectedness, and mutual deference or compromise as dominant values. In contrast, highly individualist cultures are defined oppositely. They are seen as those that emphasize independence, autonomy in choice and action, and social assertiveness" (p. 651). Many times these Western ideals, considered to be representative of emotional health, are used as actual outcome variables used to measure the success of treatment! So, if you are from a collectivist society or have been socialized to value continuity of relationships, and you are informed that if you are not successful in these individualistic ideals, then you are considered sick or deviant. Under these modalities of seeing the world, it is no wonder that females and minority youth are experiencing lower levels of self-esteem. They are not represented in the theory or the treatments being implemented. Tafarodi & Swann (1996) examined how individualism vs. collectivism relates to "global self-esteem (self-liking) and more generalized self-esteem (self-competence). They concluded that collectivists have higher levels of global self-esteem (self-liking) than generalized self-esteem (self-competence). The opposite was true for the individualists. They also concluded that in order for collectivists to report higher levels of global self-esteem, a greater amount of interdependence is expected: ". . .collectivists cultural orientation, which prescribes to the needs of others and subordination of personal needs and goals to collective ones, is especially conducive to the development of self-liking" (p. 654). One must question how much relatedness and interdependence is encouraged or even tolerated in a youth residential facility where behavioral modification models are based on level systems that embrace a worldview reflecting individual autonomy and separation as positive indices of

compliance and success. One need not come from another country to experience differing ways of relating or defining self. The question begs: should relational models and emphasis on connection and interdependence also be included as possible treatment modalities and viewed as positive traits? The answer here is, yes, whenever there are females and or minority populations included. In our rush to value autonomy we have neglected to consider other models of development which may be more meaningful for diverse populations. Perhaps we should stop pathologizing "dependence" and the relational aspects of females and minority cultures and instead embrace them as valuable indices of positive health.

The above issues are very important components to consider when deciding treatment modalities for troubled adolescent females and minority youth. These issues are especially important in residential treatment as these youths have been removed from their homes as well as alienated from their culture and now receive institutional care. It is important to state that the purpose of this research project is not simply to determine differences between male and female adolescents. The purpose is to understand how and if residential treatments affect female and minority adolescents differently than they do White males. It is argued that the initial evidence presented above is significant to justify the importance of this project. The importance of providing appropriate treatment and after-care for our troubled youths should not be ignored. Established treatment modalities and subsequent outcomes must include components that are relevant to the population they serve. Otherwise, we fall victim to treating our troubled adolescent girls and minority youths with ideological positions that are designed for a White male model of health and development.

Phase I: Outcome Study

As stated above, Phase I of this project is an outcome study. Data from this study resulted from a five state collaborative effort involving the investigation of "Alternatives to Residential Treatment" conducted by the University of South Florida (Duchnowski, Johnson, Hall, & Kutash, 1993). The state of Alaska was one of the five states chosen to participate. As a result of interviewing youth for the Duchnowski et al. study, administrators at the residential treatment facility in which youth had been adjudicated for treatment became interested in understanding the effectiveness of their program and wanted to conduct their own investigation of youth after discharge from their facility. Administrators at the residential facility entered into a Memorandum of Agreement with the University of Alaska and data collection began. Data was archived and available for analysis.

It is hypothesized that there will be significant differences between females and males on both the Achenbach YSR (1991) and the Rosenberg Self-Esteem Test (1965). It is also expected that females will report significantly more problem behaviors on the YSR and experience significant lower levels of self-esteem than males. In addition, it is hypothesized that females will report significantly lower scores on the family cohesion subscale of Faces III (Olson, Porter, & Lavee, 1985) than those of males. Finally, it is expected that Alaskan Natives will report lower levels of self-esteem, greater problem behaviors, and a lower level of family cohesion than White males.

Due to the expected small effect sizes (concerns of which are discussed below), this researcher was unsure whether significant differences would be found between Caucasian females and Alaskan Natives. When analysis was conducted by ethnicity, it was expected that Native Alaskans (both male and female) would report lower levels of self-esteem and greater problem behaviors than Whites. It was also expected that higher scores on self-esteem would be associated with less problem behaviors, greater family cohesion, later age of first sexual encounter, being male, and being White.

Table 1: Expected Group Differences in Phase I

	Males	Native Americans	Females
YSR (Problems)	moderate	higher*	highest*
FACES III (Cohesion)	low	lower*	lowest*
Self-Esteem (Self-Worth)	highest	lower*	lowest*

*Asterisks indicate significant differences between groups.

Phase I: Method

Phase I: Participants

Initial participants included 32 adolescents. Forty-four percent were female (14) and 56% (18) were male. Of the 32 youth at intake, 56% were Caucasian (9 males and 9 females) and (44%) were Alaskan Native (9 males and 5 females). All participants were adjudicated into the residential facility between 1992 and 1993. Mean age at intake was

16 years. Mean age of first sexual encounter for Alaskan females was 10.5; 11.3 years for Alaskan males; 12 years for Caucasian Females; and 12.6 years of age for Caucasian males. Annual mean income for families of the participants was approximately \$18,000. This figure is well below poverty level for Alaskan standards. Overall attrition rates across the four years resulted in 27 youths (84% of the original sample) completing the 12-month follow-up. Twenty-four youths or 75% of the original sample completed the 24 and 36-month interviews. Finally, twenty-one youths (66% of the original sample) completed the 48-month follow-up. Mean age of respondents at 48 months was 20 years old. Average length of stay at the facility was approximately 10 months to one year.

Phase I: Materials

Demographic and family background information was collected at intake and included a comprehensive history of problem behaviors as well as prior services received and previous placements.

Instruments include the Achenbach Youth Self-Report (YSR; Achenbach, 1991). The YSR is a 112 item multi-axial assessment instrument which measures "the competencies and problems" of 11-18 year-old youth in a standardized format. The YSR asks youth to report the level of problem behaviors experienced in the last six months on a Likert scale with responses ranging from (0), "Not True," to (1), "Somewhat or Sometimes-True," and (2), "Often True." The instrument is designed to produce an overall total score as well as scores on externalizing and internalizing subscales of behavior. Externalized subscales assess problem behaviors associated with aggression (item numbers; 3, 7, 16, 19, 20, 21, 23, 27, 37, 57, 68, 74, 86, 87, 93, 94, 95, 97, 104) and

delinquency (items number; 26, 39, 43, 63, 67, 72, 81, 82, 90, 101, 105) . The internalized subscales include self-reports regarding anxiety/depression (item numbers; 12, 14, 18, 31, 32, 33, 34, 35, 45, 50, 52, 71, 89, 91, 103, 112) withdrawal (item numbers; 42, 65, 69, 75, 102, 103, 111) and somatic problems (item numbers; 51, 54, 56a-56g). The YSR also includes a subscale that examines sociocultural problems reported by youth such as attention (item numbers; 1, 8, 10, 13, 17, 41, 45, 61, 62) socialization (item numbers; 1, 11, 25, 38, 48, 62, 64, 111) and thought problems (9, 40, 66, 70, 83, 84, 85). This instrument was normed on 1,315 (11-18 year-old) boys and girls. Test retest correlations of .69 were demonstrated over a 6-month interval for youths referred to a community mental health center (US Department of Health and Human Services, 1993). See Appendix A for the complete Questionnaire.

FACES III, The Family Adaptability Cohesion and Expressiveness Scale (Olson, Porter, & Lavee, 1985) is also included here. FACES III is a 20-item self-administered instrument, which assesses youth perception of family functioning in terms of cohesion (disengagement/connectedness), adaptability (rigidity/flexibility), and family type (extreme/balanced). The family cohesion subscale include items 1, 3, 5, 7, 9, 11, 13, 15, 17, and 19, while the family adaptability subscale includes items 2, 4, 6, 8, 10, 12, 14, 16, 18, and 20. The "family type" scale is a composite of the cohesion and adaptability scales and provides an overall indicator of family functioning. Questions involve asking participants to state how often their family engages in certain activities. Responses are recorded on a 5-point Likert scale with 1 being "Almost Never " to 5 "Almost Always." Internal consistency for the three subscales indicated a coefficient alpha for family cohesion of .77, while family adaptability was .62. Coefficient alpha for total scores

(family type) was .68. Face validity and content validity are reported as very good.

Correlations between family members on the cohesion scale were moderate to small with $r = .41$ and $r = .25$ for adaptability (US Department of Health and Human Services, 1993).

See Appendix B for a copy of FACES III.

The Rosenberg Self-Esteem Test (Rosenberg, 1965) is also used as an outcome measure. It is a highly reliable measure in the field and consists of 10 items that utilize a 5-point Likert scale. It is used to assess adolescents' global feelings of self-worth or self-acceptance. Participants are asked to respond to questions about how they feel about themselves. Responses ranged from (1) "Strongly Agree", (2) Agree, (3) "Not Sure", (4) "Disagree", and (5) "Strongly Disagree." The statements consist of both positive and negative affirmations. Therefore, agreement does not always reflect higher levels of self-esteem for youth. Cronbach alpha ranges from .77 to .88. Test-retest reliability is reported at .85 for 28 subjects in a two-week period and .82 for 259 subjects at a one-week interval (US Department of Health and Human Services, 1993). See Appendix C for a copy of the Rosenberg Self-Esteem Test.

Facility Treatment Modalities

See Figure 1 on the following page for a logic model depicting treatment variables associated with the facility and their short and long-term expected outcome. Treatment variables are addressed in Phase II of this thesis during the meta-analysis. They are included here in Phase I to reference the types of treatment all youth were receiving while in this facility.

Figure 1: A Logic Model Consisting Of Program Components And A Brief Description Of Short-Term Objectives Associated With Each One As Well As Long Term Goals And Outcome

Phase I: Procedure

Consent forms were obtained from both the participants as well as from their parents and or guardian. Participants included the 32 youths adjudicated into a residential treatment facility in Southeast Alaska between 1992-1993. Procedural methods involved this researcher and a colleague administering initial interviews within no more than three weeks after intake. This author completed 90% of all follow-ups conducted at the 12

month contact as well as 100% of all subsequent interviews. Annual follow-up interviews were completed as close to the original intake date as possible for a period of four years after intake. Consequently, there are five time-points available for analysis for this current project. Data are analyzed with regard to final outcome as well as assessing changes at each time point.

Phase I: Statistical Analyses

Since data collection occurred at five time-points (initial, 12, 24, 36, and 48-month interviews), analyses include the use of Repeated Measure ANCOVAs on outcome measures (self-esteem and problem behaviors) by gender and ethnicity as well as by overall group. ANCOVA was chosen over ANOVA as ANCOVA allows for a reduction in the error term by controlling for variance associated with the initial scores on self-esteem and problem behaviors in the DVs. This procedure results in a more sensitive test. This is an important consideration as small effect sizes ($\eta^2 = .01$, analogous to a Cohen's d of .20, 1988) are expected in differing levels of self-esteem between Caucasian males and females, and Alaskan Native males and females. Medium effect sizes ($\eta^2 = .059$, analogous to a Cohen's d of .50, 1988) are expected on scores associated with self-reports of problem behaviors. Eta-squared was the effect size utilized for all analyses associated with Phase I. It was chosen as it accounts for the proportion of variance accounted for in the dependent variables attributed to the independent variables. Eta-squared values of .01, .059, and .138 are analogous to Cohen's d of .20, .50 and .80 indicating small, medium, and large effect sizes respectively (Cohen, 1988).

When considering the small number of subjects, especially when broken into groups, statistical power is expected to be well below acceptable levels of .80 (Cohen, 1988). Therefore, any efforts to increase test sensitivity are especially important if we are to find any meaningful differences as significant. Initial scores on the Rosenberg Self-Esteem Test and the YSR were used as covariates while all post-treatment scores regarding self-esteem and/or problem behaviors were used as repeated measure DVs. Analyses also include the use of Multiple Regression.

Phase I: Results

Participant data was archived and available for analyses. A screening of data indicated that approximately 20% of scores within each group (Caucasian Females, Caucasian males, Alaskan females, and Alaskan males) contained some missing data points across time (with the exclusion of the initial interview). Initial descriptive statistics were analyzed to find the mean score for each group at each follow-up (e. g. Alaskan females during the 12 month). All missing scores were then replaced with the original mean score specific to that group for that follow-up. The descriptive statistics were then conducted again to determine if there were irregularities in the adjusted mean scores and to document any loss of variance in the standard deviations associated with each group. Tables 2 and 3 include a comparison of original mean scores and standard deviations with those of the adjusted mean scores and standard deviations for levels of self-esteem and reported problem behaviors for each group respectively.

Table 2 and 3 below represent the adjusted mean scores and standard deviations associated with all analyses in Phase I.

Table 2: Original and Adjusted Mean Scores and Standard Deviations For the Rosenberg Self-Esteem Test

	Esteem Initial	Esteem 12 Mon	Adj. Est. 12 Mon	Esteem 24 Mon	Adj. Est. 24 Mon	Esteem 36 Mon	Adj. Est. 36 Mon	Esteem 48 Mon	Adj. Est. 48 Mon
	Act. <u>M</u>	Act. <u>M</u>	Adj. <u>M</u>	Act. <u>M</u>	Adj. <u>M</u>	Act. <u>M</u>	Adj. <u>M</u>	Act. <u>M</u>	Adj. <u>M</u>
	Act. <u>SD</u>	Act. <u>SD</u>	Adj. <u>SD</u>	Act. <u>SD</u>	Adj. <u>SD</u>	Act. <u>SD</u>	Adj. <u>SD</u>	Act. <u>SD</u>	Adj. <u>SD</u>
Caucasi	<u>M</u> = 3.62	<u>M</u> = 4.18	<u>M</u> = 4.19	<u>M</u> = 4.37	<u>M</u> = 4.38	<u>M</u> = 4.20	<u>M</u> = 4.20	<u>M</u> = 3.6	<u>M</u> = 3.6
Female	<u>SD</u> = .606	<u>SD</u> = .620	<u>SD</u> = .537	<u>SD</u> = .670	<u>SD</u> = .410	<u>SD</u> = .141	<u>SD</u> = .05	<u>SD</u> = .44	<u>SD</u> = .44
Caucasi	<u>M</u> = 4.01	<u>M</u> = 4.10	<u>M</u> = 4.10	<u>M</u> = 4.06	<u>M</u> = 4.07	<u>M</u> = 4.27	<u>M</u> = 4.27	<u>M</u> = 4.4	<u>M</u> = 4.4
Males	<u>SD</u> = .551	<u>SD</u> = .258	<u>SD</u> = .224	<u>SD</u> = .568	<u>SD</u> = .449	<u>SD</u> = .377	<u>SD</u> = .33	<u>SD</u> = .54	<u>SD</u> = .42
Alaskan	<u>M</u> = 3.28	<u>M</u> = 3.55	<u>M</u> = 3.55	<u>M</u> = 3.25	<u>M</u> = 3.25	<u>M</u> = 3.64	<u>M</u> = 3.6	<u>M</u> = 3.7	<u>M</u> = 3.7
Female	<u>SD</u> = .610	<u>SD</u> = .681	<u>SD</u> = .589	<u>SD</u> = .714	<u>SD</u> = .618	<u>SD</u> = .802	<u>SD</u> = .80	<u>SD</u> = .69	<u>SD</u> = .69
Alaskan	<u>M</u> = 3.55	<u>M</u> = 4.01	<u>M</u> = 4.01	<u>M</u> = 3.76	<u>M</u> = 3.76	<u>M</u> = 4.08	<u>M</u> = 4.08	<u>M</u> = 4.4	<u>M</u> = 4.4
Males	<u>SD</u> = .442	<u>SD</u> = .449	<u>SD</u> = .389	<u>SD</u> = .3.1	<u>SD</u> = .221	<u>SD</u> = .769	<u>SD</u> = .554	<u>SD</u> = .34	<u>SD</u> = .27

Table 3: Original and Adjusted Mean Scores and Standard Deviations for Youth Self-Report Total Scores

	YSR Initial	YSR 12 Mon	Adj. YSR 12 Mon	YSR 24 Mon	Adj. YSR 24 Mon	YSR 36 Mon	Adj. YSR 36 Mon	YSR 48 Mon	Adj. YSR 48 Mon
	Act. <u>M</u>	Act. <u>M</u>	Adj. <u>M</u>	Act. <u>M</u>	Adj. <u>M</u>	Act. <u>M</u>	Adj. <u>M</u>	Act. <u>M</u>	Adj. <u>M</u>
	Act. <u>SD</u>	Act. <u>SD</u>	Adj. <u>SD</u>	Act. <u>SD</u>	Adj. <u>SD</u>	Act. <u>SD</u>	Adj. <u>SD</u>	Act. <u>SD</u>	Adj. <u>SD</u>
Caucasi	<u>M</u> = 73.6	<u>M</u> = 52.0	<u>M</u> = 52.0	<u>M</u> = 26.5	<u>M</u> = 22.5	<u>M</u> = 25.5	<u>M</u> = 25.5	<u>M</u> = 31.0	<u>M</u> = 31.0
Females	<u>SD</u> = 19.2	<u>SD</u> = 24.2	<u>SD</u> = 20.9	<u>SD</u> = 10.5	<u>SD</u> = 6.41	<u>SD</u> = 9.2	<u>SD</u> = 3.3	<u>SD</u> = 8.2	<u>SD</u> = 4.7
Caucasi	<u>M</u> = 56.7	<u>M</u> = 45.6	<u>M</u> = 45.6	<u>M</u> = 48.3	<u>M</u> = 48.3	<u>M</u> = 38.0	<u>M</u> = 38.0	<u>M</u> = 26.8	<u>M</u> = 26.8
Males	<u>SD</u> = 22.8	<u>SD</u> = 16.0	<u>SD</u> = 13.9	<u>SD</u> = 15.8	<u>SD</u> = 12.5	<u>SD</u> = 8.8	<u>SD</u> = 6.9	<u>SD</u> = 6.2	<u>SD</u> = 3.9
Alaskan	<u>M</u> = 66.6	<u>M</u> = 60.5	<u>M</u> = 60.5	<u>M</u> = 61.0	<u>M</u> = 61.0	<u>M</u> = 59.2	<u>M</u> = 52.9	<u>M</u> = 57.7	<u>M</u> = 57.7
Female	<u>SD</u> = 31.8	<u>SD</u> = 34.5	<u>SD</u> = 29.9	<u>SD</u> = 26.4	<u>SD</u> = 22.9	<u>SD</u> = 25.1	<u>SD</u> = 25.2	<u>SD</u> = 27.8	<u>SD</u> = 27.8
Alaskan	<u>M</u> = 61.7	<u>M</u> = 47.7	<u>M</u> = 47.7	<u>M</u> = 68.2	<u>M</u> = 68.3	<u>M</u> = 60.4	<u>M</u> = 60.4	<u>M</u> = 45.6	<u>M</u> = 45.6
Males	<u>SD</u> = 20.3	<u>SD</u> = 19.8	<u>SD</u> = 17.1	<u>SD</u> = 12.5	<u>SD</u> = 8.9	<u>SD</u> = 32.7	<u>SD</u> = 23.1	<u>SD</u> = 19.3	<u>SD</u> = 19.4

Test for Collinearity

High correlations between any two variable indicates the possible presence of collinearity. Tabachnick & Fidell (1996) state that collinearity is suspected to be present when correlations between two variables is $\geq .70$. Collinearity can represent a redundancy between variables indicating that both may be measuring the same thing. By including variance associated with the redundant variable a researcher is actually inflating the error term and weakening their analysis (Tabachnick & Fidell 1996).

Correlations between the YSR and the Rosenberg Self-Esteem Test as seen in Tables 4a - 4e (introduced during the Multiple Regression analysis) indicate the possible presence of collinearity as evidenced by the correlations ranging from $-.72$ to $-.69$. Tabachnick & Fidell (1996) inform us that when testing for collinearity and/or multicollinearity it is believed to be evident when a "conditioning index is $> .30$ and at least two variance proportions [are] $> .50$ for any given root number" (p. 87). A test for collinearity here indicated that all roots have a conditioning index of $<$ than $.30$, though root six is approaching with 24.453 . In addition, only one variance proportion associated with root 6 was $> .50$ (.96). Harlow (Psy 533 Coursepack, 2000) also informs us that collinearity is a problem when both variables are on the same side of the equation. During each analyses YSR and Esteem were used as separate outcome variables and did not reside on the same side of the equation simultaneously. Based on these guidelines it is assumed that the high correlations between the YSR and Self-Esteem are not adversely effecting the analyses.

Level of Functioning: Descriptive Analysis

To insure that more than just youth self-reports were used to determine outcome, overall functioning was also considered. It was revealed that twenty-one youths completed the 48-month follow-up, which accounted for 66% of the original sample. Mean age of respondents at year 4 was 20 years old. Of those 21 youths, it was revealed that 10 youths (48%) were either living independently with friends or roommates or with their fiancés (5 and 5 respectively), three youths (14%) were living with parents, two (10%) youths lived with their spouses, two youth were in a correction facility, one youth (5%) lived alone, one youth remained in a group home, one youth was in the army, and one youth was on the run from authorities. Of the youths that were incarcerated, one was a White male (sex offender) and one an Alaskan female (convicted of assault). The youth in the group home was an Alaskan male and the youth on the run was a White male.

Academic outcome indicated that 9 (43%) of youths had graduated from high school or received their GED (5 and 4 respectively). Seven youths (33%) had completed the 11th grade; two (10%) had completed the 10th grade, 1 (5%) completed the 9th grade, and one had dropped out of school in the 8th grade. One youth was in college, one was going to be entering a Voc-Tech program (terms of release from incarceration) and one was receiving an education in the armed forces.

Employment status at year four included 14 youths (67%) working full time; two (10%) were working part time (one was part of a work release program), two were unemployed but looking for work; one was receiving public assistance; and data was

missing on one youth due to runaway status; and the other youth was unemployed due to incarceration.

Marital and family status at year four indicated that 5 youths (24%) were engaged and 2 youths (10%) were already married. Of the 21 youths contacted, 12 youths (57%) had no children, 3 youths (14%) had one child or more (one youth had 2 children). Two youths (10%) had miscarried during their second trimester; two youths had terminated their pregnancies, one was currently pregnant. There is missing data on one youth due to runaway status. Anecdotal information from family members indicates that he is unmarried and childless.

Information from CAFAS, Child and Adolescent Functional Assessment Scale (Hodges, 1990), which was an instrument utilized by the agency to capture the perception of function as perceived by the interviewer, indicated that at year four 12 (57%) of youth reported having problems with substance abuse (four experiencing mild disruption, four moderate and four severe disruption). Seven youths (33%) were rated as having no disruptions and substance abuse data was missing on 2 participants (10%). CAFAS data also indicated that the biggest concern for youth at year four was related to their economic status. The majority of youth reported not being able to meet their basic needs in terms of food, clothing and shelter.

