GIRLS CIRCLE AND THE COUNCIL FOR BOYS AND YOUNG MEN IN SCHOOLS: EVALUATION OF TWO GENDER-SPECIFIC SUPPORT GROUPS

by

CADY V. KINTNER

A DISSERTATION

Presented to the Department of Counseling Psychology and Human Services
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

September 2020

DISSERTATION APPROVAL PAGE

Student: Cady V. Kintner

Title: Girls Circle and The Council for Boys and Young Men in Schools: Evaluation of Two Gender-Specific Support Groups

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Counseling Psychology and Human Services by:

Dr. Leslie Leve	Chairperson
Dr. David DeGarmo	Core Member
Dr. Elizabeth Stormshak	Core Member
Dr. Emily Tanner-Smith	Core Member
Dr. Nicole Giuliani	Institutional Representative

and

Janet Woodruff-Borden	Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded September 2020

DISSERTATION ABSTRACT

Cady V. Kintner

Doctor of Philosophy

Department of Counseling Psychology and Human Services

September 2020

Title: Girls Circle and The Council for Boys and Young Men in Schools: Evaluation of Two Gender-Specific Support Groups

Although there is evidence in support of gender-specific developmental trajectories for adolescent risk and protective factors, less is known about the efficacy of gender-specific interventions applied in educational contexts. Girls Circle and The Council for Boys and Young Men are widely implemented gender-specific support groups for youth ages nine to 18 that incorporate elements of relational-cultural theory, resiliency practices, and motivational interviewing into weekly closed group sessions that can be conducted during the school day. The present study is a program evaluation of these interventions that was designed in partnership with program founders, public school personnel, and the local juvenile department in a Pacific Northwest county. At the beginning of the 2015-2016 academic year, 309 (138 male, 171 female) students ages 11-19 in six schools that were already implementing Girls Circle and/or The Council were randomized to either receive the intervention in Fall term (intervention condition), Spring term (wait-list control condition), or services as usual (control condition). Participants completed surveys at the beginning of the school year, 12-15 weeks later, and again at the end of the school year (approximately nine months after baseline). Fidelity measures, attendance logs, and school records information were also collected. The study's aims were to: (a) assess the extent to

which process components (adherence, dosage, and quality of program delivery) were achieved, and (b) examine potential intervention effects on outcome variables targeted by the interventions (self-efficacy, prosocial behavior, perceived social support, body image, and school engagement). Overall, Girls Circle and The Council were both implemented with good adherence, dosage, and quality. Latent growth modeling revealed mostly null effects for the intervention outcomes as measured over the course of one academic year, with the exception of boys in The Council waitlist-control condition showing faster growth in self-efficacy than participants in other conditions. ANCOVA models assessing intervention effects for school engagement also revealed null effects for intervention. Post-hoc analyses examining potential dosage effects did not change the pattern of results. Implications for school-based implementation of these programs are discussed.

CURRICULUM VITAE

NAME OF AUTHOR: Cady V. Kintner

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Stanford University, Stanford, California

DEGREES AWARDED:

Doctor of Philosophy, Counseling Psychology, 2020, University of Oregon
Master of Science, Counseling, Family and Human Services, 2017, University of Oregon
Bachelor of Arts, Psychology, Feminist Studies, 2005, Stanford University

PROFESSIONAL EXPERIENCE:

Psychology Doctoral Intern, University Counseling Center, University of Oregon, Eugene, 2019-2020

Academic and Career Development Programming Assistant, Services for Student-Athletes, University of Oregon, Eugene, 2016-2019

Psychology Assessment Extern, Strong Integrated Behavioral Health, Eugene, Oregon, 2018-2019

Psychology Extern, Oregon State Hospital, Junction City, 2017-2018

Research Intern Supervisor, Prevention Science Institute, University of Oregon, 2017

Family Therapist, Child and Family Center, University of Oregon, Eugene, 2016-2017

Interviewer, Early Growth and Development Study (EGDS), Prevention Science Institute, University of Oregon, Eugene, 2015-2017

Practicum Intern, University Counseling Center, University of Oregon, Eugene, 2015-2016

Youth Skills Coach, Safe, Healthy Adolescent Relationships and Peers (SHARP), Oregon Social Learning Center, Eugene, Oregon, 2015-2016

University Supervisor, Family and Human Services, University of Oregon, Eugene, 2014-2016

Associate Program Director, Outward Bound California, San Francisco, 2014

Program Manager, Outward Bound California, San Francisco, 2013

Course Director and Lead Instructor, Outward Bound California, San Francisco, 2012-2018

Course Director and Lead Instructor, Outward Bound At-Risk Youth Programs, Florida and Alabama, 2009-2016

GRANTS, AWARDS, AND HONORS:

Hill M. Walker Legacy Scholarship, University of Oregon, 2018-2019

Thomas W. Vollmer Memorial Scholarship, University of Oregon, 2018-2019

General University Scholarship, University of Oregon, 2016-2019

Society for Prevention Research Travel Award, 2016-2017

College of Education Alumni Award, University of Oregon, 2015-2017

Iain More Stellar Student Award, University of Oregon, 2014-2015

PUBLICATIONS:

Kintner, C. V., Fisher, P. A., & Leve, L. D. (in press). Treatment Foster Care Oregon. In R. Ax, J. E. Louden, M. Epperson, J. Mills, M. Oliver, K. Stockdale, & A. Wilson (Eds.), *The SAGE encyclopedia of criminal psychology*. SAGE.

Leve, L. D., Harold, G. T., & Kintner, C. V. (in press). Domestic adoption. In S. Hupp & J. Jewell (Eds.), *The encyclopedia of child and adolescent development.* Hoboken, NJ: Wiley Blackwell.

Leve, L. D., Fisher, P. A., & Kintner, C. V. (2019). Foster families. In B. H. Fiese, M. Celano, K. Deater-Deckard, E. Jouriles, M. Whisman (Eds.), *APA handbook of contemporary family psychology*: Washington, DC: American Psychological Association.