Repeated Measures ANCOVAs and Follow-Up ANOVAs on Self-Esteem

For all initial analyses, alpha level was set a $p < .05$. When p was less than .001, this was noted as $<.001$. Otherwise, actual p -values were included for all tests. For all post-hoc analyses, significance was evaluated at $p < .01$. As hypothesized, descriptive

statistics indicated that Alaskan females and Alaskan males reported lower levels of self-esteem at and after intake than did Caucasian males and females with the exception of the 48-month. At 48 months there was virtually no difference in levels of self-esteem reported by the male respondents. See Figure 2 for a plotting of mean scores of levels of self esteem across time by group.

A 2 (Ethnicity) by 4 (levels of self-esteem at 12-month, 24-month, 36-month, and 48-month interviews) mixed repeated measure ANCOVA (with initial scores on the Rosenberg Self-Esteem used as the covariate) indicated that the levels of self-esteem reported by Caucasians participants were significantly higher than those reported by the Alaskan respondents $F(1, 29) = 5.31, p = .029$. The strength of this relationship as indicated by eta-squared was .15 indicating a large effect. Power associated with the between group tests was .60. In addition, the main effect for the repeated measure indicated significant differences in levels of self-esteem across the four time points $F(3,$

90) = 9.22, $p < .001$. Eta-squared for this effect was .24 with a power level of .99. The interaction associated with differing levels of self-esteem and ethnicity was approaching significance $F(3, 90) = 2.24, p = .08$. Low levels of power .55 may have contributed to this lack of significance as eta-squared indicated a medium effect size of .07. See Figure 3 for the plotting of mean scores of self-esteem by Ethnicity across time.

To better understand the meaning of the overall significance of self-esteem and ethnicity as it relates to within gender differences, a 2 (Ethnicity) by 4 (levels of self-esteem at 12-month, 24-month, 36-month, and 48-month interviews) mixed repeated measure ANCOVA was utilized to compare differences in mean scores between Alaskan females with Caucasian females and Alaskan males with Caucasian males. Significant between group main effect for female respondents was found $F(1, 11) = 13.75, p = .003$. The strength of this relationship as measured by eta-squared indicated a substantial effect

of .55 with power = .92. One way post-hoc ANOVAs at each time point indicated that Caucasian females reported significantly higher levels of self-esteem than Alaskan females in the 24 month $F(1, 12) = 17.05, p < .001$, and the 48 month interviews $F(1, 12) = 14.36, p = .003$. Power and effect size associated with these significant findings were eta-squared = .55 with power of .99 and eta-squared .50 with power of .94 respectively. Differences in mean scores associated with the 36-month data just missed a minimum significance level $F(1, 12) = 4.67, p = .052$. Again, low levels of power may have contributed to the lack of significance during the 36-month follow-up, as eta-squared indicated a large effect size of .25 but power was too low (.50) to detect even this substantial effect. No significant main effect was found on the repeated measure for females $F(3, 36) = 1.75, p = .17$, and no significant interaction was noted. See Figure 4 for a plotting of mean scores associated with female respondents across time.

When the same analysis was completed for the male respondents, there was no significant between group main effect. Eta-squared was $<.001$ and power was $.04$, $p = .96$. However, there was a significant effect on the repeated measure $F(3, 48) = 8.55$, $p < .001$, indicating the existence of meaningful increases in levels of self-esteem across time with both groups reporting virtually the same mean score at the 48 month follow-up. Eta-squared for this finding was $.35$ with a power of $.99$. Low levels of power may have played a role in the lack of significant findings associated with the interaction of self-esteem by ethnicity for males as eta-squared indicated the presence of a medium effect size of $.07$. Power levels to detect this effect as significant were very low at $.29$. See Figure 5 for a plotting of mean scores of levels of self-esteem across time for males.

Figure 5: Levels of Self-Esteem as a Function of Time for Males

When overall levels of self-esteem were analyzed by gender, descriptive statistics revealed that males reported higher levels of self-esteem than females in all follow-ups except at 24 months when females mean scores were slightly higher than those of males. A significant main effect on the repeated measure was present $F(3, 90) = 7.74, p < .001$. The strength of this relationship indicated a large effect size with eta-squared equaling .21 with power of .98. No significant between group main effect was found (eta-squared = .002 and power = .05), nor was there a significant interaction between sex and levels of self-esteem (eta-squared = .03 and power = .22). See Figure 6 for the plotting of mean scores associated with levels of self-esteem by gender across time.

Figure 6: Levels of Self-Esteem by Gender Across Time

Repeated Measures ANCOVAs and Follow-Up ANOVAs on YSR

In support of the 2nd hypothesis, overall descriptive statistics indicated that with the exception of the initial interview, Alaskan respondents reported higher levels of problem behaviors on the YSR than did Caucasian participants. A 2 (Ethnicity) by 4 (reports of problem behaviors at 12-month, 24-month, 36-month, 48-month interviews) mixed repeated measure ANCOVA (with initial scores on the Youth Self-Report used as the covariate) indicated that the levels of reported problem behaviors by Alaskan respondents were significantly higher than those reported by Caucasian participants: $F(1, 29) = 21.05$, $p < .001$. The strength of this relationship, as indicated by eta square, was .42 indicating a substantially large effect. Power associated with this effect was .99. In addition, the main effect for the within subject test indicated significant differences $F(3, 90) = 5.24$, $p = .002$ (eta-squared = .15 and power = .92). The interaction associated with differing levels of reported behaviors and ethnicity was also significant $F(3, 90) = 6.57$, $p < .001$. Levels of power for the interaction was .77 for detecting a large effect size of eta-squared = .29. See Figure 7 for mean scores associated with reports of problem behaviors by Ethnicity.

To better understand the interaction between self-reports of problem behaviors and ethnicity, post-hoc, One Way ANOVAs were conducted at 12 , 24 , 36 and 48-month interviews using $p < .01$. As stated above the more stringent alpha level of $p < .01$ was adopted for all post-hoc ANOVAs to control for the possibility of an increased family-wise error associated with the use of multiple ANOVAs. Findings indicated that Alaskan Natives reported significantly higher levels of problem behaviors than Caucasian respondents at 24 months: $F(1, 30) = 28.48, p < .001$; 36 months: $F(1, 30) = 23.54, p < .001$; and at the 48-month follow-up: $F(1, 30) = 7.67, p = .009$. Power and effect sizes associated with these findings included eta-squared at 24 months was .47 with power of .99; at 36-months eta squared = .42 with power of .98; and at 48 months effect size was .43 with power of .99. See Figure 8 for a plotting of mean scores associated with reports of problem behaviors for each group.

To gain a deeper understand of the reporting of problem behaviors, analyses were completed on between and within group differences. A 2 (Ethnicity) by 4 (reports of problem behaviors at 12-month, 24-month, 36-month, 48-month interviews) mixed repeated measure ANCOVA (with initial scores on the Youth Self-Report used as the covariate) were analyzed for female respondents only. Findings revealed that Alaskan females reported significantly higher levels of problem behaviors than Caucasian females $F(1, 12) = 9.62, p = .009$. The strength of this relationship as measured by eta-squared indicated a substantial effect size of .44 with power = .81. There was also a significant difference for the within subjects effect for females with $F(3, 36) = 4.42, p = .01$ (eta-squared = .26 with power = .83) as well as a significant interaction of reported problem behaviors between female respondents with $F(3, 36) = 3.78, p = .02$ (eta-squared = .24 with power = .77). See Figure 9 for a plotting of mean scores associated with reported problem behaviors by female respondents.

Figure 9: Reported Problem Behaviors as a Function of Female Status

Follow-up, post hoc, One way ANOVAs at $p < .01$ at each time point indicated that Alaskan females reported significantly higher levels of problem behaviors than Caucasian females at 24 months: $F(1, 12) = 18.98, p < .001$ (eta-squared = .75 with power = .99); 36 months: $F(1, 12) = 16.70, p = .002$ (eta-squared = .54 with power = .97); and the 48-month: $F(1, 12) = 8.98, p = .011$ eta-squared = .39 with power = .79).

As with the female respondents, descriptive statistics indicated that Alaskan males consistently reported higher levels of problem behaviors than did Caucasian participants. To investigate differences in male respondents, a 2 (Ethnicity) by 4 (reports of problem behaviors at 12-month, 24-month, 36-month, 48-higher levels of problem behaviors than did Caucasian participants. To investigate differences in male respondents a 2 (Ethnicity) by 4 (reports of problem behaviors at 12-month, 24-month, 36-month, 48-month interviews) mixed repeated measure ANCOVA (with initial scores on the Youth Self-Report used as the covariate) was conducted. A significant between group main effect for male respondents was found: $F(1, 16) = 7.61, p = .014$. The strength of relationship as measured by eta-squared indicated the presence of a large effect of .32 with power = .73. There was also a significant within group effect present: $F(3, 48) = 7.69, p < .001$. (eta-squared = .33 with power = .98). No significant interaction was evident even though a large effect size of eta-squared = .13 was present. Power levels to detect this effect as significant were very low at .59.

To better understand the main effect of ethnicity on problem behaviors for male respondents, post-hoc One Way ANOVAs at $p < .01$ were conducted at each time point. It was revealed that only at 24 months were findings significant at $p < .01$ with Alaskan males reporting significantly higher levels of problem behaviors than did Caucasian

males, $F(1, 16) = 15.18, p = .001$ (eta-squared = .67 with power = .99). At the 36-month follow-up, significant findings were found at $p < .05$ with an $F(1, 16) = 7.75, p = .013$ (eta-squared = .30 with power levels = .75). See Figure 10 for a plotting of mean scores associated with levels of problem behaviors for male respondents.

Figure 10: Levels of Problem Behaviors as a Function of Male Status

When overall levels of reported problem behaviors were analyzed by gender, descriptive statistics revealed that females reported higher levels of problem behaviors than males in the Initial, 12-month and 48-month follow-ups. During the 24 and 36-month follow-ups, males reported higher levels of problem behaviors than females. A 2

(Ethnicity) by 4 (reports of problem behaviors at 12-month, 24-month, 36-month, 48-month interviews) mixed repeated measure ANCOVA (with initial scores on the Youth Self-Report used as the covariate) revealed no significant main effect for differences between male respondents even though eta-squared indicated a medium effect size of .07. Power to detect this finding as significant was extremely low at .29. For the within subject effects, there was both a significant main effect for reported levels of problem behaviors across time $F(3, 90) = 5.38, p = .002$ and a significant interaction between sex and levels of self esteem $F(3, 90) = 7.64, p < .001$ was present. Eta-squared and power levels for these within subject effects included eta-squared = .15 with power of .92 and eta-squared = .20 with power of .98 respectively. Differences in reported problem behaviors by gender as viewed by ethnicity were discussed above and mean scores are represented in Figure 11. As discussed above, much of the variance for the interaction is associated with significantly higher reports of problem behaviors by the Alaskan population as compared to Caucasian youths. The interaction of problem behaviors as a function of gender are plotted in Figure 11 and are interpreted above when differences in reports of problem behaviors were investigated by group.

Figure 11: Self-Reports of Problem Behaviors by Gender

Phase I: Multiple Regression.

To understand the relationship between the prevalence of self-reported problem behaviors, feelings of family connectedness, early age of first sexual encounter, being White or being Alaskan Native, and being female or being male on levels of self esteem, standardized multiple regression was employed. The dependent variable was observed scores on the Rosenberg Self-Esteem Test at each time-point. Independent variables included, (1) overall total scores on the YSR; (2) scores on the Family Cohesion portion of FACES III [This instrument was only applicable for the initial and 12-month follow-ups as it was only completed if youth were living with their families in the last six months prior to that particular interview. At 24 months this number became too small for any statistical testing]; (3) age of first sexual encounter; (4) gender; and (5) ethnicity. Even though gender and ethnicity are dichotomized categorical variables, statistical precedence allows for their inclusion here. Convention states that approximately 5-10 subjects are needed for each IV included in this analysis (Tabachnick & Fidell, 1996). Since the sample size presented here is at or below the minimum level for participant to variable ratio, the data presented here are considered to be a pilot sample for this population. As stated above, while assessing overall outcome was important, it was also important to understand the process associated with efficacy. So Multiple Regression analysis was completed at each time point. Tables 4a through 4e include correlations, unstandardized regression coefficients (B), standardized regression coefficients (Beta), semi partial correlations (sr^2), R^2 , Adjusted R^2 , power, and unique and shared variance contributed to the dependent variable by the independent variables in the equation for each contact.

At the initial interview, R for the regression was significantly different from zero $F(5, 26) = 4.25, p = .006$. Overall, the independent variables accounted for 45% of the variance in the dependent variable, self-esteem. Only ethnicity ($sr^2 = .17$) and initial scores on the YSR ($sr^2 = .16$) contributed significantly to the prediction of self-esteem. Other sources of variability included $sr^2 = .02$ (gender), $sr^2 = .04$ (family cohesion), and $sr^2 = .01$ (age of first sexual contact). Shared variability for all independent variables accounted for 5% of the variance. Findings indicate that a negative relationship existed among ethnicity and self-esteem. This indicates that there is a significant relationship between being an Alaskan Native and reporting low levels of self-esteem. There was also a significant negative relationship between reporting of problem behaviors and levels of self-esteem indicating lower levels of self-esteem are associated with higher reports of problem behaviors. See Table 4a below for all findings associated with the Initial interview.

Table 4a: Standard Multiple Regression Of Problem Behaviors, Early Sexual Contact, Family Cohesion, Gender, And Ethnicity On Self-Esteem At Initial Interview

Variables	Esteem Initial	YSR Initial	Ethnicity	Family Cohesion	Gender	B	Beta	sr^2 (unique)
YSR Initial	-.51					-.0109	-.4286*	.161
Ethnicity	-.31	-.04				-.4850	-.4280*	.165
Fam Cohes	.30	-.32	.22			.1064	.2071	.04
Gender	.25	-.27	.14	.30		.1655	.1464	.02
Age 1st Sex	.03	-.04	-.24	.03	.11	-.0319	-.1115	.01

$$\overline{R^2} = .45$$

$$\text{Adjusted } R^2 = .34$$

$$R = .67^{**}$$

$p < .05$

Unigue variability = .40; Shared variability = .05; Female = 1 & Male = 2; Caucasian = 1 & Alaskan Native = 2

Effect Size Power

.01 .07

.059 .23

.138 .52

df(5,26)

The regression during the 12-month follow-up was also significantly different from zero with $F(5, 26) = 10.75, p < .001$. During the 12-month follow-up the percentage of variance accounted for was 67%. Again, only 2 variables contributed significantly to the prediction of self-esteem; age at first sexual contact ($sr^2 = .11$) and scores on the YSR at 12 months ($sr^2 = .52$). Other non-significant contributors included $sr^2 = .001$ (Gender), $sr^2 = .009$ (Ethnicity), and $sr^2 = .01$ (Family Cohesion). Shared variability for all independent variables accounted for another 2% of the variance. Findings indicate that higher ages at time of 1st sexual contact are associated with higher reports in self-esteem. The significant negative relationship between reporting of problem behaviors and levels of self-esteem, indicated that lower levels of self-esteem are associated with higher reports of problem behaviors. See Table 4b for all findings associated with the 12-month follow-up.

Table 4b: Standard Multiple Regression Of Problem Behaviors, Early Sexual Contact, Family Cohesion, Gender, And Ethnicity On Self-Esteem At 12 Months

Variables	Esteem 12 Mo.	YSR 12 Mo	Ethnicity	Fam Cohes 12 Mo	Gender	B	Beta	sr^2 (unique)
YSR 12 Mo	-.70					-.0183	-.7686**	.517
Ethnicity	-.32	.09				-.1051	-.1138	.009
Fam Cohes	-.18	.13	.30			.1064	-.1115	.010
Gender	.10	-.22	.14	-.10		-.0885	-.0958	.001
Age 1st Sex	.23	.18	-.24	.17	.11	.0881	.3769**	.114

$$\begin{aligned} R^2 &= .67 \\ \text{Adjusted } R^2 &= .61 \\ R &= .82*** \end{aligned}$$

$p < .05$

Unigue variability = .65; Shared variability = .02; Female = 1 & Male = 2; Caucasian = 1 & Alaskan Native = 2

Effect Size	Power
.01	.07
.059	.23
.138	.52
df(5, 26)	

The analysis at 24 months revealed that the regression was significant $F(4, 27) = 5.68, p = .002$. As stated above, it was at this time point that the Family Cohesion scale was dropped from the analysis due to only several respondents completing this instrument after the 12-month follow-up. Only reports of problem behavior at 24 months was found to be a significant contributor to self-esteem with $sr^2 = .11$ (negative correlation). Overall R^2 accounted for 45% of the variance in the dependent variable. Other sources of variability included Gender, $sr^2 = .041$ (negative correlation); Ethnicity, $sr^2 = .028$ (negative correlation); and Age at First Sexual Encounter ($sr^2 = .001$). Shared variability among all independent variables accounted for 28% of the variance. See Table 4c for all findings associated with the 24-month follow-up.

Table 4c: Standard Multiple Regression Of Problem Behaviors, Early Sexual Contact, Gender, And Ethnicity On Self-Esteem At 24 Months

Variables	Esteem 24 Month (r DV)	YSR 24 Month	Ethnicity	B	Beta	sr^2 (unique)
YSR 24 Month	-.60			-.0146	-.5408*	.106
Ethnicity	-.59	.70		-.2793	-.2553	.027
Age 1st Sex	.09	-.09	-.24	-.0098	-.0354	.001
Gender	-.07	.48	.14	.2674	.2444	.041
						$R^2 = .46$
						Adjusted $R^2 = .38$
						$R = .68^{**}$

$p < .05$

Unigue variability = .17; Shared variability = .28; Female = 1 & Male = 2; Caucasian = 1 & Alas. = 2

Effect Size Power

.01 .07

.059 .27

.138 .60

df(4, 27)

During the 36-month interview, the regression was also significant $F(4, 27) = 11.22, p < .001$. Overall R^2 accounted for 62% of the variance in the dependent variable. Reports of problem behaviors (YSR) continued to remain a significant contributor $sr^2 = .45$ (negative correlation). Gender (negative correlation) was also found to be significantly associated with self-esteem $sr^2 = .15$. Other non-significant contributors included Age at First Sexual Encounter ($sr^2 = .001$) and Ethnicity (negative correlation) ($sr^2 = .03$). No amount of shared variance was revealed, as the sum total of sr^2 was approximately .63 accounting for total R^2 . The significance of gender indicates that during the 36 month, higher levels of self-esteem are significantly associated with being male. See Table 4d for all findings associated with the 36-month follow-up.

Table 4d: Standard Multiple Regression Of Problem Behaviors, Early Sexual Contact, Gender, And Ethnicity On Self-Esteem At 36 Months

Variables	Esteem 36 Month (r DV)	YSR 36 Month	Ethnicity	B	Beta	sr^2 (unique)
YSR 36 Month	-.69			-.0213	-.9439**	.454
Ethnicity	-.33	.66		.2219	.2315	.029
Age 1st Sex	.17	-.24	-.24	-.0092	-.0380	.001
Gender	.18	.27	.14	.2674	.4128**	.152
						$R^2 = .62$
						Adjusted $R^2 = .57$
						$R = .79***$

$p < .05$

Unique variability = .63; Shared variability = 0; Female = 1 & Male = 2; Caucasian = 1 & Alaskan Native = 2

Effect Size	Power
.01	.07
.059	.27
.138	.60
$df(4, 27)$	

The analysis at 48 months resulted in the regression being significantly different from zero $F(4, 27) = 8.15, p < .001$. Only reports of problem behaviors was found to be a significant contributor to the regression equation with $sr^2 = .38$ (negative correlation). Overall R^2 accounted for 55% of the variance in the dependent variable. Additional sources of variance included Gender ($sr^2 = .019$), Age at First Sexual Encounter ($sr^2 = .006$), and Ethnicity ($sr^2 = .004$). Shared variability for all independent variables accounted for approximately 14% of the variance. See Table 4e for all findings associated with this analysis.

Table 4e: Standard Multiple Regression Of Problem Behaviors, Early Sexual Contact, Gender, And Ethnicity On Self-Esteem At 48 Months

Variables	Esteem 48 Month (r DV)	YSR 48 Month	Ethnicity	B	Beta	sr^2 (unique)
YSR 48 Month	-.72			.0182	-.6930**	.376
Ethnicity	-.34	.45		-.0645	-.0702	.004
Age 1st Sex	.02	-.10	-.24	-.0181	-.0779	.006
Gender	.16	-.05	.14	.1327	.1444	.019
Adjusted						$R^2 = .55$
						$R^2 = .48$
						$R = .74***$

$p < .05$

Unigue variability = .41; Shared variability = .14; Female = 1 & Male = 2; Caucasian = 1 & Alaskan Native = 2

Effect Size	Power
.01	.07
.059	.27
.138	.60

Phase II: Meta-Analysis

There is evidence in our field that expanded methodologies are needed if we are to capture the complexity of reality. These expanded methods, which allow for an

understanding that embraces relationships and contextual knowledge, include the use of rigorous assessment of the nature of our findings. There is also a call in our field to supply our science with multiple indices of assessing reality. For instance, Harlow, Mulaik, & Steiger, (1997) suggest that dichotomized hypothesis testing is not sufficient, nor, by its very nature, can it capture the complexity of reality. They go on to emphasize the need to include multiple ways of evaluating scientific models in addition to significance tests. At the heart of this debate is the issue of including effect sizes and power analyses within our research designs as well as when we disseminate our results. Rossi (1990) and Maddock (1999) contend that the reporting of power in most psychological journals is either non-existent or grossly underreported. They posit that incorporating and reporting effect size and establishing levels of power in our research will allow researchers to design their study with appropriate levels of power. The inclusion of these techniques will enable researchers to find meaningful differences when, in fact, they do exist.

Incorporating these methodological practices can move our science toward a more comprehensive and cumulative understanding of that which we choose to study (Rossi, in Harlow, et al., 1997; Lipsey, 1990). In addition to the use of multiple indices of effect suggested above, the use of meta-analyses is also believed to provide considerable contributions to our field in terms of establishing a comprehensive understanding of an area of study (Cook et al., 1994; Cooper & Hedges, 1994; Hunt, 1997; Rosenthal, 1991). For example, Hunt (1997) contends that meta-analysis has taken its place in contemporary research methods. He states:

meta-analysis is a means of combining the numerical results of studies with disparate, even conflicting, research methods and findings; it enables researchers to discover the consistencies in a set of seemingly inconsistent findings and to arrive at conclusions more accurate and credible than those presented in any one of the primary studies (p. 1).

Harlow (1997) succinctly explains that "...meta-analysis allows us to examine a set of findings for a phenomenon, gathered from a number of studies, and synthesize the overriding effects" (p. 9). Cooper & Hedges (1994) inform us that meta-analyses allow for the synthesis or integration of information within a field of study. To summarize, meta-analysis is not simply an exhaustive search of the literature, but a scientific and systematic way of combining outcomes so that a deeper, and perhaps, "truer" understanding is possible. This methodology is especially important in a field where controversial findings are present such as those findings associated with residential treatment (Wells, 1991; Asarnow, Aoki, & Elson, 1996; Curry, 1991). No one study can create change or establish "truth." Multiple studies must be investigated before any "true" conclusions in a field of study can be made. Meta-analysis allows us to incorporate multiple indexes of effect and examine methodological practices in a field as well as defining those variables that may be the most important predictors of outcome.

As stated above, it was expected that the effect size within this outcome study would be small (Cohen's d of .20). However, average effect size associated with this outcome study was $\eta^2 = .18$ which corresponds with a Cohen's d (1988) estimate of a large effect of .80. Average power associated with the outcome study was still inadequate even to detect this large effect at .60. Achieving the greater effect size is

believed to be related to the alarming differences in scores associated with the poorer outcome for Alaskan Females with respect to the other groups. Also, the use of the ANCOVA, which allowed for the reduction in the error term coupled with the fact that this was a longitudinal design spanning four years and including 5 time points for analysis with an attrition rate less than 34%, may have all contributed to the higher than expected effect size. Therefore, in order to understand if this effect size is typical in the field, Phase II of this project includes the completion of a meta-analysis. This practice allows for a methodological assessment of effect sizes and power levels in current published research. This information can then aid future studies to be designed with the adequate power necessary to detect these differences as statistically meaningful when they are present. That is, a meta-analysis allows for the synthesis of research findings in this area as well as providing information on the quality of such research with respect to effect size and power as suggested by those authors cited above.

In addition to an exhaustive review of the research, a thorough understanding of the treatment modalities, outcome variables, and philosophies of particular facilities under study is possible. As stated previously, Wells (1991) suggests that one of the problems associated with the evaluation of residential care is that there is a lack of consistency between agencies. She states, "There is no definitive form of residential treatment. . . practice, policy, and research." A meta-analysis helps to answer the call put forth by Wells and others and provide a comprehensive understanding of whether treatment modalities are effective for all populations receiving treatment; including female and minority adolescents. Based on the fact that research in this area is dependent on intake samples resulting in low sample sizes utilizing varying outcome measures and indices of

positive outcome, one outcome study within this area of investigation is very problematic for generalizability purposes. However, when incorporating a programmatic approach to research by completing a meta-analysis, individual outcome studies provide the data for a more comprehensive understanding.

Phase II: Procedure

Meta-analyses often involve the most recent work in a field of investigation. This meta-analysis focused on 16 different psychological journals relating to the identification and treatment of problem youth published between 1994 -1999. Specifically, the search focused on retrieving information related to youth outcome associated with those receiving residential treatment. The rationale for going back 6 years was based on the fact that the average stay in a residential treatment facility is 8 months to 1 year (Curry, 1991). Outcome studies usually include 6-months to one-year follow-ups after discharge. Given the time it takes to complete the follow-ups, do the analysis, and publish the results, going back six years seems like a reasonable time period.

Phase II: Initial Inclusion Criteria

- 1) Articles must include information related to outcome for female and minority adolescents in residential treatment.
- 2) Youth must be between the ages of 8 and 20 at the time they were receiving residential care.

- 3) Articles must contain the necessary information to calculate effect size including means, standard deviations, number of participants and/or F , t , or p -values.
- 4) Articles must include types of treatments and outcome measures utilized, as well as program goals.
- 5) Articles must express theoretical positions utilized at the facility regarding the treatment of adolescents.

Note: The inclusion criteria for number 1 listed above was modified as very few articles included outcome associated with females and/or minority adolescents. Therefore, any outcome study relating to residential treatment in general was gathered. It was this researcher's job to then try and extract appropriate information relating to minority and female adolescent outcome.

Phase II: Analyses

- 1) Assess methodological quality
- 2) Establish overall effect size and power
- 3) Identify potential mediator and or moderator variables associated with positive outcomes
- 4) Examine theoretical consensus in the field

It is expected that females and minority adolescents are not benefiting from traditional behavioral approaches used in residential care designed to reduce delinquent behaviors. This finding was supported in Phase I of this project. It is also speculated

that delinquent behaviors such as minor consuming of alcohol and/or drugs, traditionally characterized as externalizing behaviors by many theorists, may be attributable to internalizing mechanisms associated with somatic distress resulting from sexual and/or physical abuse. It may be revealed that drinking and drug use relate to both externalizing and internalizing subscales measuring problem behaviors associated with adolescents. If treatment modalities are designed to treat these types of problems as only externalizing behaviors, then there may be little hope of understanding the internalized latent construct that is driving these behaviors. Specifically, these constructs could include feelings of low self-esteem and or feelings of disconnectedness from family and community experienced by female and minority adolescents in treatment. Though it is unclear if this information will be confirmed in the journals included for this analysis.

It is also expected that few articles, if any, will report actual levels of effect size or power for their individual study. Additionally, soundness of methodological reporting of statistical information related to findings is also expected to be low resulting in few articles including the necessary information to conduct a power or effect size analysis. It is also expected that few studies will statistically control for baseline data. This is especially important when random assignment is not possible. That is, due to the fact that this is an intake sample, random assignment of participants to groups is not expected, so the use of comparison or control groups as well as using covariates to control for baseline data is very important.

Phase II: Retrieving Articles

Psychological data bases such as psycINFO and psycLIT were searched utilizing cross referenced key-words including residential treatment, treatment outcome, minority and female juvenile delinquent, follow-up, longitudinal, Native American youth, Indian youth, gender and juvenile offenders, court referred youth, emotionally disturbed youth, and at risk youth. This method resulted in many articles involving problem youth in general, but few actual outcome studies related to residential treatment were retrieved. Examination of the references associated with those articles that did relate to juvenile delinquent outcome in residential treatment revealed journal publications that were most frequently publishing work related to the topic of residential treatment for troubled adolescents.

Sixteen journals publishing literature between 1994 - 1999, then became the focus of the search. They included the *Journal of Emotional Behavioral Disorders*, *Journal of Cross Cultural Psychology*, *Journal of Adolescent Research*, *Journal of Child Welfare League of America*, *Children and Youth Services Review*, *Adolescence*, *Journal of Clinical Psychology*, *Residential Treatment for Child and Youth*, *Cultural Diversity and Mental Health*, *Journal of Child Psychology and Psychiatry*, *American Journal of Orthopsychiatry*, *Child and Youth Care Reform*, *The Canadian Journal of Nursing Research*, *Canadian Journal of Nursing Science*, *Journal of Abnormal Psychology*, and finally the *Journal of Research in Crime and Delinquency*. The Canadian Journals were included in hopes of finding outcome associated with Native American groups similar to those found in Alaska.

Phase II: Method

Approximately 3,600 different citations were included in the 16 chosen Psychological and Nursing Journals published between 1994 and 1999. These titles were examined with an expanded version of key words associated with juvenile delinquent behavior and/or treatment outcome for troubled adolescents. Approximately 20% of all titles contained relative key words resulting in the examination of some 700 abstracts. Of these 700 abstracts, 12% or 86 articles involved residential treatment for juvenile delinquents between the ages specified above. These 86 articles included several different aspects or implications of youth receiving residential treatment, and they were gathered to use in this meta-analysis. Twenty-six (30%) of the 86 articles included here involved the actual examination of outcome, at some point in time, for youth in residential treatment. These 26 articles became the focus of the methodological assessment such as establishing levels of effect size and power and looking at overall efficacy. Outcome articles were also evaluated on how well they described their sample as well as how well they included relevant information associated with their findings, such as including the type of test conducted, mean-scores, standard deviations, F , t , R^2 and /or t -values, degrees of freedom, actual and/or apriori p -levels, and non-significant findings.

In order for an article to be included in the actual synthesis of information, the research design of the original outcome study needed to either include a pre/post-test, a comparison group, a control group, or a longitudinal design as well as an overall N and/or sample size for each group. In order to transform t -scores and/or F -scores into standardized Z -scores, articles must have included the actual t or F -values and/or p -level).

All resulting $\underline{Z}$ scores were associated with one-tailed p -values. The procedures for which this meta-analysis is based rely heavily on the recommendations put forth by Rosenthal (1991) and Cooper & Hedges (1994). Formulas used to transform $\underline{t}$, $\underline{F}$, $\underline{r}$, to standardized $\underline{z}$ -scores included the following recommendations by Rosenthal (1991) and Cooper & Hedges (1994):

Transforming $\underline{t}$ -scores to $\underline{Z}$ -scores

$$(\underline{t}) / \text{SQR} [\text{df} / (\text{df}-2)] \quad (1)$$

Where t is the actual $\underline{t}$ -value and df are the degrees of freedom associated with each $\underline{t}$ -test.

$\underline{F}$ -values with $df=1$ in the numerator were transformed to $\underline{t}$ -values by taking the square root of the $\underline{F}$ -value and then applying equation number 1 where:

$$\underline{t}^2 = \underline{F} \text{ or } \underline{t} = \text{the SQR of } \underline{F} \quad (2)$$

Pearson's $\underline{r}$ s were transformed into $\underline{z}_r$ values through the formula:

$$\underline{z}_r = \ln((1 + \underline{r}) / (1 - \underline{r})) / 2 \quad (3)$$

Chi-Squared values, $\underline{X}^2$, were transformed to $\underline{z}$ -scores based on the following calculations:

$$\underline{Z} = \text{SQR} (\underline{X}^2 (1)) \quad (4)$$

Where $\underline{X}^2 (1) = (O - E)^2 / E$, and O = observed frequencies, and E = expected frequencies

Summing of $\underline{Z}$ -scores was accomplished utilizing the following procedure:

$$\underline{Z} = \underline{Z} / \text{SQR}(N) \quad (5)$$

Where $\underline{Z}$ are the $\underline{Z}$ -scores obtained through the transformations listed above, and N is the number of studies involved in the analysis (one overall $\underline{Z}$ -score was obtained for each article).

For those tests that were found to be non-significant and no *p*-values or actual test values were given, no effects were assigned. That is, *p*-values were set at .50 and *Z*-scores were set at 0. When actual *t* or *F*-values were missing, but *p*-values were indicated, then the minimum *t* or *F*-value associated with that *p*-value was assigned (Rosenthal, 1991).

Phase II: Results

Phase II: Frequencies and Demographics

Twelve (46%) of the 26 outcome articles included a designed which met the criteria for the actual meta-analysis (Ansari, et al., 1996; Day & Pal, 1994; MacNeil, 1998; Howe et al., 1998; Simmerman, 1997; Pugh, et al., 1997; Tucker et al., 1997; Burks, 1995; Kapp, 1994; Wright, 1995; Carlson, et al., 1994; and Deschenes & Greenwood, 1998). Of the remaining 14 articles, 4 (15%) were qualitative (Shennum, 1995; Miceli, 1997; Pazarat, 1999; Penzerro, 1995). Six (23%) of the articles indicated testing on frequencies and/or descriptive statistics with and without comparison or pre/post groups, but included no standard deviation scores or actual *t*, *F*, or *p*-values (Zimmerman, 1994; Stage, 1999; Daly, et al., 1998; Forand, 1999; Knorth & Elderling, 1998; Brown & Hill, 1996). Two of the articles (8%) utilized multivariate techniques such as hierarchical linear modeling (Friman et al., 1995; Smith, et al., 1996), while 1 (4%) involved scale development/measurement (Rafel, 1994), and 1 (4%) gave no indication of the type of tests utilized (Larson & Calamari 1998). Very few articles indicated the average length of stay and attrition rates associated with follow-up interviews. Also, few articles stated where youth were discharged once they were released from residential care. So these descriptives could not be included here.

As stated above, the purpose of this meta-analysis was to investigate the efficacy of residential treatment, specifically for females and minority residents. While a number of the 60 retrieved theoretical articles discussed issues associated with gender and delinquency as well as expressing concern regarding the frequency of minority offenders, few researchers addressed these issues in the design or interpretation of their research. This represents a lack of sensitivity on the part of the researchers. That is, the issues and concerns are present in the theoretical literature, yet they are not being investigated in the actual research. For instance, 22 (85%) of the 26 overall outcome articles did define their population by gender while only 14 (54%) defined their sample by ethnicity at least to the extent of including the number of non-White youth. However, of the entire 26 outcome articles, 8 (31%) considered outcome with regard to females while only 5 (19%) considered outcome for the minority portion of their sample. These numbers shrink when we examine those 12 articles that met the reporting criteria to be included in the meta-analysis.

As previously stated, only twelve (46%) of the 26 outcome studies included sufficient information so that power and Z -scores could be calculated. Of those 12 articles included in the actual statistical analysis, 6 (50%) considered female outcome and only 2 (17%) included outcome associated with minority youth. Of those that that did report results, positive outcomes were reported for both minority and female adolescents, or it was simply stated that no "significant" differences were found. This alarmingly small number of studies and lack of statistical information specifically for females and/or minorities made it unrealistic to complete a meta-analysis on only those 6 articles that included gender and those 2 that considered ethnicity. While this lack of reporting of

outcome associated with females and minority youth was expected, the degree to which it is present is somewhat astounding in light of the fact of the frequency with which these issues are appearing in the theoretical literature. So, unfortunately, this meta-analysis focused on outcome as a whole with this researcher turning over any available statistical nugget to provide clues as to what's going on for these populations. It is also interesting to note that placement after treatment was a particularly important theoretical outcome for these groups, yet this was also neglected in the reporting. In the 12 studies included for the statistical analysis only 1 included discharge placement, and only 3 (12%) of the entire 26 articles associated with outcome did so.

The total N included in the 12 articles used for the statistical analysis was 1,537 youth making an average of 128 participants per study. It should be noted that three of the studies involved federal or state funding and participants in these studies accounted for 1,001 participants. This leaves only 536 participants distributed among the other nine studies; indicating a more realistic average N of approximately 60 subjects per study. Even this relatively small average is considered to represent a large sample size for research involving intake samples of juvenile delinquent into residential treatment (Lipsey, In Cook et al., 1994). Ten articles, (83%), reported the age of their participants which was calculated to be 12.6 years old. (When only ranges of age were given, the median value of this range was used). Of the twelve articles, 10 (83%) identified their participants by gender. Within these 10 articles, there were a total of 1,238 participants with females accounting for 19% (230) of the total N. With regard to ethnicity, only 5 (42%) of the 12 studies identified their sample. There was a total of 829 participants in

these 5 studies with ethnic minorities representing 406 (49%) of all participants. See Table 5 for a complete description of the 12 articles included in the meta-analysis.

Table 5: Article Demographics, Gender, Ethnicity and Age								
Author	Year	Design	Total N	Females	Males	Caucasian	Ethnicity	Age
*Day & Pal	1994	Longit.	37	6	31	N. G.	N.G.	8.5
MacNeil	1998	Comp Group	74	32	42	N. G.	N.G.	15.2
*Howe et al.	1998	Comp Group	84	34	50	49	35	8
Simmerman	1997	Pre-Post	107	N. G.	N. G.	N. G.	N. G.	15.5
*Pugh et al.	1997	Pre-Post	246	87	159	N. G.	N. G.	8.62
*Tucker et al.	1995	Comp Group	23	23	0	11	14	15.5
*Burks	1995	Pre-Post	37	24	13	25	12	N. G.
Kapp	1994	Comp Group	563	0	563	253	310	N. G.
Wright	1995	Pre-Post	30	0	30	N. G.	N. G.	12.6
Carlson et al.	1994	Pre-Post	84	0	84	85	35	15.2
Deschenes & Greenwood	1998	Comp Group	192	N. G.	N. G.	N. G.	N. G.	13
*Ansari	1996	Pre-Post	60	24	36	N. G.	N. G.	16.5
Totals			1,537 n=12	230 (19%) n=10	1,008 (81%) n=10	423 (51%) n=5	406 (49%) n=5	12.86 n=10

N. G. = Not Given, *Includes Outcome for Females

Effect Size Analysis

As expected, no direct reports of effect size were given for any of the articles included here. It is important to point out that in several of the articles the authors expounded on the idea of the importance of effect size and limitations associated with the low number of participants, however none reported their own levels of effect size or power, and 75% of them (9) did not include enough information so that effect size could be calculated! In order to calculate the effect size of an article it must have two groups with means and standard deviations or one group with a know population mean and standard deviation as well as their associated df. Unfortunately, only 3 articles (25%) containing sixteen tests included the relevant information to calculate effect size (Howe et al., 1998; Pugh et al., 1997; and Tucker et al., 1995). All tests within any one article were

averaged to produce an overall effect for that article. Effect sizes were then combined to achieve an overall average. Effect sizes are based on Cohen's d (1988) estimates for small, medium, and large effect sizes of .20, .50 and .80 respectively. All effect sizes were calculated from computer programs developed by Dr. Rossi at the University of Rhode Island. These programs are based on the "cube root normal approximation of the noncentral F distribution (Rossi, In Harlow et al., 1997). All analysis are based on a two-tailed alpha level of .05 and all effect-sizes are based on the Cohen's d (1988) for calculating effects:

$$d = M_1 - M_2 / s \quad (6)$$

Where M_1 is the mean of the first group, and M_2 is the mean of the second group, and s is the pooled standard deviation for both groups. To correct for possible bias in effect size due to sample size Hedges's (1985) unbiased effect size estimator g^u was used on each individual effect size. Hedges's g^u , as presented by Rosenthal (1991), was estimated by:

$$g^u = c(m)g \quad (7)$$

$$\text{Where} \quad c(m) = 1 - (3/4(m) - 1) \quad (8)$$

and g^u like Cohen's $d = M_1 - M_2 / S$ and S is the pooled standard deviation for the two groups.

Each article contributed one overall effect size. When averaged, the overall effect size for the 3 articles was .26 indicating a relatively small effect. Average for Hedges's g^u was just slightly lower at .25. Ranges for effect included $d = .85$ to .06. Individual effects for each article were .34 (Tucker et al., 1995, 12 tests, $n=23$); .29 (Howe et al., 1998, 2 tests, $n=84$); .14 (Pugh et al., 1997, 2 tests, $n=246$). These findings suggest heterogeneity of effect within the small effect size range. However, no one article indicated anything

beyond a small effect size based on Cohen's (1988) estimates. In a meta-analytic inquiry into the effects of juvenile treatment, Lipsey (In Cook et al., 1994) found that the "grand mean over all the studies in their analysis [Andrew et al., 1990] was an effect size of .21 standard deviation units [Cohen's estimate], quite comparable to those found in virtually all the previous meta-analyses" (p. 87). Therefore, it is concluded that small to small-medium effect sizes are expected within this area of study. However, due to the small sample included here (n=12), the results are nonetheless interpreted with caution. The effects in the field are substantially smaller than those found in the present outcome study. Reasons for this have been discussed and it's believed that they reflect the much poorer outcome for Alaskan females. See Table 6 below for the design of the article, t-test values, df, effect size, Hedges's g^u and associated Z scores with each test as well as an average for each article and overall effect size.

Table 6: Summary of Effect Size Analysis

Author	t-value	df	effect size d	Hedges's g	z
Tucker et al., (1995)					
assertive-pre	0.2954	1, 21	0.1233	0.1118	0.2673
assertive post	1.0351	1, 21	0.4319	0.4162	0.9364
self-esteem-pre	0.4389	1, 21	0.1832	0.1765	0.3979
self-esteem-post	0.5079	1, 21	0.2121	0.2044	0.4595
success behav-pre	0.5529	1, 20	0.2357	0.2267	0.4976
success behav-post	0.2326	1, 20	0.0992	0.0954	0.2093
assertive pre-post tx	1.4671	1, 10	0.8469	0.7821	1.1736
assertive pre-post cont	0.6429	1, 9	0.3893	0.3359	0.5000
self-esteem pre-post tx	0.9217	1, 10	0.5322	0.4915	0.7374
self-esteem pre-post cont	0.6063	1, 9	0.3672	0.3357	0.4716
suc-behav pre-post tx	0.2613	1, 9	0.1582	0.1446	0.2032
suc-behav pre-post cont	0.9051	1, 9	0.5481	0.5011	0.7039
Mean Total			0.34	0.32	0.5000
Howe et al., (1998)					
tx and comp group	0.2789	1, 82	0.0609	0.0603	0.2721
	2.3381	1, 82	0.5102	0.5055	2.2811
Mean Total			0.29	0.28	1.2766
Pugh et al., (1997)					
pre and post	1.1350	1, 244	0.1447	0.1442	1.1257
	1.0990	1, 244	0.1401	0.1438	1.0897
Mean Total			.14	.14	1.1077
Overall Effect Size and Z-score			.26	.25	0.9631

Power Analysis

Once again as expected, no one article reported the level of power within their design though several authors indicated that it could be low. The power analysis was calculated using Cohen's (1988) definitions for small (.20), medium (.50), and large (.80) effect sizes. Acceptable levels of power as specified by Cohen is .80 regardless of the effect size examined. All power analyses were completed through the use of computer programs developed by Dr. Joseph Rossi at the University of Rhode Island. These power programs determine the power of a t -test based on the normal approximation to the noncentral t distribution as established by Cohen, 1988. Power for the F statistic was based on the cube root approximation of the noncentral F distribution as specified by Laubscher, 1960 (In Rossi, 1990). A two-tailed alpha level of .05 was applied to all analyses.

Many of the articles included multiple tests. Only those tests that were central to the investigation were included. This excluded any post-hoc tests and or secondary analysis that were subsequently conducted on unexpected results as well as any scale development and/or multivariate techniques for which power is not appropriate. Several articles included Pearson's r as their primary tests, therefore they were included in this analysis. Similar tests in one article with varying degrees of freedom for each test were averaged together to get an overall level of power for each effect size. For example, if an article contained three t -tests, and each of these t -tests had differing degrees of freedom, a separate analysis was conducted on each test. Then the three separate results were averaged together to get an overall level of power for that article associated with small, medium, and large effect sizes. These would then be averaged in with other tests that may

have been completed for that article resulting in one overall level of power for each effect size for that article. For this reason the article is considered to be the unit of analysis (Rossi, 1990).