ACKNOWLEDGMENTS

I wish to gratefully acknowledge Dr. Leslie Leve for her ongoing mentorship in conducting this evaluation and guidance in the preparation of this manuscript. Special thanks are due to Dr. Dave DeGarmo for the consultation he provided regarding statistical analyses, and to Dr. Emily Tanner-Smith for her close readings and thoughtful feedback on initial drafts. I also thank Dr. Nicole Giuliani and Dr. Beth Stormshak for their support throughout the dissertation process and ideas for future directions in which to extend this research.

Staff and students at the University of Oregon's Prevention Science Institute were essential to the data preparation and processing stages of this study. Large thanks to Sally Guyer and Charlotte Winter for their kindness, patience, and excellence related to data input and management. Sincere appreciation also goes to Alyssa Rayhel and Kelsey Van Brocklin for their support in managing data entry for multiple waves of paper forms, and to Rachel Kovensky and Emily Reich for their assistance in scanning close to 600 surveys.

This evaluation would not have been possible without the investment of One Circle program founders and all the effort put forth by our community partners in the Clackamas County juvenile justice department. Sincere thanks to Giovanna Taormina, Beth Hossfeld, Tanya Kramer, Mark McDonnell, and Judy Pearce. Last, but certainly not least, a huge thank you to all the students, group facilitators, teachers, and administrative staff who voluntarily completed surveys, fidelity measures, and school records forms over the course of a full academic year.

Dedicated to my family

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

During adolescence, adherence to traditional masculinity and femininity ideology can increase risk behaviors, including antisocial behavior, violence, and school dropout rates among boys, and depression, eating disorders, and low body self-image among girls (Marcotte, Fortin, Potvin, & Papillon, 2002; Millstein, Peterson, & Nightingale, 1993; Park, Goodyer, & Teasdale, 2005). Stereotypical beliefs about masculinity have been linked to boys' over-identification with being tough; instead of learning and accepting different ways of "being a man," boys may instead be more likely engage in problematic externalizing behavior in order to fit in (Hossfeld, Gibraltarik, Bowers, & Taormina, 2008; Levant et al., 2003). Meanwhile, traditional concepts of femininity often include distorted body image/higher likelihood of experiencing relational aggression, and this may lead to internalizing symptoms and eating disorders/low body self-image (McCabe & Ricciardelli, 2003).

One Circle Foundation's Girls Circle and The Council for Boys and Young Men (The Council) are gender-specific structured support groups intended to address harmful gender stereotypes, facilitate positive peer relationships, and ultimately improve mental health and behavioral outcomes for adolescents. Thousands of students have participated in these interventions since Girls Circle first began in 1994 and The Council in 2006. Yet, only one randomized controlled trial has been implemented and published (Gies et al., 2015), with limited generalizability of results. Given that allocation of curriculum hours is of key significance to students, teachers, and administrators, it is beneficial to program

creators and implementation teams to employ a rigorous approach to evaluate outcomes for programs such as Girls Circle and The Council in schools.

The present study is a program evaluation of Girls Circle and The Council as implemented in six public schools over the course of one academic year. The overall purpose of this study was to address the feasibility and impact of including Girls Circle and The Council as a part of public school services for adolescents. A randomized waitlist controlled design was employed to assess these interventions as implemented in six public schools in a county adjacent to a major Pacific Northwest city. Both process and outcome evaluation components were addressed in partnership with program founders, public school personnel, and the local juvenile department.

Girls Circle and The Council Program Models

One Circle Foundation program models are gender-specific, and integrate relational-cultural theory, resiliency practices, and social skills training, with the intent to increase participants' positive connections, competence, and personal and collective strengths (www.onecirclefoundation.org). Both Girls Circle and The Council facilitators are trained in motivational interviewing techniques, and are thus encouraged to lead group discussions by engaging youths' intrinsic motivation, rather than lecturing or demanding changes in behavior (Miller & Rollnick, 1991). According to the One Circle Foundation website (www.onecirclefoundation.org), facilitators are also trained to stimulate critical thinking and, through structured activities and guided debriefs, focus attention on youths' areas of strength rather than skill or behavior deficits. Trauma-responsive and culturally informed practices are considered inherent to the program model. The intervention was developed for youth ages 9-18. Groups are closed to new

members after initiation, and are designed to meet weekly for 1.5-2 hours for 10-12 weeks. In the present study, groups met weekly for approximately one hour for a total of 10 weeks; this is reflective of common practices when implementation occurs as part of a school day. Youth are guided in taking turns talking and listening to one another about their concerns and interests. Gender-specific themes and topics, as well as skill-based activities are introduced. For girls, this means an emphasis on positive connection, personal and collective strengths, and competence. Boys are encouraged to question stereotypical concepts about masculinity, with an emphasis on increasing boys' emotional, social, and cultural literacy. Youth in both Girls Circle and The Council are exposed to curricula and facilitated experiences that target decreasing substance use, relational aggression, and conduct problems. Positive body image, ethnic pride and respect for diversity, communication skills, healthy friendships, and prosocial behavior are also incorporated into curricula for both genders.

Program Theory

Girls Circle and The Council are both based on relational-cultural theory, resiliency practices, and motivational interviewing techniques. Relational-cultural theory was originally developed as a way to conceptualize women's relational interactions and the role that caretakers and cultural contexts have in these interactions (Miller, 1986). This theory suggests that relational characteristics and growth-fostering interactions are the primary bases from which overall psychological health and resiliency are achieved (Jordan, 2009). The seven fundamental concepts for the theory (Jordan, 2009; Miller, 1986) are the following: (a) individuals grow through and toward relationships during the course of their lives; (b) mature functioning is characterized as movement toward mutual

empathy; (c) relationship differentiation and elaboration characterize true growth; (d) mutual empathy and empowerment are the core of growth-fostering relationships; (e) authenticity is necessary for real engagement and individuals must feel as if they can be themselves without judgment; (f) in growth-fostering relationships, all people involved contribute equally and experience positive outcomes as a result of being in the relationship; and (g) relational competence and capacities are enhanced over the life span.