The power analysis included 306 separate t-tests, ANOVAs, Multiple Regressions, Pearson's r, and Chi-Square tests found in the 12 articles for an average of approximately 25 tests per article. This seems like an exorbitant amount of tests associated with only 12 articles and may represent the lack of methodological sophistication associated with this area of research often criticized by many. It should be noted that one article contributed 152 separate tests of significance. Therefore, the average number of tests per article is inflated drastically as a result of the inclusion of this article. Without this inflated contribution, average number of tests for the remaining 11 articles would be approximately 14. However, it is important to remember that all tests in individual articles were weighted according to the df associated with each test, so that individual contributions of effects were possible. See Table 7 below for the total N associated with each article, number and types of tests and overall levels of power as well as a levels of power for small, medium and large effect sizes as established by Cohen's d of .20, .50, and .80.

Table 7: Articles Included in Power Survey of Meta-Analysis

Author	Year	# of tests	Total N	t-tests	ANOVAs	MR	r	X ²	Power Sm.	Power Med.	Power Lg.	# of non-significant
Day & Pal	1994	32	37	28	4	0	0	0	0.09	0.29	0.61	4
MacNeil	1998	4	74	0	4	0	0	0	0.23	0.94	0.99	0
Howe et al.	1998	16	84	0	4	0	12	0	0.12	0.53	0.94	5
Simmerman	1997	2	107	2	0	0	0	0	0.54	0.99	0.99	0
Pugh et al.	1997	57	246	42	0	3	12	0	0.6	0.99	0.99	22
Tucker et al.	1995	3	23	3	0	0	0	0	0.07	0.21	0.45	2
Burks	1995	3	37	0	0	0	0	3	0.08	0.3	0.64	2
Kapp	1994	2	563	0	0	0	0	2	0.5	0.99	0.99	0
Wright	1995	1	30	1	0	0	0	0	0.18	0.74	0.99	0
Carlson et al.	1994	31	84	1	0	0	27	3	0.21	0.65	0.88	3
Deschenes & Greenwood	1998	152	192	149	0	0	0	3	0.21	0.79	0.98	79
Ansari et al.	1996	3	60	1	2	0	0	0	0.23	0.72	0.94	1
Sum/Percents	12	306	1537	227 (74%)	14 (5%)	3 (1%)	51 (17%)	11 (4%)	3.06	8.14	10.39	118 (39%)

Overall average power was .60

Average Power 0.25 0.68 0.87

As shown in Table 7, overall statistical power for the 12 studies was .60. Power ranged from .07 for small effects to .99 for large effects. However, when power is considered in terms of individual effect sizes of .20, .50, and .80 as established by Cohen's *d* (1988), no one study had adequate power to detect small effect sizes (.20) with an average power of .25. This indicates that nearly 75% of the time those studies with small, but meaningful, effect sizes may not be found to be statistically significant. Five (42%) of the 12 studies had adequate power to detect a medium effect size (.50) with an average power level of .68. Nine (75%) of the 12 studies had adequate power to detect large effects (.80) with an average power level of .87.

The levels of power reported here are slightly larger than those previously established in the behavioral sciences. For instance Rossi (1990) examined some 221 articles published in the *Journal of Abnormal Psychology*, the *Journal of Consulting and Clinical Psychology*, and the *Journal of Personality and Social Psychology*. Specifically,

Rossi found an overall power level of approximately .52 with average power to detect small, medium, and large effect sizes as .17, .57, and .83 respectively. He concluded that levels of power in recently published work in the behavioral sciences were lower than previous power analyses conducted in the field. He examined some 25 power surveys involving 39 journal articles publication between 1962-1989 and found the average power to be .26, .64, and .85 for small, medium, and large effect sizes respectively with an overall average of .58 (Rossi, 1990). However, he is also quick to point out that a portion of these analyses relied on Cohen's earlier assessment of effect size in 1961 which were originally slightly higher than those utilized now.

It is possible that the relatively large sample size, atypical of research in this field, may have created higher levels of power than can be realistically expected under present methodological practices associated with intake populations. However, the power levels are not so high as to be unrealistic as the total sample size included within the twelve articles was 1,537 making an average of sample size of approximately 128 per study. This sample size is reasonably high and quite respectable in the arena of univariate statistics. It is important to reiterate that power levels, with the exception of detecting large effects, remained well below the acceptable level of .80 especially when dealing with small effects which is the expected effect in this area. To illustrate, it was revealed that a total of 118 tests (39%) of the 306 tests included here indicating positive effects that did not result in significant findings. This non-significant finding is in agreement with the average level of power found here (.60). Rossi (1997) explains that the number of non-significant findings should be congruent with the overall average level of power. In this case average power was .60 and 39% of the tests did not result in significant findings. These non-significant

findings may be directly associated with studies being designed with inadequate power to reliably detect small and/or medium effect sizes. Therefore, it is concluded that levels of power associated with this area of research are grossly inadequate to detect the established small effect size indicative in the field, and moderately inadequate to detect medium effect sizes.

Actual Meta-Analysis

Overall efficacy of residential treatment associated with the research findings included here appear to be effective. No one study reported the control or comparison group or pre-test score as being more effective than the treatment group or post-scores. Overall, there were reports of positive changes associated with the reduction of problem behaviors and/or recidivism. There were also increases in self-esteem, school performance, social skills, cognitive function, and family functioning. This positive outcome is evident in the transformations of statistical tests to an overall Z -score of 8.81 significant at $p < .001$ (one tailed).

As previously stated, due to the inclusion of such a small number of studies reporting outcome specifically for females ($n=6$) and minority residents ($n=2$) made it impossible to examine these populations separately with any reliability. However an exploratory look is offered. A Z score was calculated on just those articles ($n = 6$) that included outcome with regard to females (Ansari, et al., 1996; Day & Pal, 1994; Howe et al., 1998; Pugh, et al., 1997; Tucker et al., 1997; Burks, 1995). The combined average Z -score for these articles was 3.56 with $p = .0008$ (one tailed). This indicates a positive improvement for female youths as well. This analysis must be considered with caution as

the number of articles included here is extremely small. In addition, this significant finding for females is suspect in that the fail safe N or the number of additional articles containing null effects that would threaten this significant Z -score is only 22. This is well below the tolerance level for Null effects as put forth by Rosenthal (1979). This is critical information to the position put forth here, and it should be investigated further. The concept of a fail-safe N is discussed in greater detail below. See Table 8 for average Z -scores and p -values associated with each article.

Table 8: Results of Meta-Analysis

Author	Year	Design	# of Tests	# of Non-Significant Findings	Average Z (one tail)	Average p-value
*Day & Pal	1994	Longit.	32	4	2.0615	0.0855
MacNeil	1998	Comp Groups	4	0	3.0130	0.0013
*Howe et al.	1998	Comp Group	16	5	0.9413	0.1097
Simmerman	1997	Pre & Post	2	0	2.7180	0.0100
*Pugh et al.	1997	Pre & Post	57	22	1.0134	0.1590
*Tucker et al.	1995	Comp Group	3	2	0.8115	0.3350
*Burks	1995	Pre & Post	3	2	2.0188	0.3500
Kapp	1994	Comp Group	2	0	3.7670	0.0100
Wright	1995	Pre & Post	1	0	2.5650	0.0100
Carlson et al.	1994	Pre & Post	31	3	1.3832	0.0528
Deschenes & Greenwood	1998	Comp Group	152	79	8.4250	0.3520
*Ansari et al.	1996	Pre-Post	3	1	1.8659	0.1210
Sums			306	118 (39%)	30.53	1.5963
Mean Z and p values					2.5441	0.1330
Meta-Analytic Z-Score					8.81	

It should be noted that this analysis is not representative of all research conducted in the field, as it includes only that research which was published between the established time frame and within those journals associated with adolescent research. In addition, one

must also be concerned with unpublished literature. This is known in the field as the file drawer problem. It represents the number of studies showing Null results that were then not submitted for publication or were rejected (Rosenthal, 1991). Understanding the file drawer problem is especially important in light of the fact that expected effect sizes in the field are small (.26 here), and power levels to detect these differences as significant were inadequate (mean power = .60 with power to detect small effects at .25). Within a meta-analysis this problem is addressed through the calculation of a fail-safe N (Rosenthal, 1991). Essentially, this calculation determines the number of unpublished studies that would have to be sitting in the dungeons of the file drawer, possibly committed there as a result of Type II errors which are rarely considered in our field (Rossi, 1990). The fail-safe N calculates the number of these non-significant studies that would be required to make the mean $\underline{Z}$ -score found during the meta-analysis no longer significant indicating a lack of efficacy concerning treatment. Rosenthal (1991) explains "that journals are filled with the 5% of studies that show Type I error, while the file drawers are filled with the 95% of the studies that show non-significance (e.g. $p > .05$)" (p. 104). The following formula for calculating fail-safe N as presented by Rosenthal (1991) is as follows.

$$X = K[K\underline{Z}^2 - 2.706] / 2.706 \quad (9)$$

Where X is the number of non-significant studies, and $\underline{Z}$ is the average $\underline{Z}$ -score found in the meta-analysis, and K represents the total number of studies included in the analysis. In this case $K = 12$ and $\underline{Z} = 2.544$.

Results indicate that it would require 333 studies conducted (between 1994 - 1999) with non-significant findings to create a non-significant $\underline{Z}$ -score. This is nearly 28 times the number of articles (12) utilized in this statistical analysis. If one were to consider this

figure utilizing the logic suggested by Rosenthal (1991); then at the most conservative level, 95% of studies conducted in this area are sitting in file drawers. So if the twelve articles included here represent the magic 5% that were published, then that would mean that there are an additional 228 outcome studies sitting in file drawers. The fail-safe N calculated here indicated 333 articles would be necessary to overturn the significant Z . Therefore, it would take an additional 105 articles ($228 + 105$) associated with this very conservative estimate to threaten the results presented here. At best, the chances of this occurring seem unconceivable. Even to the most conservative critic, it seems very remote. In addition, 333 articles exceeds the suggested tolerance level for null effects as established by Rosenthal (1979) which is calculated to be $(5 \times \text{number of studies} + 10)$. Therefore, it is concluded that there are positive effects for youth receiving residential treatment as compared to comparison groups and pre-tests scores.

Discussion

Discussion of Phase I: Outcome Study

Through the use of repeated measure ANCOVAs, Multiple Regression and Descriptive statistics, it became evident that the "overall" results revealed that outcome for this sample of youth in residential treatment have experienced positive results as measured by significant reductions in problem behaviors and higher levels of self-esteem from intake to year four, and increased levels of functioning. This is just what the agency wanted to hear, and the overall analyses supported this conclusion. Simply put, when we look at the overall results, we see improved functioning, higher levels of self-esteem, and a decrease in overall problem behaviors from intake to year four.

However, when we look beyond overall outcome, we see no significant difference in levels of self-esteem from intake to year four for either group of females. The overall significance was relevant for the male population. In addition, the Alaskan males and Caucasian males and females all experienced significant decreases in reported problem behaviors across time. The Alaskan females did not report significant decreases in problem behaviors. When the examination of outcome moves into the individual groups, it becomes evident that not all youth are benefiting equally from treatment. The Alaskan sample continued to report higher levels of problem behaviors and lower levels of self-esteem at all follow-ups with the exception in year four when Alaskan males report virtually no difference in self-esteem as compared to White males. Outcomes for Alaskan females are especially problematic. Alaskan females continue to report significantly lower levels of self-esteem and higher levels of problem behaviors than Whites (female or male) or even Alaskan Males. They did not experience any significant increases in self-esteem, nor did they achieve any significant decreases in the level of problem behaviors across time. Caucasian females experienced a significant decrease in problem behaviors over the course of the study, whereas Alaskan females changed very little in their reporting of problem behaviors. As stated above, neither group of females experienced significant increases in self-esteem across time.

While outcome for Caucasian females is less positive than it is for males, it appears that race and gender may be interacting to create an even poorer outcome for minority females. Being female and being an Alaskan Native may make it doubly difficult to experience a positive outcome after residing in a residential setting. It is speculated that low levels of self-esteem associated with the female respondents, as

evidenced by the lack of significant gains across time as well as being significantly lower at several time points when compared to males, is indicative of a lack of treatment modalities addressing specific issues associated with female offenders. If agencies are not addressing the issues specific to the female offender as discussed by Miller et al. (1995); Mennen & Meadow (1993); Gutierrez & Todd (1997) Yandow (1990) Kaplan et al. (1991) and others, then treatments may not be available to help these young women cope with the issues and consequences of possible sexual abuse and being denied a relational way of being.

Another possible explanation for the poorer outcomes associated with the Alaskan females may be found in their multiple identity of being female, being a Native American, and being a juvenile delinquent separated from family, culture, and worldview. Do these multiple identities compound the already high levels of stress and isolation that Alaskan females may be experiencing associated with being adjudicated as a juvenile delinquent? Could this compounding effect explain the poorer outcome? Fisher, Storck, & Bacon (1999) inform us that through years of adversity associated with being an American Indian, research is suggesting that "American Indians [may be] the most severely disadvantaged population in the United States" (p. 295). It is possible that as a result of being an Alaskan Native at a historical time in which Native cultures are in a state of transition, and youth are struggling to fit in with the dominant culture, and at the same time trying to maintain connections associated with a collectivist culture, that Alaskan females sense of self is being eroded by these pressures resulting in lower levels of self-esteem and greater reports of problem behaviors. Dukes and Martinez (1994) investigated the interaction between race and gender on self-esteem and found that Native American

females adolescents reported the lowest levels of self-esteem than any of the other groups including African American and Asian females! They concluded that "discrimination (both sexist and racists) attacks both global and specific aspects of self-esteem among females. . ." (p. 111). This would help to explain why Alaskan males improved significantly in terms of greater self-esteem and lower problem behaviors throughout the course of the study. The duality of being female and a minority may be interacting to create even lower levels of self-esteem than that experienced by non-minority females or minority males. In addition, the contradictory outcome between Alaskan males and Alaskan females is evidence of the idea that low levels of self-esteem associated with Alaskan females is not just an artifact of the population. For if this were the case, Alaskan males would not have improved as they did.

As depicted in Figure 1, the treatments associated with this facility include a Positive Peer Culture (PPC). This is a youth directed group process. Group members use confrontational and assertive strategies toward another member of the group as a method of having that youth take responsibility for their actions. It is believed that this process will result in the abandonment of personal negative values resulting in the adoption of positive values as represented within the "group culture." This confrontational and assertive "treatment" is a good example of a treatment modality based on Western individualist ideals. It may work well for the White male population, however it may do little to help females and minority youths feel connected to the group as a result of this practice. In fact, especially if females have experienced abuse, it is very easy to see how being confronted by male peers (Yandow, 1990) can create disconnection and compound already low feelings of self-worth.

This treatment modality may be especially detrimental to feelings of well-being for Alaskan females. Based on a collectivist approach, this "treatment" may be viewed as damaging to the only relationships available to Alaskan females who are, many times, hundreds of miles away from their family and culture. Based on self in relation theory as proposed by Kaplan, Gleason, & Klein (1991), if meaningful relational processes are missing within the residential facility, and they are substituting relationships based on confrontational styles of assertiveness, then females, and especially Alaskan females, may be prevented from developing a positive sense of self. This loss in meaningful relationships and emotional support is then manifested as low levels of self-esteem and higher perceptions of problem behaviors.

Multiple Regression analysis informs us that there was a significant negative correlation between self-esteem and problem behaviors. High levels of problem behaviors would significantly predict lower levels of self-esteem. Problem behaviors remained a significant inverse predictor of self-esteem at all time points. Without positive gains in self-esteem, females continued to view themselves as having more problems. One could argue that this relationship is yet another artifact or a redundancy between the two instruments; in that as problem behaviors go up, self-esteem goes down. It is reasonable to assume that engaging in delinquent behavior would make us feel bad about ourselves. In contrast, there is increasing evidence to suggest that juvenile delinquency may help to increase self-esteem. For instance, Deschenes & Greenwood (1998) cite Rosenberg, Schooler, and Schoenbach (1989) who state that "delinquency may increase self-esteem among youths who are lower in socioeconomic status" (p. 268). This very interesting phenomenon and its implications for treatment have not been investigated here. The

purpose of its inclusion is to provide evidence for the idea that low levels of self-esteem and high problem behaviors can vary and are therefore not an artifact between the two variables.

The remaining treatments associated with this facility include behavioral modification techniques such as the use of a level system, educational day treatment, and individual counseling. All of these overarching treatments, including PPC, are designed to increase self-esteem and efficacy as well as feelings of connectedness with others. These are then believed to result in the valuing of self and others leading to a reduction in problem behaviors and higher functioning. This projected outcome was not evident for Alaskan female and to a lesser degree Caucasian females. Group processes may have made it impossible for females to experience greater connections while compounding already low feelings of self-worth.

Overall Multiple Regression analysis supported the hypothesis that being White and being male would be a significant predictor of self-esteem. Within this sample the "mean" age of first sexual contact was 10.5 years old for Alaskan females and 11.3 years Alaskan males. This is astounding! At minimum it suggests inappropriate sexual activities. At its worse, it is evidence of possible sexual abuse. Age of first sexual contact for Caucasians was also low; Caucasian females at age 12 and for Caucasian males it was 12.6 years of age. Multiple Regression analysis revealed that in the 12-month interview, age of first sexual encounter was a significant predictor of self-esteem. The higher the age, the higher the self-esteem. Being closely involved with this organization it is my testimony that no specific treatment was incorporated into this facility to address the issues associated with early sexual contact or sexual abuse. Based on the contributions of

Gutierrez & Todd (1997), Mennen & Meadow (1993), Miller et al. (1995), and Yandow (1990), the lack of treatment with respect to issues of possible sexual abuse, especially in light of the fact that the Alaskan sample was well below the norm for sexual encounters, may have significantly contributed to the poorer prognosis for females and especially for Alaskan females.

If resiliency is related to self esteem as suggested by Brooks (1994), then the lack of "righting of self" for the Alaskan females may be representative of this relationship. Resiliency can be thought of as an adaptive process involving coping behaviors that allows one to deal with things as they come in such a way that future upsets are not compounded by previous challenges. Yet, it seems that Alaskan females, and, to some degree Caucasian females, are not experiencing this coping strategy or have not yet been taught of its possible existence within them. Thus, they remain stuck in a reciprocal negative feedback loop. Addressing this relationship would seem to be of paramount importance when treating female juvenile delinquents. That is, the low levels of self-esteem resulting in a lack of resilience coupled with the desire to remain connected to primary relationships, and then being in an environment that does not directly foster the substitution of new meaningful relationships, appears to have contributed to the somewhat stagnant recovery of Alaskan females as they try to adjust from being labeled a juvenile delinquent removed from their home, family and culture.

Addressing these issues in treatment and helping youth to adjust after treatment with aftercare follow-ups is extremely important (Curry, 1991; Blackman & Chowdhury, 1991). The importance of aftercare, or the lack there of, is evident in the falling scores on self-esteem and increases in problem behaviors in year two--one year after discharge for

many youths. If this outcome study had not continued to collect data beyond the two year period, this "righting" trend may have been missed. That is, it appears that most youths, except for the Alaskan females, experienced an improvement after a relatively long readjustment period. As previously stated, follow-up treatment may be paramount to the recovery of our female and minority youth.

There was no evidence to suggest that family cohesion plays a significant role in self-esteem. This is surprising in that much of the literature regarding residential treatment expounds on the importance of the relationship between family connectedness and positive outcome (Nobel & Gibson, 1994; Ponce, 1995; Burks, 1995; Stage, 1999; Kapp, 1994; Wright, 1995; Deschenes & Greenwood). It may be that the cohesion scale of FACES III was not a good indicator of need or desire for family connectedness. As stated in the methods section, this instrument was only included in the analysis at intake and at 12 months. This was because the protocol of the study involved completing FACES III only if youth had previously lived with family members during the last six months. In retrospect this was a poor choice. The instrument should have been administered whether the youth had lived with their family or not. At intake, a positive correlation between family cohesion and self-esteem of .29 was approaching significance at .054. This relationship indicated a Beta of .21 with a t -value of 1.285, $p = .21$. However, at the 12-month follow-up this correlation indicated a negative non-significant relationship between self-esteem and family cohesion. The reason for this is unclear. It is speculated that perhaps youth were trying to adapt to the loss of family relationships by denying their existence. Or perhaps, it indicates the presence of a suppressor variable influencing outcome.

Without a control group it is impossible to make any statement as to whether the overall decrease in problem behaviors and increases in self-esteem observed here are actually due to the treatment or are simply a matter of maturation or the result of some other mediating variable. However through the use of a standardized measure such as the Achenbach Youth Self-Report (1991) some meaningful comparisons regarding outcome can be made. For instance, even though all youth improved on the overall scores on the YSR, when compared to norms established by Achenbach (1991), Alaskan females remained in the clinical range with an average total mean score of ≥ 60 for all time points except for a mean of 52 during the 48-month follow-up. Achenbach (1991) defines this range (60) as the beginning of the clinical range indicating the presence of deviant behavior or thought processes. This level of problem behaviors was only evident for Caucasian females at intake. This escalated pattern in mean scores was also evident for Alaskan males in that they began within the clinical range and then dipped at the 12-month interview with a mean score of 47.71 (still in treatment). They then escalated back up to the clinical range during the 24 and 36-month follow-ups (after treatment). However, scores for the Alaskan males were below the clinical range at 48 months and were closer to those reported by White males with a mean score in the low forties. White respondents did not enter the clinical range as defined by Achenbach (1991) after intake. This indicates that the Alaskan population may be having greater problems with adjusting to life after discharge. This finding also supports the need for aftercare, especially with minority youths.

In summary, after four years of tracking these youths it appears that many in this study improved in several areas of functioning. Most no longer experience clinical levels

of problem behaviors and have experienced greater levels of self-esteem. The majority of those contacted are managing to live independently without evidence of any long term inpatient treatment or incarceration. Many are responsibly managing relationships (including children) and work. However, most have not really reached a level of functioning that would be considered normal by society (living in poverty with many experiencing substance abuse problems). They report violating some rules of society (such as drinking while driving or using illegal substances) although most have not yet experienced significant problems with authorities. It is worrisome that the majority of youth report severe problems meeting their basic economic needs. It is also disturbing that Alaskan females showed very little improvement during or after treatment. With the exception of Alaskan females there is reason for guarded optimism about the future of these now young adults.

The findings here indicate that agencies need to understand the effectiveness of their treatment with respect to the populations that they are serving. If we as researchers lump our participants and the overall effects into one conclusion, it is highly possible that important components associated with outcome may be missed especially when working with diverse populations. Superficial analysis of our treatment populations is not enough to understand the efficacy of treatment. We must peel back the layers of our overall findings and seek a deeper and more comprehensive understanding of the actual meaning behind our overall "significance".