Although few studies have explicitly examined the effectiveness of relational-cultural theory in comparison with more widely implemented approaches (e.g., cognitive behavioral therapy and other behavioral theoretical orientations), existing research on relational-cultural theory-based interventions suggests their effectiveness in improving resiliency factors among youth and adolescents of color (Lenz, Speciale, & Aguilar, 2012; Morray & Liang, 2005; Sparks, 2004). The findings from these studies suggest relational-cultural theory as a promising approach for improving empathy in cross-cultural relationships between Arab and Israeli youth (Morray & Liang, 2005) and improving at-risk African American and Latina adolescent girls' personal strength, authenticity, and motivation to change relationship patterns (Lenz Speciale, & Aguilar, 2012; Sparks, 2004).

Other studies have used relational-cultural theory as an exploratory framework to better understand the relational needs of middle school girls. One case study explored how applying relational-cultural theory practices in school counseling settings with early adolescent girls (ages 11-14) showed promise for expanding students' relational capacities, social skills, and self-regulation (Tucker, Smith-Adcock, & Trepal, 2011). Another demonstrated how relational-cultural theory could be used in counseling a group

of middle school girls (Cannon, Hammer, & Curtin, 2012). In this study, female adolescent participants in a relational-cultural theory-based skills group were able to acknowledge and address social stratification in their peer groups, build empathy for other group participants, and elicit change-talk. Although both of these studies were explorative in nature, they provide preliminary support for applications of relational-cultural theory in individual and group-based settings as part of a typical school day.

Within the relational-cultural approach, Girls Circle and The Council incorporate resiliency practices. Resilience is generally defined as the process of adapting well in the face of adversity, trauma, tragedy, threats or significant sources of stress (American Psychological Association, 2018). As proposed by Benard (2004), resiliency development in youth is reliant on high expectations, caring and support, and meaningful participation within their communities. In addition, positive identification with one's cultural, ethnic, or racial group is believed to also increase resiliency traits. Girls Circle and The Council's curricula incorporate activities specifically aimed at increasing youth resiliency, as well as a group format that encourages mutual empathy and support. Resiliency development in adolescence is associated with stronger prosocial skills, coping abilities, and decreased relational aggression and victimization (Donnon, 2010). Adolescent girls engaging in Young Women Leaders Program, a group mentoring programs that, like the Girls Circle and The Council, encourages mutual empathy and support, reported improved social and relational skills, self-regulation, and self-understanding upon group completion (Deutsch et al., 2017).

Girls Circle and The Council facilitators also apply motivational interviewing techniques that are intended to promote resiliency and protective factors. The primary

goal of motivational interviewing is to help resolve the ambivalence that prevents individuals from creating and sustaining positive change in their lives (Miller & Rollnick, 1991). The five key principles of motivational interviewing are as follows: (a) express empathy through reflective listening; (b) develop discrepancy between individuals' goals or values and their current behavior; (c) avoid argument and direct confrontation; (d) adjust to client resistance rather than opposing it directly; and, (e) support self-efficacy and optimism. Motivational interviewing is a counseling method initially developed as a means of treating adults with substance use issues (Miller & Rollnick, 1991). Since then, it has been broadly applied to a range of other populations of various ages and presenting concerns, with promising results (Gayes & Steele, 2014).

Motivational interviewing has also been shown helpful for individual and group interventions targeting adolescent behavioral outcomes. Meta-analysis of motivational interviewing interventions targeting adolescent health behaviors provides support for its effectiveness in improving outcomes such as substance use, sexual health, and physical activity in teens (Cushing et al., 2014). Motivational interviewing may also be a promising tactic to facilitate engagement in adolescent mental health settings; adolescents with anxiety and mood disorder randomized to attend support groups in which facilitators used motivational interviewing techniques were more likely to have higher session attendance than control conditions (Dean et al., 2016).

In addition to motivational interviewing approaches employed by facilitators, Girls Circle and The Council logic models also specifically target respectful communication and supportive relationships between group members. Respectful communication between teachers and adolescent students has been connected to

improved school engagement and self-efficacy (Holloman & Yates, 2013) and perceived peer social support can buffer adolescents against negative psychosocial symptoms such as depression, anxiety, and low self-esteem (Demaray & Malecki, 2002).

With approaches grounded in relational-cultural theory, resiliency practices, and motivational interviewing facilitation techniques, Girls Circle and The Council program models emphasize empathic connections between group members as a driving force behind girls' and boys' psychological and behavioral health. Integral to the change process are attentive listening, respectful communication, and positive relationships between group members, and facilitator-driven motivational interviewing practices encouraging self-reflection and authenticity. Then, within the context of a supportive and reflective group setting, curriculum elements target gender-specific risk factors and positive identity formation.

Previous Research

Although both Girls Circle and The Council are widely implemented in thousands of school-based and juvenile justice settings (www.onecirclefoundation.org), methodologically rigorous analyses of these programs' efficacy have yet to be widely published in peer-reviewed journals; Steese et al.'s 2006 article in *Adolescence* is the sole evaluation of either program published in a peer-reviewed journal. Using a pre-post-test single group design, Steese et al. (2006) found promising results for Girls Circle participants (improved body image, perceived social support, and self-efficacy). Further studies of both interventions have been conducted primarily by consultants specializing in program evaluation, and, with the exception of Gies et al.'s 2015 report (a 2-year follow-up randomized control trial examining Girls Circle outcomes for girls on

probation in Cook County), all other formal evaluations have relied on pre-post-test single group designs. These reports have primarily focused on Girls Circle; and results include positive short-term results for self-efficacy, body image, and perceived social support (Irvine, 2005), interpersonal skills and relationships (Roa et al., 2007), and perceived body image, communicating needs to adults, and self-efficacy (Irvine & Roa, 2010). A program evaluation of a Girls Circle juvenile justice sample (Gies et al., 2015) used a randomized controlled design, but found no significant intervention effects for recidivism, psychosocial assets, school/aspirations, nor perceived body image. However, there were fidelity and attrition concerns in Gies' report, and significant effects for dosage emerged. As attendance in the Girls Circle group increased, average condom use, educational aspirations, and educational expectations significantly increased; conversely, average self-control scores significantly decreased. The authors explain this unexpected decrease in self-control by postulating that, while Girls Circle participants may have initially felt in control of their lives, after revealing and acknowledging their past behaviors in the group setting, they may have further reflected on their actual self-control more accurately during the posttest, and thus rated themselves lower.

Compared to Girls Circle, The Council has had fewer evaluation reports released (two are included on the One Circle Foundation website), and existing research has not yet been published in peer-reviewed journals. The first formal evaluation of The Council (Gray et al., 2008) used a single group pre-post-test design, evaluated 93 participants of diverse ages and who received the intervention in a range of settings, and showed significant increases in school engagement scores and no differences on other study variables. The second report (Mankowski et al., 2011) evaluated young men in juvenile

correctional facilities, and was intended to be a randomized experimental design. However, due to complications with implementation, randomization did not occur, significant between group differences at baseline emerged, and no changes were found between intervention and control groups. There was, however, a significant effect for intervention dosage on decreasing the rate of increase of traditional masculinity beliefs. Perhaps because Girls Circle was created first, there have been fewer impact studies released on The Council, and less is known about its short- or long-term outcomes for boys and young men.

In sum, results from existing evaluations of Girls Circle and The Council provide some indications of program efficacy, but have limited generalizability and the study designs have been mixed in their methodological rigor. Single group pre-post-test studies examining Girls Circle provide some evidence of participants' improvement in self-efficacy, body image, perceived social support, interpersonal skills, relationship skills, and ability to communicate needs to adults. However, in the absence of control or comparison groups, we cannot draw rigorous conclusions about the degree to which these factors might have improved on their own without participation in these programs. With inclusion of a control or comparison group, there is better interpretation of effect sizes. Especially when assignment to treatment and control groups is randomized, there is increased likelihood that improvement is due to the intervention rather than to confounding variables or developmental trends over time. Additionally, without continued data collection beyond immediate post-intervention surveys, the implications of these studies are limited to short-term outcomes. It is possible that participants in the

interventions may experience positive results later on, or that initial positive effects may diminish over time.

Results from the one randomized controlled trial evaluating Girls Circle showed no intervention effects, although it did find dosage effects for average condom use, educational aspirations, and educational expectations. Meanwhile, only two studies have evaluated The Council; one single group design found a short-term increase in school engagement, while the other found no intervention effects, but a dosage effect for a decreased rate of increase in traditional masculinity beliefs. Existing research of these two interventions has been limited by weak evaluation designs, including issues with implementation and fidelity, and high levels of attrition. Additionally, although both programs are widely implemented in schools, no existing studies have focused specifically on school-based programs. Given the likely potential benefits and wide dissemination of Girls Circle and The Council in schools, juvenile justice settings, and youth-serving non-profits, it is surprising that more rigorous evaluations of these programs have not yet been conducted. Evaluations with rigorous experimental designs could lead to increased uptake of these interventions, especially for partnering agencies looking to add or improve upon evidence-based programs in their services.

Integration of Process Evaluation Components

Including process evaluation components alongside outcome evaluation analyses is important because many program evaluations that show null effects for interventions typically also include implementation challenges that result in reduced service delivery (Patton, 2012). Assessing service delivery requires comparing intervention

implementation in actuality with what was intended by the program creators and the study design.

In general, service delivery is thought to include the following three core components: (a) adherence (extent to which programs are delivered as designed); (b) dosage (extent to which participants received the intended amount of the intervention over the intended time period); and, (c) quality with which the program was delivered (Gresham, 2014). All three components are important to consider in the context of program evaluation (Cook & Campbell, 1979; Durlak & DuPre, 2008; Power et al., 2005; Sanetti et al., 2011). If programs are not delivered with their core components to the appropriate participants, with the correct materials, and in the intended contexts (i.e., poor adherence), any observed intervention effects would rightly be called into question (Cook & Campbell, 1979). Second, if participants do not receive the sufficient level of services (i.e., insufficient dosage), it is unlikely that programs will be effective in achieving intended outcomes (Power et al., 2005). Finally, if the program is delivered in an unskillful manner (i.e., poor quality), it is also less likely to meet intended outcomes (Sanetti et al., 2011).

Existing outcome evaluations of Girls Circle and The Council indicate a need for including process evaluation components in this and future studies employing rigorous designs and quantitative analyses in assessing program efficacy and effectiveness. As noted in the previous section, significant dosage effects have emerged for both Girls Circle (increased average condom use, educational aspirations, and educational expectations; Gies et al., 2015) and The Council (decreased rate of increase for traditional masculinity beliefs; Mankowski et al., 2011). Gies et al. (2015) also conducted an

extensive process evaluation of Girls Circle that indicated an overall low degree of adherence, dosage, and program quality; these findings accompanied null effects for the intervention condition. Given that challenges can arise at various stages of program implementation and implicate multiple arms of service delivery, the present study prioritizes addressing process evaluation components prior to conducting outcome evaluation analyses.

CHAPTER II

THE PRESENT STUDY

Overview

The intent of the present study is to provide an independent program evaluation of Girls Circle and The Council as they are conducted in co-ed public schools. The project was conducted in collaboration with One Circle Foundation founders and the juvenile justice department local to the evaluated schools; the juvenile justice department was responsible for training/supporting facilitators, collecting data, and de-identifying data. Research personnel had no ties to evaluation outcomes and conducted data analysis independently from obligations to program stakeholders.