Phase II: Meta-Analysis Discussion

The meta-analysis revealed that youths receiving residential treatment do experience improvement in several areas of functioning. It was found that 333 articles containing null-results would be necessary to conclude no positive effect existed. It was also established that average effect size for this area of study is expected to be small. This was supported by other meta-analytic inquiries previously conducted prior to 1990 that were summarized by Lipsey (In Cook et al., 1994). Lipsey concludes that small effect sizes should be expected. It was also concluded here that power levels to detect this effect size are grossly inadequate. This was also the case for Lipsey. He states that a median sample size of 60 does "not yield sufficient statistical power for an individual study to find statistical significance for effect sizes in the range of .10 - .20 standard deviation units" (p. 126). Therefore, inconsistent findings associated with the efficacy of residential treatment as put forth by Wells (1991), Asarnow, Aoki, & Elson (1996), Curry (1991), Blackman, Eustace & Chowdhury (1991) and others, may be more statistical in nature than they are for varying treatment modalities and/or the use of differing outcome variables. That is, if overall effects in this area are small and we have grossly low levels of power to detect such differences as "significant," and we rely solely on dichotomized testing for meaning, then it is no wonder that no consensus exists in the field. We are capitalizing on chance. Our narrow methodological practices are preventing us from seeing the forest (efficacy) through the trees. It's like trying to find your keys on a darkened beach with a pin-light. You might get lucky and find them. But, not finding them does not mean that they are not there. Why not use a bigger light!

One of the limitations of this meta-analysis is that it did not focus on discriminating differing outcomes based on specific outcome measures or treatment modalities. For instance, a meta-analyst might compare outcome using only scores retrieved from the Achenbach Child Behavior Checklist or YSR (1991), or they could examine outcome based on frequencies of recidivism only. Another might look at a particular treatment and include only those outcomes associated with that modality. In addition, the approach here did not statistically establish the effect on outcome associated with mediating and/or moderating variables. Though these variables were certainly discussed in relation to their theoretical implications for treatment, and it was established that many times these all important variables, with respect to the populations investigated here, are ignored in the treatment as well as in the reporting of outcome. Completing this more specific type of meta-analysis is a much more sophisticated approach likened to the differing levels of sophistication associated between doing a Multiple Regression analysis and a Canonical Correlation. Both methods provide information regarding relationships, but the latter provides a much more micro look at that phenomenon under investigation. This type of analysis is also critical, and it is suggested that future directions should include the use of a more in depth approach.

The current meta-analysis explored outcome as a whole based on levels of improvement on a variety of measures. Variables assessing outcome within the 26 outcome articles were very diverse with duplication occurring mostly with behavioral checklists such as the Achenbach CBCL (1991); overall performance while in treatment (level system); least restrictive environment after treatment; social skills; assertiveness; recidivism; school performance; and self-esteem ratings. Specifically, outcome variables

included the following: youth perspectives or satisfaction with services (Shennum, 1995); Achenbach Child Behavior Check List, CBCL (Day & Pal, 1994); overall level of functioning (Zimmerman, 1994); social competency (Macneil, 1998); emotional and social skills, sexual acting out, and levels of aggression (Howe et. al., 1998); level system and CBCL (Larson & Calamari, 1998); school performance and previous placement (Simmernan, 1997); CBCL, WISC, behavioral files (Pugh et al., 1997); story telling (Micell, 1987); self-esteem, assertiveness, CBCL (Tucker et al., 1995); recidivism, length of time in discharge placement (Burke, 1995); discharge to least restrictive environment (Stage, 1999); school performance after discharge (Rafel, 1994); post treatment imprisonment (Kapp, 1994); social skills, length of time in discharge placement, recidivism (Wright, 1995); completion of treatment, behavioral discharge records (Carlson et al., 1994); decrease in problem behaviors, increase in positive attitudes toward school and community (Pazarat, 1999); achievement of goals (Brown & Hill, 1996); school performance and changes in attitude toward school (Smith et al., (1996); overall performance while in treatment (Knorth & Elderling, 1998); Global Assessment Scale (overall functioning), restrictiveness of living environment, CBCL, and least restrictive environment at discharge (Daly et al., 1998); self-esteem, family functioning, social and coping skills (Deschenes & Greenwood, 1998); youth satisfaction, isolation from family and friends, and personal control (Friman, 1995); overall performance in treatment based on level system (Ansari, 1996); and finally, attachment (Penzerro, 1995). As one can see outcome variables vary wildly. This finding was evident in the literature regarding criticism in the field as put forth by Zimmerman (1996), Pratt & Moreland (1996), O'Mally (1996) and others. It is speculated that the use of numerous outcome variables

may have hindered efforts to have an accumulative knowledge regarding efficacy. Many times measures were artifacts of the agency that do not allow for comparison across facilities let alone groups. However, methodological issues contributing to the inconsistencies in the field should not be underestimated.

Specific outcome variables used to assess efficacy for females and minority youth (of which only six included the appropriate information to be included in the statistical portion of the meta-analysis) included the following: attachment (Ansari, 1996); school performance after discharge (Rafel, 1994); discharge to least restrictive environment (Stage, 1999); Achenbach Child Behavior Check List, CBCL (Day & Pal, 1994); overall level of functioning (Zimmerman, 1994); emotional and social skills, sexual acting out, and levels of aggression (Howe et. al., 1998); CBCL, WISC, behavioral files (Pugh et al., 1997), story telling (Micell, 1987); self-esteem, assertiveness, CBCL (Tucker et al., 1995); recidivism, length of time in discharge placement (Burke, 1995); overall performance while in treatment (Knorth & Elderling, 1998); and post treatment imprisonment (Kapp, 1994). As one can see, most of the outcome variables do not address many of the issues associated with females and minority youths discussed above. However, it is hopeful that levels of attachment and sexual acting out behaviors were included. These variables may be more relative to outcome associated with these populations. See appendix D for citations associated with all 86 articles retrieved relating to residential treatment in general.

Contributions and Future Directions

While the above findings do make a contribution to the field, it is nonetheless, very disappointing to this researcher that the original purpose of this project could not be accomplished: understanding outcome specifically for females and minority recipients of residential care. This is especially disheartening in light of the fact that when Phase I of this project was conducted, it appeared that, overall, youth were experiencing greater reduction in problem behaviors and higher levels of self esteem at most time points. However, with closer examination, it was revealed that minority females were not improving at the same rate as their peers including Alaskan males. Alaskan females remained in or very close to the clinical range for problem behaviors throughout the duration of the study. While the primary purpose of a meta-analysis is to synthesize, it seems that a major contribution of this meta-analysis has been revealing the neglect in the field in identifying outcome for females and minority youth. It has also highlighted methodological problems evident in the paucity of reporting indices of effect including simple means and standard deviations.

As a result we are only slightly more informed about outcome associated with females or minority adolescents than we were when this project began. The significant, but suspect, Z-score associated with those articles addressing outcome for females, while interpreted with caution, is an indication that we should be more tenacious in our mapping of treatment efficacy by going beyond identifying treatment modalities which are "believed" to work for all because we have found significance. It is equally important that our search for "truth" should be to empirically understand how those treatments or in this case placements impact groups of youth differently. Is it as Burt (1993) suggests? Are we

simply to accept one mode of psychological development as being superior simply because it is the dominant view? Does the significant Z -score found here only apply to outcome for the dominant White males? It appears that it does. But we don't really know for sure, because researchers did not ask. These are questions that must be addressed in future investigations as we seem to be missing the point. For example, Pratt and Moreland (1996) in their very extensive critique of outcome studies related to residential care, including the use of meta-analyses, expound on the importance of identifying those factors that may "hinder improvement." However, they completely neglected to consider the importance of race as an outcome predictor. Had they done so, this would have helped to incorporate the differing needs and worldviews associated with developing in a collectivist and/or individualistic environment and their implications for outcome.

One could speculate that the reason for not actively investigating outcome for specific groups within this area of research is related to the small number of participants in each study. One could go on to say that breaking up this already small sample size into even smaller groups would create lower levels of statistical power and limit statistical applications. However, if researchers have an understanding of the expected effect size for a field, then studies could be designed with appropriate levels of power to detect that effect. In addition, research designs in the field could benefit from utilizing longitudinal data that statistically increases the number of participants in each cell of the design, and reduce error through the statistical use of ANCOVA or MANCOVA designs allowing for greater effect. If these practices were more widely adopted, then perhaps we could gain a better understanding of outcome for all youth receiving residential treatment.

The results here indicate that at the very least it is time that, as researchers, we must fully define our samples with regard to ethnicity, SES, age, sex, attrition, length of stay and aftercare placement. We must also responsibly report all indices of effect including our descriptive statistics from which our analyses are structured upon. One of the criticism of meta-analytic procedures put forth by Pratt & Moreland (1996) is that many times meta-analysts are "forced to estimate, assume, and infer data that have not been reported, but are nonetheless crucial for the purpose of comparison" (p. 7). If researchers are comprehensively reporting their statistics, this criticism could be eradicated allowing for greater inclusion for synthesis resulting in a more cumulative truth.

As Rossi (1997) illustrated with his meta-analytic work within spontaneous recovery theory, the loss of cumulative knowledge in psychological research can and has resulted in the "death" of important theories, which if properly investigated may have contributed greatly to the area of study and brought all those involved closer to the "truth." Simply saying that we have found significance is not enough. As researchers we must be asking what is the size of the effect? How confident are we of that effect? How well did our design allow us to detect that or any other effect? Who is being affected? How are they being affected? Who is not being affected? Why, what is the context of the effect? If we are saying that something works without understanding for whom it may be working for, then this is analogous to committing a Type I error. That is, if a proportion of our participants are not benefiting from treatment, and researchers' are saying that they are because they have found overall significance, then we have committed an ethical Type I error. The overall p value may not represent the only truth--we may find that group means

have different "truths," and dichotomized testing in and of itself may not always prove to be the most appropriate way to proceed. As Rossi quotes (1997), "The Truth is Out There" (p. 175). As researchers, it is our ethical responsibility to find "it" through the most appropriate method/s possible and report it comprehensively. This researcher will continue to look for the "truth" associated with positive outcome for females and minority problem youth. As stated earlier, a more sophisticated employment of meta-analytic procedures is needed. However, if efficacy can not be fully understood through the use of quantitative methods alone (due to a lack of comprehensive reporting), then perhaps the field and our investigations would benefit from the inclusion of qualitative methodologies. This integration of method may help bring us even closer to the "truth."

The overall contributions of this thesis may help to produce a more cumulative understanding regarding effective treatment modalities and research practices regarding efficacy of residential treatment for female and ethnic minority adolescents. As such, its contributions can be far reaching. However, it must be reiterated that many questions remain unanswered and much work remains to be done.

Phase III: Epilogue

The final Phase of this project is important to include as it represents a reflection of the work that I have just completed. The mental action inherent in critiquing something can lead to a new way of thinking and understanding that may have not been evident prior to the reflection. So, in this case, the product of the reflection, the Epilogue, emerges, simultaneously as a prologue. A prologue, as defined by Webster's is considered "a preliminary act or course of action

foreshadowing greater events." This meandering in thought is included here to illustrate the fundamental connections between what was, what is, and what is to come. The course of action foreshadowing "greater events" within this paper begins with a personal reflection and goes on to discuss issues associated with our traditional approaches to scientific inquiry. This discussion includes some of the challenges of this approach as offered by philosophers of science and others advocating change within our scientific paradigm and existing methodological practices. This process will then lead to an identification of my own theoretical position with regard to ontological and epistemological positions as they are now. Issues of congruence will be discussed with regard to my own philosophical position and my resulting product: my thesis.

Personal Reflection

Throughout this project I have tried to remain congruent with what is believed to be true for me as a woman and what my science tells me is true. As a beginning methodologist it has been a challenge to bring these two worldviews into focus: for one does not always allow for the other. Throughout this paper there has been a concerted effort to include multiple indices of effect with multiple methods as well as multiple interpretations so that a more comprehensive truth may emerge. To accomplish this task I have searched and re-searched for meaning and connections to help explain the statistical analysis. For me, including the connections and the meanings of our discoveries is equally as important as engaging in the scientific process. The use of such expanded techniques was an attempt to honor my own way of knowing and ontological position that

includes a metaphysical view of reality. Specifically, my philosophical position includes ideas associated with the "new science" in which it is believed that there is an underlying reality based on wholeness and an inherent connectedness to all things. This worldview is an expanded one. To work as a scientist within this ontology requires an expanded scientific paradigm. That is, if we are to holistically understand reality (that which we choose to study) then we must include "all of evidence" (Baars, In Harman, 1991 p. 34). The meta-analysis included here exemplified this type of integration and holistic synthesis as that proposed by many seeking an expanded scientific paradigm. My position is encompassed in the words of Walt Whitman below:

*A vast similitude interlocks all,
All spheres, grown, ungrown, small, large, suns, moons,
planets,
All distances of place however wide,
All distances of times, all inanimate forms,
All souls, all living bodies though they be ever so different,
or in different worlds,
All gaseous, watery, vegetable, mineral processes, the fishes
the brutes,
All nations, colors, barbarisms, civilization, languages,
All identities that have existed or may exist on this globe,
or any globe,
All lives, and deaths, all of the past, present, future,
This vast similitude spans them, and always has spann'd,
And shall forever span them and compactly hold and enclose them.
-Walt Whitman*

*Leaves of Grass
In Space, Time and Medicine (Dossey, 1982)*

Traditional Scientific Views

There is nothing which can better deserve our patronage than the promotion of science and literature. Knowledge is in every country the surest basis of happiness.

*George Washington,
Address to congress, January 8, 1790*

In Sagan(1996): The Demon Haunted World: Science as a Candle in the Dark

Our existing scientific paradigm was founded on a logical positivist approach encompassing a mechanistic model of reality. Within this paradigm, it is believed that reality "is driven by immutable laws" (Guba, 1990, p. 19). As such, it can be objectively defined and measured to the point that predication and control are possible. This approach represents a realist ontology, which posits that reality is external--it is out there--separate from human consciousness. Guba (1990) goes on to define this ontological approach as "Knowledge of these laws and mechanism is conventionally summarized in the form of time and context-free generalizations. Some of these generalizations [then] take the form of cause and effect laws" (p. 20).

According to Guba (1990) and many others, the ways of knowing within a realist's ontological position involves a "dualist/objectivist" approach. It is believed that the observer is able to remain detached and objective (value free) as the "object" under study is subjected to experimental manipulation designed to control for any extraneous or confounding variables; including the consciousness of both the observer and the observed as well as the context under which the testing is taking place. Packard and Polifroni (1990) describe a realist perspective as a "philosophical stance regarding the nature of truth . . .[which states] the picture science gives of the world is true, is real, and that the entities postulated do, in fact, exist. In addition, real things are in no sense condition by their being known" (p. 159).

We could place this pure realist's ontological position on the far-left side of a continuum. A circle at the center interrupts the ontological continuum. Beyond the circle, at the right end of the continuum, we could place those holding the anti-realist perspective (Burbank, NUR 660, 1999). Within the relativist/constructivist philosophy "reality exists [only] in the form of multiple mental constructions, socially and experientially based, local and specific, dependent for their form and content on the persons who hold them" (Guba, 1990, p. 27). Objectivity and external reality are not separate from subjective experience and universal laws do not exist. Subjectivity is "truth." Truth is dependent (relative) on an individual's or collective consciousness. In contrast to a realist perspective, reality is believed to be constructed relative to the individual, and ways of knowing involve subjective inquiries which may better align themselves to qualitative methodologies. In addition, reality from the relativist position is found "within" the mind and consciousness plays a fundamental role.

The realist perspective has undergone multiple challenges especially with regard to objectivity. These challenges, stemming greatly from multiple works by Thomas Kuhn (In Klemke, Hollinger, Wyss Rudge, & Kline, 1998) as well as from many feminist scientists have resulted in an expansion of the realist's perspective. While it is beyond the scope of this reflection to discuss all of them here, the revised ontologies have questioned the issue regarding the achievement of objectivity and a detached reality. This controversy has resulted in the formation of revised philosophical positions allowing for a modified realist's ontological philosophy. The revised ontologies are residing more toward the right

side of the continuum. A realist, whether they be a post-positivist, critical theorist (feminist theory), or a modified objectivist will always hold the belief that reality exists "out there" distinct from the individual's experience within it. Although it may be recognized that we can never really "know" complete truth and that objectivity is not really always possible (modified objectivist's position), to be a realist is to remain to the left of the circle.

This same progression toward the circle is also evident for those in the relativist's camp; except their movement has been from the far right of the continuum toward the left. So, in search of a comprehensive truth, the realists are moving toward the right and the relativists are moving toward the left. As a scientific community we are coming closer and closer toward the circle. The circle may represent a place of unification. Like other scientific revolutions that result in greater explanation of reality by encompassing multiple theories, perhaps there exists a unifying theory of reality that will explain the positions of both the realist and the relativist contained within the circle (Burbank, 1999). Perhaps, there is a distinct (objective) reality and as Burbank (1999) has suggested, consciousness has created it. Is this the message in the circle? Is this the movement of the future? These are the ontological "dances" that are currently taking place in our scientific community.

When science is spoken of here I am primarily referring to my own discipline of psychology, though these debates permeate all of the sciences. Looking at the history of our science, it is astounding to see how much of our theories regarding what science is today (or is not) is reflected in the socio-cultural

lives of the scientists themselves. For example, one can see how the Western ideals of separation and autonomy are manifested in the positivist view (objectivity). This ontological position has resulted in the "belief" that a complete objective reality could be achieved and of course measured (because we are separate and autonomous). It even goes as far as to declare that control/domination of Nature is possible and preferable. Keller & Longino (1996); Landrine (1995); and Unger and Crawford (1996); as well as many other feminist theorists inform us that through the historic exclusion of women as scientists and only recently as groups of interest in scientific investigation, the history of our science is based on male values and is documented by male history (his-story). For instance, through the use of critical theory, which seeks to understand phenomenon based on the sociopolitical context in which it occurs, Landrine et al. (1995) emphasizes the importance of honoring the feminist psychologists' concern regarding their reluctance to conjoin with the practices of traditional scientific method. Many feminist psychologists perceive traditional methods (rightly so in many instances) as exploitive and serve to "recapitulate patriarchal forms of relation" resulting in the stratification of gender, class, and culture. In addition, feminist psychologists posit that traditional methods devalue subjective experience and encourage, if not demand, the [removal] "of behavior from its context for the sake of control" (Landrine 1995, p. 56). Nonetheless, Landrine et al. (1995) also suggest that feminist psychology must rise above the fear of possible further marginalization resulting from "methodological deviance" and "must adopt a revised methodology that consists of traditional and of person-centered [contextual] approaches" (p. 58).

Harlow, Muliak, & Steiger (1997) also suggest that traditional dichotomized hypothesis testing is not sufficient, nor, by its very nature, can it capture the complexity of reality. They go on to emphasize the need to include multiple ways of evaluating scientific models in addition to significance tests. Social and behavioral sciences are often criticized for their lack of consensus and cumulative theoretical positions (Meehl, 1978; Rossi, 1997). Perhaps our lack of consensus is in part due to the fragmentation of the theorists themselves resulting from adherence to methodological dogma based firmly on the idea that an objective reality exists: culminating in the belief that as detached scientists we can capture it without altering it through our manipulation and observation and present it as truth. For example, the scientific enterprise of psychological research is firmly embedded in the idea of dichotomized significance testing (Rosnow & Rosenthal, 1989). This method results in dichotomized conclusions based solely on their significance without consideration given to the importance of context. These conclusions, at times, can be both damaging and prejudicial depending on whom is being deemed different from whom. While this method may be very important and valid in some contexts, it should not be considered an end all in scientific inquiry. It can be seen that this paradigm, though useful in many areas of investigation, is not comprehensive nor does it allow for the interpretation of meaning in the absence of "significance."

Presently, our post-positivist positions provide us with a linear representation of reality. It is "truth" based on the classifications (stratification) of people into groups. This method allows for the examination of differences.

The paradigm excludes or "controls" understanding in context and does not allow for the examination of the meaning of the differences, nor does it include conscious processes in the analysis. In essence, we have traded "richness" of knowledge (subjective experience) for what we believed would give us consistent findings (objectivity) (Dossey, 1982). While the examination of some differences may be important to understanding, without interpreting the reasons those differences exist, we have missed finding the truth--we know the symptoms, but not "cause."

As stated, this method is valid in many contexts, and it is not suggested here that the traditional scientific baby should be thrown out with the bath water. As Harlow et al. (1997) points out, the use of multivariate techniques and the inclusion of multiple indices of effect (expanded methods) will help us achieve a richer and more contextual understanding of our investigations. However, as suggested by Baars (In Harman, 1991) we must include "all the evidence." As we move toward the circle in search of unification, all of the evidence includes the subjective contributions of both the observed and the observer.

Wheelis (1971), in his philosophical discussion on the relativity of knowledge as proposed by Karl Popper, reminds us that our perspectives as theorists color that which we observe or measure. He goes on to say that as a result, "Induction is, therefore, a myth and no number of confirming instances, however large, can establish truth. Our best established laws, therefore are conjectures, very well founded perhaps but never, so far as we can know, finally true" (p. 91). While this perspective may be adamantly rebuffed by many, it has

nonetheless been partially resolved through the movement toward the circle by critical and modified objectivists, to name but a few.

This discussion leads us to the point that we can no longer study differences for the sake of finding differences. From a critical perspective (associated with feminist scientific inquiry), if we are to effect change we must identify the social constructions and theoretical and philosophical positions that perpetuate the existence of those differences. To do this, we must be looking for the connections that explain the meanings for the existence of those differences. For example, Unger (In Landrine, 1995) explains that in our attempt to include the issue of diversity in both our methods and our participants (data base) we must be wary so as to not measure diversity simply in terms of differences as we have done with gender. Unger proposes that we must investigate the "meaning" of differences as they relate to the situational context in which they occur. We can not accomplish what Unger and other feminist psychologists' suggest when we rely primarily on group difference statistics and omnibus F tests.

If many traditional methodological approaches do not allow for the examination of complex, diverse, and interrelated latent realities inherent in human beings and society, then it seems that we must broaden our methods so that they can capture the richness of reality that extends beyond logical positivism. We must go "behind" the tapestry to understand the complex patterns that make up the "observable" side of the tapestry (Miller & Burbank, 1994). After all, isn't that what psychological (scientific) inquiry is all about? Shouldn't it be equally important to incorporate experimental methods, philosophical and theoretical

orientations that embrace the study of complex realities within a contextual framework? It seems that it should!