The present study evaluated Girls Circle and The Council groups that took place in six public middle and high schools during the 2015-2016 academic year. Schools were identified and selected based upon whether they were currently implementing Girls Circle and The Council; only schools already implementing Girls Circle or The Council were invited to participate in data collection for the program evaluation. A randomized waitlist-control design was applied at the student level to assign participants to the intervention support groups or to school services as usual. Analyses incorporate survey data completed by participants at three time points over the course of the academic year, fidelity and attendance measures, and school records information. Outcome variables included in analyses were self-efficacy, prosocial behavior, perceived social support, body image, and school engagement. These variables were selected based on their relation to target outcomes of Girls Circle and The Council curricula and because, as

protective factors, they were anticipated to be enhanced by the strengths-based design of both interventions.

Below are the process and outcome goals of this study, along with corresponding specific objectives to measure the success of each goal:

Process Goals

1. Adherence: Assess the extent to which Girls Circle and The Council programs were delivered as designed (i.e., evaluate feasibility of program implementation in schools).

a) Determine whether programs were delivered with all core components to the appropriate populations.

b) Determine whether the correct protocols, techniques, and materials were used in the correct locations/contexts (format component of fidelity).

2. Dosage: Assess the extent to which individuals in the intervention and waitlist-control conditions participated in each program in terms of sessions attended.

c) Determine the number of Girls Circle and The Council sessions attended by participants in each condition.

d) Determine the length of intervention provided.

3. Quality: Assess the quality of program delivery.

e) Evaluate the extent to which facilitation elements were achieved (process component of fidelity).

f) Evaluate the degree to which facilitator impressions of the Girls Circle and The Council programs were positive.

Outcome Goals

4. Overall Effectiveness of Offering Groups in Schools: Assess the effectiveness of participation in either Girls Circle or The Council in comparison with school services as usual as measured over the course of one academic year.

g) Determine whether intervention and/or waitlist-control participants demonstrated improved psychosocial assets (prosocial skills, self-efficacy, perceived social support, and body satisfaction) compared to control group participants as measured by self-report surveys completed at three time points over the course of the 2015-2016 academic year.

h) Determine whether intervention and/or waitlist-control participants differed from control group participants in school engagement as measured by school records submitted post-program completion at the end of the 2015-2016 academic year.

5. Effectiveness of Girls Circle and The Council Program Models: Assess for differences in effectiveness by program type (i.e., intervention effects may differ between Girls Circle and The Council).

i) Evaluate whether group type (Girls Circle vs. The Council) is a significant predictor for improved psychosocial assets (prosocial skills, self-efficacy, perceived social support, and body satisfaction).

j) Evaluate whether group type (Girls Circle vs. The Council) is a significant predictor for school engagement.

6. Dosage Effects: Determine whether the number of Girls Circle and/or The Council sessions attended impacts change in psychosocial assets and school engagement.

k) Incorporate Girls Circle and The Council session attendance data into outcome analyses; examine for potential main and moderating effects of dosage.

CHAPTER III

METHODS

Overview

This study is a program evaluation integrating process and outcome evaluation components. The process evaluation uses quantitative and qualitative methods to provide context to program implementation in schools, including attendance, fidelity measures, and facilitator impressions. The outcome evaluation uses a randomized waitlist-control design to assess program outcomes as measured over the course of an academic year.

Participants

At the beginning of the 2015-16 school year, 309 youth (66.7% White, 17.0% multiethnic, 2.7% American Indian/Alaska Native, 1.5% Native Hawaiian/Pacific Islander, 1.5% Asian American, 1.1% Black/African American, 9.5% missing/refuse to answer; 16.5% Hispanic/Latino/a) attending one of six public schools implementing Girls Circle and/or The Council were randomly assigned by the juvenile department to intervention, waitlist-control, or school services as usual. At baseline, youth ranged in age from 11-19 years old (M = 14.13, SD = 2.32), and self-reported their sexual orientation (73.8% straight, 7.8% bisexual, 3.6% other, 2.6% unsure/questioning, 0.3% gay/lesbian, 12.0% missing/refuse to answer) and gender identity. Demographic information for the intervention, waitlist-control, and control groups is reported in Table 1.

Table 1

Demographic information for the intervention group, waitlist-control group, control group, and overall sample.

	Intervention (n = 71)		Waitlist-control (n = 92)		Control (n = 146)		Overall sample (n= 309)	
	N	%	*N*	%	*N*	%	*N*	%
Middle Schools	35	49.3	39	42.4	56	38.3	130	42.1
School 1	18	25.4	15	16.3	37	25.3	70	22.7
School 2	4	5.6	19	20.7	12	8.2	35	11.3
School 3	13	18.3	5	5.4	7	4.8	25	8.1
High Schools	36	50.7	53	57.6	90	61.7	179	57.9
School 4	0	-	23	25.0	56	38.4	79	25.6
School 5	21	29.6	9	9.8	13	8.9	43	13.9
School 6	15	21.1	21	22.8	21	14.4	57	18.4
Gender: Girls	41	57.7	37	40.2	93	63.7	171	55.3
Ethnicity								
White	44	68.8	46	67.6	86	65.2	176	66.7
Multiethnic	10	15.6	11	16.2	24	18.2	45	17.0
American Indian/ Alaska Native	3	4.7	2	2.9	2	1.5	7	2.7
Native Hawaiian/ Pacific Islander	0	-	0	-	4	3.0	4	1.5
Asian American	0	-	1	1.5	3	2.3	4	1.5
Black/African American	1	1.6	0	-	2	1.5	3	1.1
Missing/refuse to answer	13	18.3	32	34.8	25	17.1	70	22.7
Hispanic/Latino/a	13	18.8	11	17.7	19	14.6	43	16.5
Sexual orientation								
Heterosexual/ straight	58	81.7	58	63.0	112	76.7	228	73.8
Bisexual	9	12.7	5	5.4	10	6.8	24	7.8
Other	0	-	3	3.3	8	5.5	11	3.6
Unsure/questioning	1	1.4	2	2.2	5	3.4	8	2.6
Gay or lesbian	1	1.4	0	-	0	-	1	0.3
Missing/refuse to answer	2	2.8	24	26.1	11	7.5	37	12.0
Age	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
	13.93	2.28	13.93	2.39	14.34	2.31	14.13	2.32

Note. n = number of participants assigned to each condition. *M(SD)* = mean(standard deviation).

For the purposes of randomized assignment to single-gender intervention groups, all 309 participants were identified by school facilitators as either male (n = 138) or female (n = 171). Three participants also self-identified as transgender, and six also self-identified as "other" with regard to gender identity. Participants who identified as transgender were assigned to gendered condition types congruent with their gender identity. Of the students who identified as "other", one was randomly assigned to a girls waitlist-control group, one was assigned to a boys waitlist-control group, and four were assigned to the control condition.

The six participating public schools (three middle schools, three high schools) were located in a Pacific Northwest county adjacent to a major regional city. Schools were selected for inclusion in consultation with program founders and the county juvenile department. Schools were selected for inclusion if they had already been running Girls Circle and/or The Council group for at least one academic year prior to baseline, could staff each group with a program-certified facilitator, and agreed to study participation. Eight schools were originally included, but data were never received by the research team from two of these schools. One of these two schools did administer baseline surveys, but struggled with low recruitment numbers overall, severe attrition issues in the intervention condition, and a majority of students opting out of completing surveys in all three conditions. This school thus did not collect any further data after baseline and data were not provided to the research site for use in the program evaluation. The second school that was dropped from the study planned to begin administering surveys, but data were never submitted to the research team. Figure 1 provides a flowchart of study participants at each point of data collection.

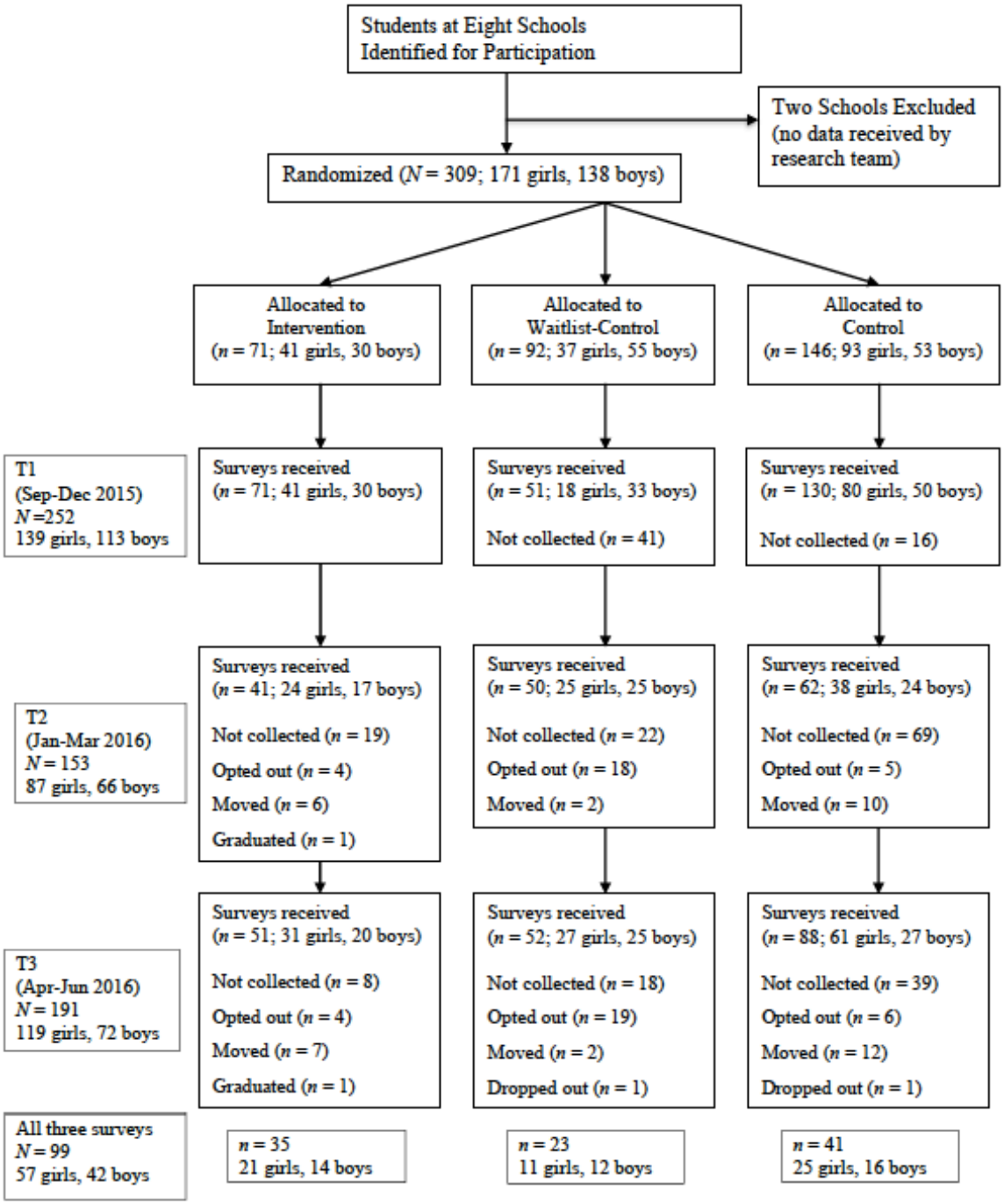

Figure 1. CONSORT flowchart of study participants, point of random assignment, and missing data at each stage of study for intervention, waitlist-control, and control groups.