On a more controversial note, perhaps the lack of cumulative knowledge in many sciences is an indication that not just the methodology is at debate, but perhaps it indicates that a scientific reformation is necessary. That is, contradictions in findings may be the result of anomalies that exist due to a representation of reality that is incomplete. Historically, anomalies leading to contradictory findings in science have lead to scientific revolutions. For example, monumental shifts in scientific thought were only possible when existing knowledge was challenged by contradictory scientific findings. For example, findings associated with Copernicus (1543, *The Revolution of Celestial Orbs*); Galileo (1638, *Discourse on Two New Sciences*); Newton (1687, *Principia*); and Einstein's theories of relativity (1905 and 1915) rocked the beliefs of what was known to be true at the time (Kuhn, 1970). The work of these scientists and their findings resulted in science entering (kicking and screaming) into a new worldview, each, leading to new scientific revolutions and, ultimately, an expanded view of reality.

Kuhn (1970) (In Mosedale, 1979), supports this view of science and suggests that perhaps the progression of science is not found in its accumulative knowledge at all, but is based on its theoretical revisions. According to Kuhn (1970), cumulative science is referred to as "normal science." Within this state, knowledge "is predicated on the assumption that the scientific community knows what the world is like" (In Mosedale, 1979, p. 223). The unexplained anomalies

challenge "normal science" resulting in shifts in thought that lead to further scrutiny by other scientists as they try to maintain continuity of thought. This intensive inquiry may then lead to discard older theories in favor of a new way of thought. Hopefully, the product of this process is an understanding of some phenomenon that is more unifying (explaining more of reality) than that which was previously possible with the former theoretical positions. This revision then extends and expands science's ability to represent a more holistic and comprehensive representation of truth and knowing.

While it is true that I see myself as a realist, I also see myself as close to the circle as my science will allow me to get. Burbank (1999) has proposed, and I concur, that perhaps there is a unifying whole within the circle that borders these seemingly mutually exclusive truths. Suppose we were to say that yes, reality exists, but collective consciousness and subjective experience have created it (Burbank 1999). If this were the case, then a unifying theory might be possible. This new theory would have to explain both the realist and the relativist's perspective with one overarching philosophical and theoretical perspective. To deny either ontological position merely serves to alienate us from each other, our methods, our theories, as well as from the potentiality of our science to reflect a more holistic representation of reality. Keller (1996), Miller & Burbank (1994) and others express this integration as a need for "logical congruence" between "philosophy, theory, and method" (Miller & Burbank, 1994). Miller & Burbank (1994) go on to contend that, "Philosophy, theory, and method are interrelated and fit together like pieces of a puzzle" (p. 704). Reality, itself, may also fit together

like pieces of a puzzle. Therefore, excluding or simply denying the existence of pieces does not allow for the entire picture to emerge. We need to be collecting and connecting our pieces, not placing them on opposite sides of the puzzle or trying to analyze a part (a side) as if it represents a whole. Integration, blending, and expansion resulting in "logical congruence" is what's needed (Miller & Burbank, 1994).

Now, let's suppose that this unifying theory may be explained by employing the fundamental principals inherent in the emergence of the "New Science" that is influencing our technology and intellectual thought today (Harman 1991; Bohm, 1980; Zohar, 1994; Zokav, 1979). While the ideas inherent in the new science will be more extensively explored as will this Epilogue during my dissertation, this idea will be briefly presented here. To call this philosophical position a "new science" is a misnomer. For the "new science" has its origins within classical physics. Within the "new science," complex patterns and relationships among phenomenon are fundamental not the mechanistic nature of that phenomenon. The Newtonian mechanistic model would give credence to the idea that within an objective universe, the examination of smaller parts from a larger whole would allow us to observe mechanical like interactions. We would see more clearly the mechanistic nature of that which we study, thereby confirming that law. However, when phenomenon is examined within the context of quantum physics (the study of subatomic particles) and relativity, this assumption does not hold true. For instance, light was found to present itself as both a particle and a wave and behaved as such depending on

how it was measured. This duality of light illustrated the presence of the Heisenberg's uncertainty principal where "any attempt to gain knowledge [observe phenomenon] at the atomic level [photons] involved inevitable changes in what was being observed" (Miller & Burbank, 1994, p. 705). When light was measured in one way, it behaved as a wave. When it was measured in another way, it behaved as a particle. This illustrated the fact that through the act of observation (measurement) that which we observed is changed by that very act. Therefore, as stated by Miller and Burbank (1994), "It was not possible, even in principle, to view the smallest material realm with completeness and certainty" (p. 705).

Also, the behavior of subatomic particles themselves did not coincide with a mechanistic model. In fact, they behave the opposite to that which was expected within this model. For example, John Bell "analyzed the logic of certain measurements on systems in which two particles interacted and were then separated" (Harman, 1991, p. 8). He found that two particles, that were once part of a larger particle, continued to act as a single particle even though they were separated by time and space. When one particle was spun, the other also spun simultaneously. This event directly confirmed the assumption that within a quantum universe there exists a connection to all things uninterrupted by time and space. Zukav (1979) quotes one of the leading theorists of quantum physics, Henry Stapp (1971) who describe the physical world of quantum mechanics as:

... not a structure built out of independently existing unanalyzable entities, but rather a web of relationships between elements whose meanings arise wholly from their relationships to the whole (p. 103).

This provides evidence that all particles within the universe are connected and interaction within any system influences all other systems. Stapp goes on to contend that Bell's theorem also holds true at the macro level. He states:

"The important thing about Bell's theorem is that it puts the dilemma posed by quantum phenomenon clearly at the realm of macroscopic phenomena. . .[it] shows that our ordinary ideas about the world are somehow profoundly deficient even on the macroscopic level" (In Zukav, 1979, p. 306).

Thus, it is concluded that "if we are to speak of an objective reality behind the appearances, then that reality must be interconnected over all space and time" (Harman, 1991, p. 8). These perspectives are in direct opposition to the concept of an objective universe in which all things are separate and individual and remain unaffected by observation. So we have a dilemma. If we are to modify our ideas regarding the existence of an object reality within our investigations [based on the evidence inherent in classical physics], then we must also give up the idea of a detached separate science based on objective observations. Alternatively, we must adopt a view a science as a holistic science as well as a reality that is inherently connected and changed by the observer and the observed (Harman, 1991). Harman (1991) quotes the work of physicist John Wheeler (1987) who

eloquently describes the necessary consequences to scientific ontology related to such findings and states:

"The idea that behind the appearances. . . is a more real world that is empty and dead in which pieces of matter move mindlessly about forming chance combinations; the idea that the mind is somehow separate from all of this, observing it passively as mirror--these ideas can no longer be taken as the basis for any ultimate belief about reality" (p. 9).

Now, for the role of consciousness in the creation of reality. A quantum view of reality includes the occurrence of action at a distance, or non-locality such as that exemplified in Bell's theorem with the simultaneous action of the spinning particles. It is also explained when a person realizes (feels) as though a loved one is in trouble and then this "message" is confirmed in "reality." This quantum view exemplifies the existence of synchronicity, where seemingly meaningless coincidences occur that may result in monumental life or scientific shifts in thought. Since all matter consists of essentially the same "particles", then as humans, we too are quantum beings, living and interacting in a quantum universe with the fundamental characteristics of sharing and experiencing connectedness. Harman (1991) concludes then that quantum theory is also inherent in consciousness. He explains the ideas of Roger Sperry's (1987) work regarding "order" as a "consciousness revolution." Harman (1991) states that "order" is observed in the physical world, but [that order represented in physical reality] is never free from the possibility of a "downward causation" [from consciousness]

"down to the physical" (p. 100). It can be conceived here that Harman and Sperry are suggesting that through this downward causation process, consciousness is an active participant in the creation of reality. Zukav (1979) also suggests that consciousness is a quantum experience. In his discussion of how protons "knew" how to respond to the wave particle duality of light, he highlights evidence which suggests that through their connections in a quantum universe, photons may be organic and therefore be conscious. This consciousness then creates an event, which was observed by the scientist, in this case, by Thomas Young in 1803. Zukav (1979) quotes physicist E. H. Walker (1970) who supports this contention and states that:

"Consciousness may be associated with all quantum mechanical processes. . . .since everything that occurs is ultimately the result of one or more quantum mechanical events, the universe is inhabited by an almost unlimited number of rather discrete conscious, usually nonthinking entities that are responsible for the detailed working of the universe" (p. 88).

Therefore, Zukav (1979) as well as many others not included here, conclude that when something is organic, it has the knowledge to process information "and to act accordingly." Photons are responding intelligently to the different experiments as established by the physicist. Thereby, they are creating a reality (viewing light as a wave or a particle) through their organic or conscious processes. Of course this logic is based on the premise that everything organic has consciousness. Physicists state that this organic nature of photons presents

itself in reality as a wave (a probability wave) (Zukav, 1979). This wave is then measured as a tendency or probability of producing some event. Even if that event does not occur, the potential or probability of it occurring remains within the collective whole. Physicists refer to this wave probability as "patterns of organic energy" (Zukav, 1979 p. 90). Zukav (1979) quotes Heisenberg's (1958) description of the impact of the probability wave in physics as meaning:

"...a tendency for something. It was a quantitative version of the old concept of "potentia" in Aristotelian philosophy. It introduced something standing in the middle between the idea of an event and the actual event a strange kind of physical reality just in the middle between possibility and reality" (p. 90).

Based on this brief argument which will be further discussed as part on my dissertation work, it is concluded that living organisms are organic and exist in a quantum reality. Our consciousness, through its connections with the collective whole is directly linked to these patterns of organic energy that are actively creating that which we know as reality and even that reality that exists as a potential. These potentials or probability waves are what David Bohm (1980), one of the most prominent theoretical physicists and a former associate of Einstein, describes as similar to that of a holograph--where any minute piece of the three dimensional image can produce the entire image. In this way, Bohm describes the universe as an "indivisible whole" where the "classical view of bit and pieces of building blocks has given way to the concept of patterns, process, and

interrelatedness" (Dossey, 1982, p. 104). Bohm (1980) refers to the existence of this indivisible whole "enfolded" in the universe in the form of movement. According to Bohm (1980) universal movement is found in electromagnetic waves, sound waves, electron beams and other form of movements, and all are an expression of "implicate order." Simply put, "implicate order" is encapsulated within the universe as infinite energy (Bohm 1980). This encapsulation of energy (movement) within the universe leads us to perceive a reality that is isolated and disconnected. Bohm (1980) claims that this is an illusion--a distortion "of the underlying behind-the scene oneness and unity, which is an intrinsic quality of the world" (Dossey, 1982, p. 104). This is analogous to the tapestry metaphor as put forth by Miller and Burbank (1994) which asks us to seek the interrelatedness and connections of things (threads) to help us understand the whole.

The above arguments extend the idea of consciousness as being something beyond simple electro-chemical activities in the brain (a reductionist perspective) to one where a collective consciousness is a causal agent in the creation of reality. To illustrate, Dossey (1982) quotes Nobel Prize winning physicist Eugene Wigner's (1967) ideas on consciousness as stating:

"... if the mind could not affect the physical world but was only effected by it, this would be the only known example in modern physics of such a one-way interaction. For in modern physics, one-way interactions are not known to occur (Dossey, 1982, p. 208).

The findings presented above illustrate the existence of a reciprocal interaction between consciousness and physical reality. Collective consciousness is seen to extend and permeate all that is organic. This interaction is then manifested as a collective reality. Dossey (1982) documents what Willis Harman (1991) describes as four characteristics of consciousness that are necessary to explain phenomenon that has been observed. They include the idea that mind is "spatially extended, mind is temporally extended, ultimately mind is predominant over the physical [in that the physical is the result of a the collective mind] and finally minds are joined" [to the collective whole] (p. 209).

To recapitulate, perhaps the circle is the place where seemingly opposites meet and reconnect with the underlying wholeness inherent within the universe. Perhaps within this blending we can establish a truth that explains both positions and have what Niels Bohr referred to as a "a great truth:" A truth where one truth explains and supports the other seemingly mutually exclusive position. From a quantum perspective it makes perfect sense that both positions are correct. We as humans merely need to alter our perspective to include that of wholeness. In support of this closing statement, Zohar and Marshall (1994) suggest that a quantum reality replaces the notion of opposites colliding and crashing into one another with one and state, "When two quantum systems meet their particles tend to stay somewhat separate and maintain shades of their original identities, while their wave aspects merge, giving rise to entirely new system that enfolds the originals" (p. 54). The parts have simply been allowed to represent a greater whole. This is the sort of integration our science needs. While there are still

many problems which need to be resolved within this inchoate presentation of our what our science has been, what it is, and what it could be, nonetheless, expansion on these ideas by scientists, educators and philosophers may result in greater congruence within our, theories, our philosophies, and our methods as suggested by many of those cited here. We, as scientists, could have a greater comprehensive corresponding truth of reality.

Based on this position put forth here, the question must be asked. What would my work look like if I were adhering to a paradigm such as the one presented above? First of all, greater consideration would have be given to my position as the observer, and an inclusion of a comprehensive understanding of the subjective experience of both myself and the youth's with whom I have been working. My interpretation of findings would include a prominent emphasis on the subjective or conscious experiences of youth rather than on the existence of differences within and between groups as expressed on standardized measures. In essence, I would be looking for subjective relationships with regard to feelings of all youths. Then compare how those feelings are experienced differently for different groups of youths. What are the connections they share? For instance, could we speculate that by completing 95% of all interviews myself and following these youth for four years, may have resulted in a bond being established between them and I? Could this bond have effected the ways in which they would respond? I think yes. But I am also left "uncertain" as to what that effect might be. Could it be that they wanted to present themselves in the best light so that they underreported their problem behaviors? Or was their reporting more honest

because we had established trust in each other, knowing, through history, that what they shared with me had remained confidential? I think the latter. However, without including this connection within my investigation a complete interpretation of this effect is not possible.

Incorporating qualitative material beyond the mostly structured interview may have allowed for a greater understanding of those things that youth feared, were happy with, confused about, or sorry for. Outcome variables would have been associated with subjective aspects of a comprehensive sense of well-being. Though this was attempted through the use of an outcome measure based on subjective reporting of feelings of worth, it was still primarily quantitatively driven. Multiple measures of subjective outcomes including both quantitative and qualitative measures would have been more appropriate. This expansion of evidence would have allowed a more comprehensive picture to emerge.

I would have also have investigated feelings of connectedness and belonging in a more direct way. If we as beings are fundamentally connected with all that is--yet we do not experience that connectedness as something salient--can reengagement in feelings of connectedness result in better outcomes? Is the reason that Alaskan youth are reporting greater problem behaviors and lower levels of self-esteem the result of being both physically and spiritually disconnected from tradition and cultural beliefs resulting in delinquent behavior? As Harman (1991) explains, when working with indigenous people a Western ontological perspective of separateness and autonomy may be in direct opposition to a Native people's worldview involving connectedness and wholeness. As

scientists we need to utilize methods that allow for the inclusion of diverse perspectives especially when defining outcome. In a Western, White, male, worldview, successful adolescent development is defined by a break in primary relationships resulting in an independent and autonomous being. This "norm" may be supported in a society for White males, but for females and other alternative populations, this may not be considered positive emotional health. Alaskan Natives suffer greatly when separated from their Villages. To have their success based on their ability to detach from their ways of knowing is an attack on their spirit and an invalidation of who they are. This particular thread was evident throughout my work.

Finally, if a collective consciousness has a downward causation (evidenced in a material reality) as suggested by Sperry and Harman (1991), then instead of utilizing a model of development where one-size-fits-all based on separating and autonomizing self from others--from reality--from Nature, we could focus on treatments that would work within this downward causation paradigm and begin changing feelings of worth, as well as focusing on reestablishing feelings of connectedness, to one's culture, family, and traditions that are salient to the individual. The inclusion of consciousness in this way of knowing is essential. If "downward causation" exists, then problem youth should be able to play a prominent role in their own "righting." We should be supporting the maintenance of relationships and responsibility of self and nature as critical indices of health, instead of defining health as the mere absence of delinquent behavior. We must be investigating how disconnections from self,

others, and nature are impacting the world of our youth. If, as human beings, our infinite self resides in a collective whole, then what ills have been produced by adhering to Western Eurocentric bias based on objective detachment and exploitation?

In conclusion, it is proposed here that most mainstream philosophies associated with a post-positivist perspective and their inherent methodologies are compressed to the point that integration of complex experience (consciousness and subjectivity) cannot be included, let alone, adequately measured. Our beliefs as well as our methods need to be supplemented, "unfolded and expanded" with encompassing ontological perspectives and statistical procedures that will embrace qualitative and contextual methods such as those found in feminist and/or multicultural psychology, and nursing sciences. Miller & Burbank (1994) inform us that we need to establish congruence within our theory, our philosophies, and our methods. They go on to contend that as scientists, we should attempt to establish truth in such a way that it expands our existing knowledge base. This expansion should result in the incorporation of multiple perspectives resulting in a broadening of method and application that would not have been possible without the integration. Essentially they are talking about an epistemological revolution within our sciences.

The question remains to be answered. How can we incorporate multiple perspectives, such as those found residing on opposite sides of the ontological continuum, and be left with congruence? The work that needs to be done is to somehow interact with these two differing scientific paradigms (connections and

separateness; objectivity and consciousness) in such a way that allows one philosophical position to encompass the position of both. Through the guidance of my committee members as well as my professors, it is hoped that future work will help to further our understanding of expanded scientific worldview. As suggested by Dr Burbank, perhaps the answer resides within the circle. We are not alone in our search for "truth" as evidenced by all those authors cited here, but the work must continue even in the face of opposition. Let us think of the words spoken by Machiavelli as he attempted to effect change:

"It must be considered that there is nothing more difficult to carry out, nor more doubtful of success, nor more dangerous to handle than to initiate a new order of things. For the reformer has enemies [adversaries] in all those who profit by the old order, and only lukewarm defenders in all those who would profit by the new order, this lukewarmness arising partly from fear of their adversaries who have laws [tradition] in their favor; and partly [from the] incredulity of man[human]kind, who do not truly believe in anything new until they have had actual experience of it."

Machiavelli, *The Prince*

I'm finally done ☺, Thank you for being my audience and my guides.

References

- Achenbach, T. M. (1991). Manual for the Youth Self-Report and 1991 profile. Burlington Vermont: University of Vermont Department of Psychiatry.
- Asarnow, J. R., Aoki, W., & Elson, S. (1996). Children in residential treatment: a follow up study. Journal of Clinical Child Psychology, 25 (2), 209-214.
- Blackman, M. Eustace, J. & Chowdhury, T. (1991). Adolescent residential treatment: A one to three year follow-up. Canadian Journal of Psychiatry 36, (7), 472-479.
- Bohm, D. (1980). Wholeness and implicate order. London: Routledge & Kegan Paul.
- Brooks, R. B. (1994). Children at risk: Fostering resilience and hope. American Journal of Orthopsychiatry, 64, (4), 545-553.
- Burbank, P. (1999). NUR660 fall semester: Philosophical foundations of health care research. Department of Nursing, University of Rhode Island.
- Burks, J. (1995). Edgewood childrens' center outcome study 1991/1992. Residential Treatment for Children and Youth, 13 (2), 31-40.
- Burnett, R., & Rathbun, C. (1993). Discovery and treatment of adolescent sexual offenders in a residential treatment center. Residential Treatment for Children and Youth, 11 (2), 57-64.
- Burt, H. (1993). Issues in art therapy with the culturally displaced American Indian youth. Arts In Psychotherapy, 20 (2), 143-151.
- Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. Journal of Abnormal and Social Psychology, 65, 145-153.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences: Second edition. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cook, T. D., Cooper, H., Cordray, D. S., Hartmann, H., Hedges, L. V., Light, R. J., Louis, T. A., & Mostelller, F. (1994). Meta-Analysis for explanation: A casebook. New York: Russell Sage Publications.

Cooper, H., & Hedges, L. V. (1994). The handbook of research synthesis. New York: Russell Sage Publications.

Curry, J. (1991). Outcome research on residential treatment: Implications and suggested directions. American Journal of Orthopsychiatry, 61 (3), 348-357

Deschenes, E. P., & Greenwood, P. W. (1998). Alternative placement for juvenile offenders: Results from the evaluation of the Nokomis challenge program. Journal of research in Crime and Delinquency, 35 (3), 267-293.

Duchnowski, A., Johnson, M. K., Hall, K. S., & Kutash, K. (1993). The alternatives to residential treatment study: Initial findings. Journal of Emotional and Behavioral Disorders, 1 (1), 17-26.

Dukes, R. L., & Martinez, R. (1994). The impact of ethnicity on self-esteem among adolescents. Adolescence, 29 (113), 105-114.

Enns, C. Z. (1994). On teaching about the cultural relativism of psychological constructs. Teaching of Psychology, 21 (4), 205-211.

Epps, K. (1991). The residential treatment of adolescent sex offenders. Issues in Criminology and Legal Psychology, 1 (17), 58-67.

Fisher, P. A., Storck, M., Bacon, J. G. (1999). In the eye of the beholder: Risk and protective factors in rural American Indian and Caucasian adolescents. American Journal of Orthopsychiatry, 69 (3) 294 - 304.

Friedman, S. S. (1995). Beyond White and Other: Relationality and narratives of race in feminist discourse. Signs, 21 496-503.

Gutierrez, S. E., Felipe-Russo, N., & Urbanski, L. (1994). Sociocultural and psychological factors in American Indian drug use: Implications for treatment. The International Journal for Addictions, 29 (14), 1761-1786.

Gutierrez, S. E., & Todd, M. (1997). The impact of childhood abuse on treatment outcomes of substance users. Professional Psychology Research and Practice, 28 (4), 348-354.

Guba, E. G. (1990). The paradigm dialogue. Sage Publications.

Harlow, L. L., Muliak, S. A., & Steiger, J. H. (1997). What if there were no significant tests? Mahwah, New Jersey: Lawrence Erlbaum Associates.

Harlow, L. L. (2000). Psychology 533 spring semester: Multivariate methods in psychology coursepack. Department of Psychology, University of Rhode Island.

Harman, W. W. (1991). A re-examination of the metaphysical foundations of modern science: Causality issues in contemporary science. Sausalito, California: Institute of Noetic Sciences.

Hiller, M., L. (1996). Effectiveness of a specialized women's intervention in a residential treatment program. Substance Use and Misuse, 31 (6), 771-783.

Hsu, D. Y. F. (1995). Internalize culture, culturocentrism, and transcendence. Counseling Psychologist, 23, 4-24.

- Hodges, K. (1990). Child and adolescent functional assessment scale. Vanderbilt Children Health Services Evaluation Project.
- Hunt, M. (1997). How science takes stock. New York: Russell Sage Publications.
- Kaplan, G. A., Gleason N., & Klein, R. Women's self-development in late adolescence. In Women's growth in connection, 1991. New York: The Guilford Press.
- Keller, F. E., & Longino, H. E. (1996). Feminism and science. New York: Oxford University Press.
- Klemke, E. D., Hollinger, R., Wyss Rudge, D., & Kline, D. (1998). Introductory readings in the philosophy of science. Amherst, New York: Prometheus Books.
- Kuhn, T. S. (1970). The structure of scientific revolutions. International Encyclopedia of Unified Science 2nd Edition. The University of Chicago Press.
- Landrine, H., Klonoff, E. A., & Brown-Collins, A. (1992). Cultural Diversity and methodology in feminist psychology: Critique, proposal, and empirical example. Psychology of Women Quarterly, 16, 145-163.
- Landrine, H. (1995). Introduction: Cultural diversity, contextualism, and feminist psychology. In H Landrine (Ed) Bringing cultural diversity to feminist psychology.
- Lott, B. 1994. Women's lives: Themes and variations in gender learning. Second Edition. Pacific Grove: Brooks/Cole Publishing.
- Lipsey, M. W. (1990). Design sensitivity: Statistical power for experimental research. Newbury Park, California: Sage Publications.
- Maddock, J. E. (1999). Statistical power and effect size in the Field of Health Psychology. Unpublished doctoral dissertation, University of Rhode Island.

Meehl, P. E. (1978). Theoretical Risks and Tabular Asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. Journal of Counseling and Clinical Psychology, 46, 806-834.

Mennen, F. E., & Meadow, D. (1993). The relationship of sexual abuse to symptom levels in emotionally disturbed girls. Child and Adolescent Social Work Journal, 10 (4), 319-328.

Miller, A. (1986). Thou shall not be aware: Societies betrayal of the child. New York, New York: Meridian Publishers.

Miller, D., Trapani, C., Fejes-Mendoza, K., Eggleston, C., & Dwiggins, D. (1995). Adolescence 30, (118), 429-435.

Miller, J. B. (1976). Toward a New Psychology for Women. Boston: Beacon Press.

Miller, J. R., & Burbank, P. M. (1994). Congruence among philosophy, theory, and method in nursing research. In workgroup of European nurse researchers, The contribution of nursing research: Past-Present-Future, vol. 2, Proceedings at the 7th Biennial Conference, Oslo, Norway: Falch Hurtigtrykk, 704-709.

Mosedale, F. E. (1979). Philosophy of science: The wide range of interaction. New Jersey: Prentice Hall.

Nobel, D. N., & Gibson, D. (1994). Family values in action: Family connectedness for children in substitute care. Child and Youth Care Forum, 23 (5), 315-328.

O'Mally, F. (1996). Critical issues in short-term residential treatment. Residential Treatment for Children and Youth, 14 (1), 61-74.

- Packard, S. A. & Polifroni, C. E. (1990). The nature of scientific truth. Nursing Science Quarterly, 5 (4), winter. 158-163.
- Pfeiffer, S. I., & Strzelecki, S. C. (1990). Inpatient psychiatric treatment of children and adolescents: A review of outcome studies. Journal of the American Academy of Child and Adolescent Psychiatry, 29 (6), 847-853.
- Ponce, D. E. (1995). Value orientation: Clinical applications in a multi-culture residential treatment center for children and youth. Residential Treatment for Child and Youth, 12 (4), 29-42.
- Pratt, S. I., & Moreland, K. L. (1996). Introduction to treatment outcome: Historical perspective and current issues. Residential Treatment for Child and Youth, 13 (4), 1996, 1-27.
- Rosenberg, M. (1965). Society and the Adolescent Self Image. Princeton, New Jersey: Princeton University Press.
- Rosenthal, R. (1979). The "file drawer" problem and tolerance for null results. Psychological Bulletin, 86, 638-641.
- Rosenthal, R. (1991). Meta-analytic procedures for social research. Newbury Park: Sage Publications.
- Rosnow, R. L. & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. American Psychologist, 44, 1276-1284.
- Rossi, J. S. (1990). Statistical power in psychological research. What have we gained in 20 years? Journal of Consulting and Clinical Psychology, 58, 1276-1284.

- Sagan, C. (1996). The demon haunted world: Science as a candle in the dark. Random House Publishers.
- Scheidlinger, S. (1992). Therapist gender in child-adolescent treatment groups. Journal of Child and Adolescent Group Therapy, 2 (2), 104-108.
- Small, R., Kennedy, K., & Bender, B. (1991). Critical issues for practice in residential treatment: The view from within. American Journal of Orthopsychiatry, 61 (3), 327-337.
- Stern, L. (1991) Conceptions of separation and connection in female adolescents. In C. Gilligan, C., Lyons, N., & Hammer, T. (Eds.), Making Connections: The Relational World of Female Adolescents at Emma Willard School (p. 73-87). Cambridge: Harvard University Press.
- Surrey, J. L. (1991). The self-in-relation: A theory of women's development. In Women's Growth and Development, 1991. New York: The Guilford Press.
- Tabachnick, B. G., & Fidell, L., S. (1996). Using multivariate statistics: Third edition. New York: HarperCollins Publishers.
- Tafarodi, R. W., & Swann, W. B. (1996). Individualism-collectivism and global self-esteem. Journal of Cross Cultural Psychology, 27 (6), 651-672.
- Unger, R., & Crawford, M. (1996). Women and gender: A feminist psychology. United States: McGraw Publishing.
- US Department of Health and Human Services: Center for Substance Abuse Prevention (1993). Measurement in prevention: A manual on selecting and using instruments to evaluate prevention programs. DHHS Publication No. (SMA) 93-2041.

Valliant, P. (1993). Cognitive and behavioral therapy with adolescent males in residential treatment care. Journal of Child and Youth Care, 8 (3), 41-49.

Wheelis, A. (1971). The end of the modern age. New York: Harper & Row.

Wells, K. (1991). Long-term residential treatment for children: Introduction. American Journal of Orthopsychiatry, 61 (3), 324-326.

Wright, N. (1995). Social skills training for conduct- disorder boys in residential treatment. Residential Treatment for Children and Youth, 12 (4), 15-28.

Yandow, V. (Video, 1990). Alcohol and drug problems in women. The Center for Interdisciplinary Studies: Battelboro, Vermont.

Zimmerman, P., D. (1996). A comparison of commonly used treatment measures. Residential Treatment for Children and Youth, 13 (4), 49-69.

Zohar, D., & Marshall, I. (1994). The quantum society. New York: Morrow and Company Inc

Zukav, G. (1979). The dancing wu li masters: An overview of the new physics. New York.

Appendix A

Achenbach Youth Self-Report

YOUTH SELF-REPORT FOR AGES 11-18

For others use only
ID #

YOUR NAME			PARENTS' USUAL TYPE OF WORK, even if not working now (Please be specific—for example, auto mechanic, high school teacher, homemaker, laborer, lathe operator, shoe salesman, army sergeant.) FATHER'S TYPE OF WORK: _____ MOTHER'S TYPE OF WORK: _____ Please fill out this form to reflect <i>your</i> views, even if other people might not agree. Feel free to write additional comments beside each item and in the spaces provided on pages 2 and 4.
YOUR SEX ☐ Boy ☐ Girl	YOUR AGE	ETHNIC GROUP OR RACE	
TODAY'S DATE Mo _____ Day _____ Yr _____		YOUR BIRTHDATE Mo _____ Day _____ Yr _____	
GRADE IN SCHOOL NOT ATTENDING SCHOOL ☐	IF YOU ARE WORKING, STATE TYPE OF WORK		

- I. Please list the sports you most like to take part in. For example: swimming, baseball, skating, skate boarding, bike riding, fishing, etc.

☐ None

a. _____
b. _____
c. _____

Compared to others of your age, about how much time do you spend in each?

Less Than Average	Average	More Than Average
☐	☐	☐
☐	☐	☐
☐	☐	☐

Compared to others of your age, how well do you do each one?

Below Average	Average	Above Average
☐	☐	☐
☐	☐	☐
☐	☐	☐

- II. Please list your favorite hobbies, activities, and games, other than sports. For example: cards, books, piano, autos, crafts, etc. (Do not include listening to radio or TV.)

☐ None

a. _____
b. _____
c. _____

Compared to others of your age, about how much time do you spend in each?

Less Than Average	Average	More Than Average
☐	☐	☐
☐	☐	☐
☐	☐	☐

Compared to others of your age, how well do you do each one?

Below Average	Average	Above Average
☐	☐	☐
☐	☐	☐
☐	☐	☐

- III. Please list any organizations, clubs, teams or groups you belong to.

☐ None

a. _____
b. _____
c. _____

Compared to others of your age, how active are you in each?

Less Active	Average	More Active
☐	☐	☐
☐	☐	☐
☐	☐	☐

- IV. Please list any jobs or chores you have. For example: Paper route, babysitting, making bed, working in store, etc. (Include both paid and unpaid jobs and chores.)

☐ None

a. _____
b. _____
c. _____

Compared to others of your age, how well do you carry them out?

Below Average	Average	Above Average
☐	☐	☐
☐	☐	☐
☐	☐	☐

- V. 1. About how many close friends do you have? ☐ None ☐ 1 ☐ 2 or 3 ☐ 4 or more
(Do not include brothers & sisters)
2. About how many times a week do you do things with any friends outside of regular school hours?
(Do not include brothers & sisters) ☐ less than 1 ☐ 1 or 2 ☐ 3 or more

VI. Compared to others of your age, how well do you:

	Worse	About the same	Better	
a. Get along with your brothers & sisters?	☐	☐	☐	☐ I have no brothers or sisters
b. Get along with other kids?	☐	☐	☐	
c. Get along with your parents?	☐	☐	☐	
d. Do things by yourself?	☐	☐	☐	

VII. Performance in academic subjects. ☐ I do not go to school because _____

	Failing	Below Average	Average	Above Average
a. English or Language Arts	☐	☐	☐	☐
b. History or Social Studies	☐	☐	☐	☐
c. Arithmetic or Math	☐	☐	☐	☐
d. Science	☐	☐	☐	☐
Other academic subjects – for example: computer courses, foreign language, business. Do not include gym, shop, driver's ed., etc.	☐	☐	☐	☐
e. _____	☐	☐	☐	☐
f. _____	☐	☐	☐	☐
g. _____	☐	☐	☐	☐

Do you have any illness, physical disability, or handicap? ☐ No ☐ Yes—please describe

Please describe any concerns or problems you have about school:

Please describe any other concerns you have:

Please describe the best things about yourself:

very true or often true of you. Circle the 1 if the item is somewhat or sometimes true of you. If the item is not true of you, circle the 0.

0 = Not True 1 = Somewhat or Sometimes True 2 = Very True or Often True

0 1 2 1. I act too young for my age
0 1 2 2. I have an allergy (describe): _____

0 1 2 3. I argue a lot
0 1 2 4. I have asthma
0 1 2 5. I act like the opposite sex
0 1 2 6. I like animals
0 1 2 7. I brag
0 1 2 8. I have trouble concentrating
or paying attention
0 1 2 9. I can't get my mind off certain thoughts
(describe): _____

0 1 2 10. I have trouble sitting still
0 1 2 11. I'm too dependent on adults
0 1 2 12. I feel lonely
0 1 2 13. I feel confused or in a fog
0 1 2 14. I cry a lot
0 1 2 15. I am pretty honest
0 1 2 16. I am mean to others
0 1 2 17. I daydream a lot
0 1 2 18. I deliberately try to hurt or kill myself
0 1 2 19. I try to get a lot of attention
0 1 2 20. I destroy my own things
0 1 2 21. I destroy things belonging to others
0 1 2 22. I disobey my parents
0 1 2 23. I disobey at school
0 1 2 24. I don't eat as well as I should
0 1 2 25. I don't get along with other kids
0 1 2 26. I don't feel guilty after doing
something I shouldn't
0 1 2 27. I am jealous of others
0 1 2 28. I am willing to help others
when they need help
0 1 2 29. I am afraid of certain animals, situations,
or places, other than school
(describe): _____

0 1 2 30. I am afraid of going to school
0 1 2 31. I am afraid I might think or
do something bad
0 1 2 32. I feel that I have to be perfect
0 1 2 33. I feel that no one loves me
0 1 2 34. I feel that others are out to get me
0 1 2 35. I feel worthless or inferior
0 1 2 36. I accidentally get hurt a lot
0 1 2 37. I get in many fights
0 1 2 38. I get teased a lot
0 1 2 39. I hang around with kids who get in trouble

0 1 2 40. I hear sounds or voices that other people
think aren't there (describe): _____

0 1 2 41. I act without stopping to think
0 1 2 42. I would rather be alone than with others
0 1 2 43. I lie or cheat
0 1 2 44. I bite my fingernails
0 1 2 45. I am nervous or tense
0 1 2 46. Parts of my body twitch or
make nervous movements (describe): _____

0 1 2 47. I have nightmares
0 1 2 48. I am not liked by other kids
0 1 2 49. I can do certain things better
than most kids
0 1 2 50. I am too fearful or anxious
0 1 2 51. I feel dizzy
0 1 2 52. I feel too guilty
0 1 2 53. I eat too much
0 1 2 54. I feel overtired
0 1 2 55. I am overweight
56. Physical problems without known medical
cause:
0 1 2 a. Aches or pains (not headaches)
0 1 2 b. Headaches
0 1 2 c. Nausea, feel sick
0 1 2 d. Problems with eyes (describe): _____

0 1 2 e. Rashes or other skin problems
0 1 2 f. Stomachaches or cramps
0 1 2 g. Vomiting, throwing up
0 1 2 h. Other (describe): _____

0 1 2 57. I physically attack people
0 1 2 58. I pick my skin or other parts of my body
(describe): _____

0 1 2 59. I can be pretty friendly
0 1 2 60. I like to try new things
0 1 2 61. My school work is poor
0 1 2 62. I am poorly coordinated or clumsy
0 1 2 63. I would rather be with older
kids than with kids my own age

- 0 1 2 64. I would rather be with younger kids than with kids my own age
- 0 1 2 65. I refuse to talk
- 0 1 2 66. I repeat certain actions over and over (describe): _____

- 0 1 2 67. I run away from home
- 0 1 2 68. I scream a lot
- 0 1 2 69. I am secretive or keep things to myself
- 0 1 2 70. I see things that other people think aren't there (describe): _____

- 0 1 2 71. I am self-conscious or easily embarrassed
- 0 1 2 72. I set fires
- 0 1 2 73. I can work well with my hands
- 0 1 2 74. I show off or clown
- 0 1 2 75. I am shy
- 0 1 2 76. I sleep less than most kids
- 0 1 2 77. I sleep more than most kids during day and/or night (describe): _____

- 0 1 2 78. I have a good imagination
- 0 1 2 79. I have a speech problem (describe): _____

- 0 1 2 80. I stand up for my rights
- 0 1 2 81. I steal at home
- 0 1 2 82. I steal from places other than home

- 0 1 2 83. I store up things I don't need (describe): _____

- 0 1 2 84. I do things other people think are strange (describe): _____

- 0 1 2 85. I have thoughts that other people would think are strange (describe): _____

- 0 1 2 86. I am stubborn
- 0 1 2 87. My moods or feelings change suddenly
- 0 1 2 88. I enjoy being with other people
- 0 1 2 89. I am suspicious

- 0 1 2 90. I swear or use dirty language
- 0 1 2 91. I think about killing myself
- 0 1 2 92. I like to make others laugh

- 0 1 2 93. I talk too much
- 0 1 2 94. I tease others a lot

- 0 1 2 95. I have a hot temper

- 0 1 2 96. I think about sex too much

- 0 1 2 97. I threaten to hurt people

- 0 1 2 98. I like to help others

- 0 1 2 99. I am too concerned about being neat or clean

- 0 1 2 100. I have trouble sleeping (describe): _____

- 0 1 2 101. I cut classes or skip school

- 0 1 2 102. I don't have much energy

- 0 1 2 103. I am unhappy, sad, or depressed

- 0 1 2 104. I am louder than other kids

- 0 1 2 105. I use alcohol or drugs for nonmedical purposes (describe): _____

- 0 1 2 106. I try to be fair to others

- 0 1 2 107. I enjoy a good joke

- 0 1 2 108. I like to take life easy

- 0 1 2 109. I try to help other people when I can

- 0 1 2 110. I wish I were of the opposite sex

- 0 1 2 111. I keep from getting involved with others

- 0 1 2 112. I worry a lot

Please write down anything else that describes your feelings, behavior, or interests

Appendix B

The Family Adaptability, Cohesion, and Expressiveness Scale

FACES III

FACES III* (Actual)

Now I'd like to ask you some questions about your family. I'm going to read activities that families sometimes do. Then I'll ask you to tell me how often YOUR family does this activity. For example, "Family members ask each other for help." SHOW CARD. Would you say that this happens in your family: almost never, once in awhile, sometimes, frequently, or almost always? RECORD RESPONSE AND CONTINUE.

	Almost never	Once in awhile	Some- times	Frequ- ently	Almost always
1. Family members ask each other for help.....	1	2	3	4	5
2. In solving problems, the children's suggestions are followed.....	1	2	3	4	5
3. We approve of each other's friends.....	1	2	3	4	5
4. Children have a say in their discipline.....	1	2	3	4	5
5. We like to do things with just our immediate family.....	1	2	3	4	5
6. Different persons act as leaders in our family.	1	2	3	4	5
7. Family members feel closer to other family members than to people outside the family....	1	2	3	4	5
8. Our family changes its way of handling tasks...	1	2	3	4	5
9. Family members like to spend free time with each other.....	1	2	3	4	5
10. Parent(s) and children discuss punishment together.....	1	2	3	4	5
11. Family members feel very close to each other...	1	2	3	4	5
12. The children make the decisions in our family..	1	2	3	4	5
13. When our family gets together for activities, everybody is present.....	1	2	3	4	5
14. Rules change in our family.....	1	2	3	4	5
15. We can easily think of things to do together as a family.....	1	2	3	4	5
16. We shift household responsibilities from person to person.....	1	2	3	4	5
17. Family members consult other family members on their decisions.....	1	2	3	4	5
18. It is hard to identify the leader(s) in our family.....	1	2	3	4	5
19. Family togetherness is very important.....	1	2	3	4	5
20. It is hard to tell who does which household chores.....	1	2	3	4	5

* Olson, D.H., Portner, J., & Lavee, Y. (1985)

Appendix C

Rosenberg Self-Esteem Test

SELF-ESTEEM TEST^{*}

I am now going to ask you several questions about how you feel about yourself. After I read each statement, please tell me whether you: strongly agree, agree, are not sure, disagree, or strongly disagree with it. SHOW CARD.

	STRONGLY AGREE	AGREE	NOT SURE	DISAGREE	STRONGLY DISAGREE
1. I feel that I'm a person of worth, at least on an equal basis with others.	1	2	3	4	5
2. I feel that I have a number of good qualities.	1	2	3	4	5
3. All in all, I am inclined to feel that I am a failure.	1	2	3	4	5
4. I am able to do things as well as most other people.	1	2	3	4	5
5. I feel I do not have much to be proud of.	1	2	3	4	5
6. I take a positive attitude toward myself.	1	2	3	4	5
7. On the whole, I am satisfied with myself.	1	2	3	4	5
8. I wish I could have more respect for myself.	1	2	3	4	5
9. I certainly feel useless at times.	1	2	3	4	5
10. At times I think I am no good at all.	1	2	3	4	5

^{*} Rosenberg, M. (1965)

Appendix D

All References Retrieved for Meta-Analysis

Appendix D

Articles Retrieved for Meta-Analysis

Abbey, A. A., Nicholas, K. B., & Bieber, S. L. (1997). Predicting runaways upon admission to an adolescent treatment center. Residential Treatment for Children and Youth, 15 (2), 73-85.

Anderson, K. (1998). A Canadian child welfare agency for urban natives: The clients speak. Child Welfare League of America, LXXVII (4), 441-460.

Ansari, A. A., Gouthro, S., Ahmad, K., & Steele, C. (1996). Hospital-based behavior modification program for adolescents: Evaluation and predictors of outcome. Adolescence, 31 (122), 469-476.

Bautista de Domanico, Y., Crawford, I., & De Wolfe, A. S. (1994). Ethnic identity and self-concept in Mexican-American adolescents: Is bicultural identity related to stress or better adjustment? Child and Youth Care Forum, 23 (3), 197-206.

Beker, J. (1999). Community-based residential child and youth care work: Opportunity or oxymoron? Child and Youth Care Forum, 28 (4), 257-259.

Beker, J., & Magnuson, D. (1996). Residential education in the United States: Some questions, some suggestions, and a look ahead. Residential Treatment for Children and Youth, 13 (3), 117-124.

Bochner, S. (1994). Cross-cultural differences in the self-concept: A Test of Hofstede's individualism/collectivism distinction. Journal of Cross-Cultural Psychology, 25 (2), 273-283.

Brewer, V. E., & Smith, M. D. (1995). Gender inequality and rates of female homicide victimization across U.S. cities. Journal of Crime and Delinquency, 32 (2), 175-190.

Brooks, R. B. (1994). Children at risk: Fostering resilience and hope. American Journal of Orthopsychiatry, 64 (4), 545-553.

Brown, A. J. (1996). The meaning of respect: A First Nations perspective. Canadian Journal of Nursing Research (4) , 95-109.

Brown, R. A., & Hill, B. A. (1996). Opportunity for change: Exploring an alternative to residential treatment. Child Welfare League of America, LXXV (1), 35-56.

Burks, J. (1995). Edgewood Childrens' Center outcome study 1991/1992. Residential Treatment for Children and Youth, 13 (2), 31-60.

Burt, H., & Dtati, A. T. R. (1993). Issues in art therapy with the culturally displaced American Indian Youth. The Arts in Psychotherapy, 20, 143-151.

Burton, V. S., Jr., Cullen, F. T., Evans, T. D., Alarid, L. F., & Dunaway, R. G. (1998). Gender, self-control, and crime. Journal of Research in Crime and Delinquency, 35 (2), 123-147.

Butler, S. M., Mackay, S. A., & Dickens, S. E. (1995). Maternal and adolescent ratings of psychopathology in young offender and non-clinical males. Canadian Journal of Nursing Research, (3), 333-342.

Carlson, B. E., Barr, W. B., & Young, K. J. (1994). Factors associated with treatment outcomes of male adolescents. Residential Treatment for Children and Youth, 12 (1), 39-58.

Clarke, H. F. (1997). Research in nursing and cultural diversity: Working with First Nations people. Canadian Journal of Nursing Research, (2), 11-25.

Conley, D. J. (1994). Adding color to a black and white picture: Using qualitative data to explain racial disproportionality in the juvenile justice system. Journal of Research in Crime and Delinquency, 31 (2), 135-148.

Cozens, W. R. (1996). Outcome measurement: Who ordered it, anyway? Residential Treatment for Children and Youth, 14 (2), 45-59.

Cozens, W. R. (1998). Why do you collect data? Residential Treatment for Children and Youth, 15 (4), 25-43.