Eligibility Criteria

Girls and boys were eligible for participation in the study if they were: (a) age 11-19 years old at enrollment; (b) enrolled in one of six middle or high schools that were

offering Girls Circle or The Council; and (c) fluent in either English or Spanish. Girls and boys were not eligible for the study if they (a) were medically fragile; (b) had a significant developmental disability; (c) graduated from high school or had a GED; or (d) were detained or confined in a penal institution at enrollment.

Prior to commencing, the study was reviewed by the IRB at the evaluation site and the research was deemed to be exempt due to it being a program evaluation of an existing school-based educational program and the lack of any identifiable data at the evaluation site. In the summer prior to the study, schools sent caregivers a letter providing them the opportunity to remove their children from the research evaluation and the groups. Youth in all conditions were provided with assent forms prior to completing surveys and were given opportunity to decline to participate.

Randomization

Girls and boys meeting eligibility criteria were assigned to participate in Girls Circle/The Council during the Fall or Winter of the 2015-2016 school year, or to classes as usual (e.g., health class, physical education, study hall, etc.). Participants opting out of the research study could still continue to participate in their assigned condition. The local county juvenile department assisted schools in randomizing the youth equally into the three conditions: intervention (participated in the intervention during Fall term), waitlist control (participated in the intervention during Winter term), and control (no opportunity for intervention participation). School administrators were provided with guidelines for randomizing students to the three conditions, and each school had at least one intervention group and one wait-list control group. Schools identified a list of students meeting eligibility criteria, then randomly selected from that initial list to form

intervention, control, and waitlist-control groups. Students were randomized to condition based on schools' pre-existing random assignment for class schedules; for example, if all eighth graders in a given school were typically in health class one term, PE class another term, and study hall another term, half of the students randomly assigned to study hall in the fall term might be assigned to the intervention condition.

Intervention, Waitlist-Control, and Control Conditions

Participants in the Girls Circle and The Council conditions (intervention and waitlist-control) met weekly for 10 weeks during the school day in single gender, closed groups. Group sessions typically lasted for one hour each, for a total of 10 curriculum hours per student. Participants only attended groups at the school they were enrolled in, though groups did include youth of different ages and grades within each school.

Adult facilitators (one to two per group, constant over the 10 weeks) led the groups, following the One Circle Foundation curriculum for Girls Circle or The Council. Each group had a single facilitator, although facilitators often directed multiple groups, sometimes at different schools. All facilitators participating in this study had been leading groups in these schools for at least one academic year previous to the initiation of the study, were the same gender as the participants in their groups, and all had completed a One Circle Foundation certified facilitator training prior to becoming facilitators.

Control group participants received school services as usual. Students in this condition were not assigned to participate in a Girls Circle or The Council group during the 2015-2016 academic year and instead attended regularly scheduled classes/study hall periods during the time intervention groups were being conducted. Control group participants had not received the Girls Circle or The Council intervention in a previous

academic year; they were, however, eligible to participate in future years following the conclusion of data collection.

Program founders and the juvenile department consulted with the research team in order to determine final scales and constructs to include in participant surveys, as well as to receive guidance on how to implement a randomized waitlist-control study design of the One Circle interventions. However, members of the research team were not involved in designing or implementing the content of the intervention, in training or communicating with intervention facilitators after the study began, or in the data collection activities.

Assessment Procedures and Measures

Youth self-report surveys. Youth participants in all three conditions completed a 75-item survey at the beginning of the school year, 12-15 weeks later, and again at the end of the school year (approximately nine months after baseline). Surveys were available in English and Spanish (two participants completed Spanish versions). Surveys were initially written in English, then translated to Spanish by a juvenile department staff member. Surveys measured nine outcome variables hypothesized to be improved by participation in the intervention: school engagement, self-efficacy, prosocial skills, perceived social support, body satisfaction, conduct problems, relational aggression, substance use, and ethnic pride and respect for diversity. The present study focuses on a subset of these self-report variables (prosocial skills, self-efficacy, perceived social support, body satisfaction) based on their direct relation to program theory and curricula.

Coding survey administration ranges. Although the initial research protocol called for participants in all three conditions to complete surveys at three points during

the school year (baseline/T1 = Fall 2015, T2 = Winter 2016, T3 = Spring 2016), the actual survey administration dates varied by school. Survey administration and collection relied on a coordinated team effort by juvenile department staff, school administrators, teachers, and intervention group facilitators adding this responsibility to their pre-existing workloads. Additionally, groups started and ended at different times due to facilitator availability and differing academic calendars. For these reasons, some flexibility was provided to allow for differing survey administration and intervention start dates.

In its raw form, when the data were presented to the research team by the community partners, baseline (T1) survey dates ranged from September 7, 2015 to February 2, 2016 (with one case dated in April). T2 dates ranged from December 11 to June 17, 2016. T3 dates ranged from May 14 to June 16, 2016. For participants who completed all three surveys (n = 100), T1 dates ranged from September 28 to November 30, 2015, T2 dates ranged from December 11, 2015 to March 31, 2015, and T3 dates ranged from May 7 - June 18, 2016. These ranges were considerably narrower than the time point ranges for the sample as a whole. For analysis purposes, a decision rule was applied by the research team to account for the wide variety in survey administration dates for participants who completed two or fewer surveys, while also more accurately assessing individual change over time over the course of the academic year. For participants who completed two or fewer surveys, a survey was coded as T1 if it was dated from September - December 2015, T2 if dated January - March 2016, and T3 if dated April - June 2016. These date ranges were chosen based on their alignment with assessment and intervention dates of participants completing all three surveys and by their alignment with Fall/Winter/Spring term ranges of the school district.