Curry, J. F. (1995). The current status of research in residential treatment. Residential Treatment for Children and Youth, 12 (3), 1-17.

Daly, D. L., Schmidt, M. D., Spellman, D. F., Criste, T. R., Dinges, K., & Teare, J. F. (1998). The Boys Town Residential Treatment Center: Treatment implementation and preliminary outcomes. Child and Youth Care Forum, 27 (4), 267-279.

Day, D. M., Pal, A., & Goldberg, K. (1994). Assessing the post-residential functioning of latency-aged conduct disordered children. Residential Treatment for Children and Youth, 11 (3), 45-61.

De Groot, A., Koot, H. M., & Verhulst, F. C. (1996). Cross-cultural generalizability of the youth self-report and teacher's report form cross-informant syndromes. Journal of Abnormal Child Psychology, 24 (5), 651-665.

Dery, M., Toupin, J., Pauze, R., Mercier, H., & Fortin, L. (1999). Neuropsychological characteristics of adolescents with conduct disorder:

Association with attention-deficit-hyperactivity and aggression. Journal of Abnormal Child Psychology, 27 (3), 225-236.

Deschenes, E. P., & Greenwood, P. W. (1998). Alternative placements for juvenile offenders: Results from the evaluation of the Nokomis Challenge Program. Journal of Research in Crime and Delinquency, 35 (3), 267-294.

Dukes, R. L., & Martinez, R. (1994). The impact of ethnicity on self-esteem among adolescents. Adolescence, 29 (113), 105-115.

Eagle, R. S. (1994). The separation experience of children in long-term care: Theory, research, and implications for practice. American Journal of Orthopsychiatry, 64 (3), 421-431.

Edwards, J. K. (1994). Children in residential treatment: How many, what kind? Do we really know? Residential Treatment for Children and Youth, 12 (1), 85-99.

English, R. A. (1996). Some implications of the African-American experience for the development of residential education in the United States. Residential Treatment for Children and Youth, 13 (3), 63-67.

Fisher, P. A., Starck, M., & Bacon, J. G. (1999). In the eye of the beholder: Risk and protective factors in rural American Indians and Caucasian adolescents. American Journal of Orthopsychiatry, 69 (3), 294-304.

Green, R. S., & Newman, F. L. (1996). Criteria for selecting instruments to assess treatment outcomes. Residential Treatment for Children and Youth, 13 (4), 29-45.

Halverson, A. M. (1995). The importance of caring and attachment in direct practice with adolescents. Child and Youth Care Forum, 24 (3), 169-173.

Howe, T. R., Tepper, F. L., & Parke, R. D. (1998). The emotional understanding an peer relations of abused children in residential treatment. Residential Treatment for Children and Youth, 15 (3), 69-81.

Johnson, M. M. (1999). Multiple dimensions of family centered practice in residential group care: Implications regarding the roles of stakeholders. Child and Youth Care Forum, 28 (2), 123-131.

Johnson, M. M. (1999). Managing perceptions: A new paradigm for residential group care. Child and Youth Care Forum, 28 (3), 165-178.

Kapp, S. A., Schwartz, I., & Epstein, I. (1994). Adult imprisonment of males released from a residential childcare: A longitudinal study. Residential Treatment for Children and Youth, 12 (2), 19-35.

Kashima, Y. (1995). Introduction to the special section on culture and self. Journal of Cross-Cultural Psychology, 26 (6), 603-605.

Kelly, T. D., & Thorndike-Christ, T. (1997). Self-concept in residential treatment facilities. Residential Treatment for Children and Youth, 15 (1), 11-25.

Knorth, E. J., & Eldering, L. (1998). Immigrant adolescents in residential group care and treatment settings: Research and experience in the Netherlands. Child and Youth Care Forum, 27 (4), 237-251.

Kuperminc, G. P., Allen, J. P., & Arthur, M. W. (1996). Autonomy, relatedness, and male adolescent delinquency: Toward a multidimensional view of social competence. Journal of Adolescent Research, 11 (4), 397-420.

Larson, J. D., Calamari, J. E., West, J. G., & Frevert, T. A. (1998).

Aggression management with disruptive adolescents in the residential setting: Integration of a cognitive-behavioral component. Residential Treatment for Children and Youth, 15 (4), 1 –9.

Lynse, M., & Levy G. D. (1997). Differences in ethnic identity in Native American adolescents as a function of school context. Journal of Adolescent Research, 12 (3), 372-388.

MacEachron, A. E. (1994). Supervision in tribal and state child welfare agencies: Professionalization, responsibilities, training needs, and satisfaction. Child Welfare League of America, LXXIII (2), 117-128.

MacNeil, G. (1998). Social competence among adolescents in residential settings: Validation of an empirically-based inventory. Residential Treatment for Children and Youth, 16 (1), 79-89.

Mannes, M. (1995). Factors and events leading to the passage of the Indian Child Welfare Act. Child Welfare League of America, LXXIV (1), 265-282.

McCord, D. M. (1995). Toward a typology of wilderness-based residential treatment program participants. Residential Treatment for Children and Youth, 12 (4), 51-59.

Mears, D. P., Ploeger, M., & Warr, M. (1998). Explaining the gender gap in delinquency: Peer influence and moral evaluations of behavior. Journal of Research in Crime and Delinquency, 35 (3), 251-266.

Miceli, D. M. (1997). Qualitative analysis of emotionally disturbed boys in residential summer camp. Residential Treatment for Children and Youth, 14 (3), 29-79.

Miller, D., Trapani, C., Fejes-Mendoza, K., Eggleston, C., & Dwiggins, D. (1995). Adolescent female offenders: Unique considerations. Adolescence, 30 (118), 429-435.

Mordock, J. B. (1994). The search for and identity: A call for observational-inductive research methods in residential treatment. Residential Treatment for Children and Youth, 12 (1), 1-23.

Nansel, T. R., Raines, S., Jackson, D. L., Teal, C. R., Force, R. C., Klingporn, M. J., & Burdsal, C. A. (1998). A survey of residential treatment centers' outcome research practices. Residential Treatment for Children and Youth, 15 (4), 45-58.

Noble, D. N., & Gibson, D. (1994). Family values in action: Family connectedness for children in substitute care. Child and Youth Care Forum, 23 (5), 315-329.

Noss, J. (1998). Physiological implications for the treatment of children in residential settings. Residential Treatment for Children and Youth, 15 (4), 61-67.

O'Malley, F. (1996). Critical issues in short-term residential treatment. Residential Treatment for Children and Youth, 14 (1), 61-73.

Palmer, S., & Cooke, W. (1996). Understanding and countering racism with first nations children in out-of-home care. Child Welfare League of America, LXXV (6), 709-725.

Paternoster, R., & Mazerolle, P. (1994). General strain theory and delinquency: A replication and extension. Journal of Crime and Delinquency, 31 (3), 235-263.

Pazaratz, D. (1999). An impressionable evaluation of the efficacy of a residential treatment facility for emotionally disturbed youth. Residential Treatment for Children and Youth, 16 (3), 15-35.

Ponce, D. E. (1995). Value Orientation: Clinical applications in a multi-cultural residential treatment center for children and youth. Residential Treatment for Children and Youth, 12 (4), 29-42.

Pratt, D. I., & Moreland, K. L. (1996). Introduction to treatment outcome: Historical perspectives and current issues. Residential Treatment for Children and Youth, 13 (4), 1-27.

Pugh, R. H., Tepper, F. L., Halpern-Felsher, B. L., Howe, T. R., Tomlinson-Keasey, C., & Parke, R. D. (1997). Changes in abused children's social and cognitive skills from intake to discharge in a residential treatment center. Residential Treatment for Children and Youth, 14 (3), 65-81.

Rafal, A. S. (1994). The adolescent adjustment scale. Residential Treatment for Children and Youth, 12 (2), 49-65.

Russell, G. L., Fujino, D. C., Sue, S. Cheung, M., & Snowden, L. R. (1996). The effects of therapist-client ethnic match in the assessment of mental health functioning. Journal of Cross-Cultural Psychology, 27 (5), 598-615.

Schulman, M. (1996). The caring profile: A values modification program for adolescents in residential facilities. Residential Treatment for Children and Youth, 14 (1), 9-23.

Shennum, W. A., & Carlo, P. (1995). A look at residential treatment from the child's point of view. Residential Treatment for Children and Youth, 12 (3), 31-42.

Sherraden, M. S., & Segal, U. A. (1996). Multicultural issues in child welfare. Children and Youth Services Review, 18 (6), 497-504.

Simmerman, S. (1997). A comparison of educational programs at an adolescent residential treatment center. Residential Treatment for Children and Youth, 15 (1), 1-9.

Stage, S. A. (1999). Predicting adolescents' discharge status following residential treatment. Residential Treatment for Children and Youth, 16 (3), 37-55.

Stevens, J. W. (1997). African American female adolescent identity: A three-dimensional perspective. Child Welfare League of America, LXXVI (1), 145-172.

Summerville, M. B., Kaslow, N. J., Abbate, M. F., & Cronan, S. (1994). Psychopathology, family functioning, and cognitive style in urban adolescents with suicide attempts. Journal of Abnormal Child Psychology, 22 (2), 221-235.

Tafarodi, R. W., & Swann, W. B., Jr. (1996). Individualism-collectivism and global self-esteem: Evidence for a cultural trade-off. Journal of Cross-Cultural Psychology, 27 (6), 651-672.

Terpstra, J. (1998). Child caring institutions: A fact sheet. Residential Treatment for Children and Youth, 16 (2), 33-34.

Thompson, R. W., Smith, G. L., Osgood, D. W., Dowd, T. P., Friman, P. C., & Daly, D. L. (1996). Residential Care: A study of short- and long-term educational effects. Children and Youth Services Review, 18 (3), 221-242.

Tucker, C. M., Herman, K. C., Brady, B. A., & Fraser, K. P. (1995). Operation positive expression: A behavior change program for adolescent halfway house residents. Residential Treatment for Children and Youth, 13 (2), 67-79.

Van Gijseghem, H., & Gauthier, M. (1994). Links between sexual abuse in childhood and behavioural disorders in adolescent girls: A multivariate approach. Canadian Journal of Behavioural Science, 26 (3), 339-352.

Watkins, D., & Regmi, M. (1996). Within culture and gender differences in self-concept. Journal of Cross-Cultural Psychology, 27 (6), 692-699.

Watkins, D., & Cheung, S. (1995). Culture, gender, and response bias: An Analysis of responses to the self-description questionnaire. Journal of Cross-Cultural Psychology, 26 (5), 490-504.

Whittaker, J. K. (2000). The future of residential group care. Child Welfare League of America, LXXIX (1), 59-73.

Whittaker, J. K., & Pfeiffer, S. I. (1994). Research priorities for residential group child care. Child Welfare League of America, LXXIII (5), 583-601.

Wright, N. A. (1995). Social skills training for conduct-disordered boys in residential treatment: A promising approach. Residential Treatment for Children and Youth, 12 (4), 15-28.

Wright, S. C., Taylor, D. M., & Ruggiero, K. M. (1996). Examining the potential for academic achievement among Inuit children: Comparisons on the raven coloured progressive matrices. Journal of Cross-Cultural Psychology, 27 (6), 733-753.

Young, T. J., & French, L. A. (1997). Homicide rates among Native American childhood: The status integration hypothesis. Adolescence, 32 (125), 57-59.

Zimmerman, D. P. (1994). A pilot demographic study of population changes in a residential treatment center. Residential Treatment for Children and Youth, 11 (3), 17-33.

Zimmerman, D. P. (1996). A comparison of commonly used treatment measures. Residential Treatment for Children and Youth, 13 (4), 49-69.

Zimmerman, D. P., & Dillard, J. (1994). The Rorschach assessment of sexually abused children in residential treatment: A research note. Residential Treatment for Children and Youth, 12 (1), 59-72.

Bibliography

Achenbach, T. M. (1991). Manual for the Youth Self-Report and 1991 profile. Burlington Vermont: University of Vermont Department of Psychiatry.

Asarnow, J. R., Aoki, W., & Elson, S. (1996). Children in residential treatment: a follow up study. Journal of Clinical Child Psychology, 25 (2), 209-214.

Blackman, M. Eustace, J. & Chowdhury, T. (1991). Adolescent residential treatment: A one to three year follow-up. Canadian Journal of Psychiatry 36, (7), 472-479.

Bohm, D. (1980). Wholeness and implicate order. London: Routledge & Kegan Paul.

Brooks, R. B. (1994). Children at risk: Fostering resilience and hope. American Journal of Orthopsychiatry, 64, (4), 545-553.

Burbank, P. (1999). NUR660 fall semester: Philosophical foundations of health care research. Department of Nursing, University of Rhode Island.

Burks, J. (1995). Edgewood childrens' center outcome study 1991/1992. Residential Treatment for Children and Youth, 13 (2), 31-40.

Burnett, R., & Rathbun, C. (1993). Discovery and treatment of adolescent sexual offenders in a residential treatment center. Residential Treatment for Children and Youth, 11 (2), 57-64.

Burt, H. (1993). Issues in art therapy with the culturally displaced American Indian youth. Arts In Psychotherapy, 20 (2), 143-151.

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. Journal of Abnormal and Social Psychology, 65, 145-153.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences: Second edition. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cook, T. D., Cooper, H. Cordray, D. S., Hartmann, H., Hedges, L. V., Light, R. J., Louis, T. A., & Mostelller, F. (1994). Meta-Analysis for explanation: A casebook. New York: Russell Sage Publications.

Cooper, H., & Hedges, L. V. (1994). The handbook of research synthesis. New York: Russell Sage Publications.

Curry, J. (1991). Outcome research on residential treatment: Implications and suggested directions. American Journal of Orthopsychiatry, 61 (3), 348-357

Deschenes, E. P., & Greenwood, P. W. (1998). Alternative placement for juvenile offenders: Results from the evaluation of the Nokomis challenge program. Journal of research in Crime and Delinquency, 35 (3), 267-293.

Duchnowski, A., Johnson, M. K., Hall, K. S., & Kutash, K. (1993). The alternatives to residential treatment study: Initial findings. Journal of Emotional and Behavioral Disorders, 1 (1), 17-26.

Dukes, R. L., & Martinez, R. (1994). The impact of ethnicity on self-esteem among adolescents. Adolescence, 29 (113), 105-114.

Enns, C. Z. (1994). On teaching about the cultural relativism of psychological constructs. Teaching of Psychology, 21 (4), 205-211.

Epps, K. (1991). The residential treatment of adolescent sex offenders. Issues in Criminology and Legal Psychology, 1 (17), 58-67.

Fisher, P. A., Storck, M., Bacon, J. G. (1999). In the eye of the beholder: Risk and protective factors in rural American Indian and Caucasian adolescents. American Journal of Orthopsychiatry, 69 (3) 294 - 304.

Friedman, S. S. (1995). Beyond White and Other: Relationality and narratives of race in feminist discourse. Signs, 21 496-503.

Gutierrez, S. E., Felipe-Russo, N., & Urbanski, L. (1994). Sociocultural and psychological factors in American Indian drug use: Implications for treatment. The International Journal for Addictions, 29 (14), 1761-1786.

Gutierrez, S. E., & Todd, M. (1997). The impact of childhood abuse on treatment outcomes of substance users. Professional Psychology Research and Practice, 28 (4), 348-354.

Guba, E. G. (1990). The paradigm dialogue. Sage Publications.

Harlow, L. L., Muliak, S. A., & Steiger, J. H. (1997). What if there were no significant tests? Mahwah, New Jersey: Lawrence Erlbaum Associates.

Harlow, L. L. (2000). Psychology 533 spring semester: Multivariate methods in psychology coursepack. Department of Psychology, University of Rhode Island.

Harman, W. W. (1991). A re-examination of the metaphysical foundations of modern science: Causality issues in contemporary science. Sausalito, California: Institute of Noetic Sciences.

Hiller, M., L. (1996). Effectiveness of a specialized women's intervention in a residential treatment program. Substance Use and Misuse, 31 (6), 771-783.

Hsu, D. Y. F. (1995). Internalize culture, culturocentrism, and transcendence. Counseling Psychologist, 23, 4-24.

Hodges, K. (1990). Child and adolescent functional assessment scale. Vanderbilt Children Health Services Evaluation Project.

Hunt, M. (1997). How science takes stock. New York: Russell Sage Publications.

Kaplan, G. A., Gleason N., & Klein, R. Women's self-development in late adolescence. In Women's growth in connection, 1991. New York: The Guilford Press.

Keller, F. E., & Longino, H. E. (1996). Feminism and science. New York: Oxford University Press.

Klemke, E. D., Hollinger, R., Wyss Rudge, D., & Kline, D. (1998). Introductory readings in the philosophy of science. Amherst, New York: Prometheus Books.

Kuhn, T. S. (1970). The structure of scientific revolutions. International Encyclopedia of Unified Science 2nd Edition. The University of Chicago Press.

Landrine, H., Klonoff, E. A., & Brown-Collins, A. (1992). Cultural Diversity and methodology in feminist psychology: Critique, proposal, and empirical example. Psychology of Women Quarterly, 16, 145-163.

Landrine, H. (1995). Introduction: Cultural diversity, contextualism, and feminist psychology. In H Landrine (Ed) Bringing cultural diversity to feminist psychology.

Lott, B. 1994. Women's lives: Themes and variations in gender learning, Second Edition. Pacific Grove: Brooks/Cole Publishing.

Lipsey, M. W. (1990). Design sensitivity: Statistical power for experimental research. Newbury Park, California: Sage Publications.

Maddock, J. E. (1999). Statistical power and effect size in the Field of Health Psychology. Unpublished doctoral dissertation, University of Rhode Island.

Meehl, P. E. (1978). Theoretical Risks and Tabular Asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. Journal of Counseling and Clinical Psychology, 46, 806-834.

Mennen, F. E., & Meadow, D. (1993). The relationship of sexual abuse to symptom levels in emotionally disturbed girls. Child and Adolescent Social Work Journal, 10 (4), 319-328.

Miller, A. (1986). Thou shall not be aware: Societies betrayal of the child. New York, New York: Meridian Publishers.

Miller, D., Trapani, C., Fejes-Mendoza, K., Eggleston, C., & Dwiggins, D. (1995). Adolescence 30, (118), 429-435.

Miller, J. B. (1976). Toward a New Psychology for Women. Boston: Beacon Press.

Miller, J. R., & Burbank, P. M. (1994). Congruence among philosophy, theory, and method in nursing research. In workgroup of European nurse researchers, The contribution of nursing research: Past-Present-Future, vol. 2, Proceedings at the 7th Biennial Conference, Oslo, Norway: Falch Hurtigtrykk, 704-709.

Mosedale, F. E. (1979). Philosophy of science: The wide range of interaction. New Jersey: Prentice Hall.

Nobel, D. N., & Gibson, D. (1994). Family values in action: Family connectedness for children in substitute care. Child and Youth Care Forum, 23 (5), 315-328.

O'Mally, F. (1996). Critical issues in short-term residential treatment. Residential Treatment for Children and Youth, 14 (1), 61-74.

Packard, S. A. & Polifroni, C. E. (1990). The nature of scientific truth. Nursing Science Quarterly, 5 (4), winter. 158-163.

Pfeiffer, S. I., & Strzelecki, S. C. (1990). Inpatient psychiatric treatment of children and adolescents: A review of outcome studies. Journal of the American Academy of Child and Adolescent Psychiatry, 29 (6), 847-853.

Ponce, D. E. (1995). Value orientation: Clinical applications in a multi-culture residential treatment center for children and youth. Residential Treatment for Child and Youth, 12 (4), 29-42.

Pratt, S. I., & Moreland, K. L. (1996). Introduction to treatment outcome: Historical perspective and current issues. Residential Treatment for Child and Youth, 13 (4), 1996, 1-27.

Rosenberg, M. (1965). Society and the Adolescent Self Image. Princeton, New Jersey: Princeton University Press.

Rosenthal, R. (1979). The "file drawer" problem and tolerance for null results. Psychological Bulletin, 86, 638-641.

Rosenthal, R. (1991). Meta-analytic procedures for social research. Newbury Park: Sage Publications.

Rosnow, R. L. & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. American Psychologist, 44, 1276-1284.

Rossi, J. S. (1990). Statistical power in psychological research. What have we gained in 20 years? Journal of Consulting and Clinical Psychology, 58, 1276-1284.

Sagan, C. (1996). The demon haunted world: Science as a candle in the dark. Random House Publishers.

Scheidlinger, S. (1992). Therapist gender in child-adolescent treatment groups. Journal of Child and Adolescent Group Therapy, 2 (2), 104-108.

Small, R., Kennedy, K., & Bender, B. (1991). Critical issues for practice in residential treatment: The view from within. American Journal of Orthopsychiatry, 61 (3), 327-337.

Stern, L. (1991) Conceptions of separation and connection in female adolescents. In C. Gilligan, C., Lyons, N., & Hammer, T. (Eds.), Making Connections: The Relational World of Female Adolescents at Emma Willard School (p. 73-87). Cambridge: Harvard University Press.

Surrey, J. L. (1991). The self-in-relation: A theory of women's development. In Women's Growth and Development, 1991. New York: The Guilford Press.

Tabachnick, B. G., & Fidell, L. S. (1996). Using multivariate statistics: Third edition. New York: HarperCollins Publishers.

Tafarodi, R. W., & Swann, W. B. (1996). Individualism-collectivism and global self-esteem. Journal of Cross Cultural Psychology, 27 (6), 651-672.

Unger, R., & Crawford, M. (1996). Women and gender: A feminist psychology. United States: McGraw Publishing.

US Department of Health and Human Services: Center for Substance Abuse Prevention (1993). Measurement in prevention: A manual on selecting and using instruments to evaluate prevention programs. DHHS Publication No. (SMA) 93-2041.

Valliant, P. (1993). Cognitive and behavioral therapy with adolescent males in residential treatment care. Journal of Child and Youth Care, 8 (3), 41-49.

Wheelis, A. (1971). The end of the modern age. New York: Harper & Row.

Wells, K. (1991). Long-term residential treatment for children: Introduction. American Journal of Orthopsychiatry, 61 (3), 324-326.

Wright, N. (1995). Social skills training for conduct- disorder boys in residential treatment. Residential Treatment for Children and Youth, 12 (4), 15-28.

Yandow, V. (Video, 1990). Alcohol and drug problems in women. The Center for Interdisciplinary Studies: Battelboro, Vermont.

Zimmerman, P., D. (1996). A comparison of commonly used treatment measures. Residential Treatment for Children and Youth, 13 (4), 49-69.

Zohar, D., & Marshall, I. (1994). The quantum society. New York: Morrow and Company Inc

Zukav, G. (1979). The dancing wu li masters: An overview of the new physics. New York.