Attendance, fidelity, and school records data. In addition to surveys completed by youth participants, data were also collected from Girls Circle and The Council group facilitators and participating schools. Facilitators completed weekly attendance logs and fidelity checklists each week that groups were running. School records were obtained from school administrators at the end of the 2015-2016 academic year. Data gathered from school records contained information about the participants' attendance, discipline referrals, suspensions/expulsions, academic performance, classroom behavior, and social/peer relations.

Demographic information. Demographic information was provided by participant self-report at each survey administration period. Baseline responses were typically used in analysis unless data were missing at baseline but provided at a later administration date.

Age. Participants reported their age in years, as well as birth month and birth year.

Hispanic or Latino/a. In response to the question, "Do you consider yourself Hispanic or Latino/a?" participants selected one of the following options; yes, no, refuse to answer.

Race. Participants selected one of the following options regarding race: American Indian/Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Black or African American, White, More than one race, Unsure or refuse to answer.

Gender identity. Participants selected one of the following options regarding gender identity: male, female, transgender, unsure, other.[1]

[1] As noted in the *Participants* section, for the purposes of randomized assignment to single-gender intervention groups, all participants were also identified by school facilitators as either male or female.

Sexual orientation. Participants selected one of the following options regarding sexual orientation: heterosexual or straight, gay or lesbian, bisexual, unsure/questioning, other.

Dosage and fidelity. Girls Circle and The Council facilitators completed weekly attendance logs and fidelity checklists for intervention and waitlist-control groups.

Dosage was measured by the total number of group sessions attended by each student.

Fidelity was measured by checklists provided by the One Circle Foundation. Facilitators marked "yes" or "no" next to each item to indicate whether specific program elements and activities occurred during each session. Checklist items included completion of each step in the Girls Circle/The Council format (e.g., opening ritual, theme introduction, check-in with use of talking piece, activity, closing ritual) as well as facilitation elements (e.g., fostered space that is emotionally, culturally and physically safe, engaged girls in critical thinking and decision-making process, learning was transferred to real life circumstances, etc.). Each "yes" response was coded as a 1 and totaled, with higher scores indicating higher intervention fidelity. Girls Circle fidelity checklists included 13 items; scores for each Girls Circle group ranged from six to 13 (α = .60). The Council fidelity checklists included 15 items; scores for each The Council group ranged from nine to 15 (α = .66).

Outcome variables. Self-efficacy, prosocial skills, perceived social support, and body satisfaction scores were measured by pre-existing scales included in the self-report paper surveys that were administered to youth participants three times over the course of the academic year. These four variables were measured by scales that have been widely

used in studies assessing adolescent samples and have demonstrated sensitivity to change in previous studies (Berscheid, Hatfield, & Bohrnstedt, 1972; Orpinas, 1993; Schwarzer & Jerusalem, 1995; Zimet, Dahlem, Zimet, & Farley, 1998). Cronbach's alpha (α) values for each self-report measure are reported for the full sample, as well as for boys and girls separately. Overall, α values for these variables indicate good internal reliability at baseline, T2, and T3.

The fifth outcome variable included in this study, school engagement, was also assessed via youth self-report. It was intended to be measured by four items from the One Circle Foundation's larger school engagement survey, which is adapted from the Colorado Foundation for Families and Children (2006) school engagement measure. However, due to face validity and internal reliability concerns (see process evaluation results for further discussion), it was dropped from statistical analysis. Instead, school engagement was measured by a construct created from school records items.

Self-efficacy was measured with the 10-item Schwarzer's Self-Efficacy Scale (Schwarzer & Jerusalem, 1995), which measures youth's self-efficacy both in dealing with everyday situations as well as more stressful events. It is comprised of statements someone might use to describe themselves and encourages youth to rate their agreement with that statement from not at all true to exactly true. Scale scores were based on 10 summed items from a four-point scale (0-3), with higher scores indicating higher self-efficacy. Sample scores on this measure ranged from 0-30 at each time point. At baseline, full sample α = .91 (girls α = .92, boys α = .89), T2 full sample α = .93 (girls α = .92, boys α = .95), and T3 full sample α = .93 (girls α = .90, boys α = .95).

Prosocial skills were measured by the eight-item Modified Aggression Scale (Orpinas, 1993), which measures youths' frequency of engaging in various prosocial behaviors, such as helping someone stay out of a fight or cooperating with peers. Scale scores were based on summed items on a four-point scale (0-3), with 0 indicating zero times engaging in each prosocial behavior to three indicating the youth engaged in the behavior five or more times; higher scores indicated greater levels of care and cooperating behavior. Of note is that the standard version of this scale questions youth as to their frequency of prosocial behaviors over the course of one week. However, youth in this study were instead asked how often they engaged in each scale item over the course of 30 days. Sample scores ranged from 0-22 at baseline and 0-24 at T2 and T3. At baseline, full sample $\alpha = .74$ (girls $\alpha = .70$, boys $\alpha = .77$), T2 full sample $\alpha = .83$ (girls $\alpha = .80$, boys $\alpha = .86$), and T3 full sample $\alpha = .80$ (girls $\alpha = .78$, boys $\alpha = .77$).

Perceived social support was measured by the 12-item Multidimensional Scale of Perceived Social Support (Zimet et al., 1998), which assesses youth perceptions of social support from friends and family. Survey items included statements about supportive people, and youth were prompted to select the level to which they agreed or disagreed with those statements. Examples of these statements included "There is a special person who is around when I am in need" and "My family is willing to help me make decisions." Scale scores were based on a seven-point scale (1-7), with higher scores indicating higher levels of perceived social support. Sample scores ranged from 12-84 at all time points. At baseline, full sample $\alpha = .95$ (girls $\alpha = .95$, boys $\alpha = .95$), T2 full sample $\alpha = .96$ (girls $\alpha = .96$, boys $\alpha = .97$), and T3 full sample $\alpha = .96$ (girls $\alpha = .95$, boys $\alpha = .96$).

Body satisfaction was measured using a modified version of the Body Parts Satisfaction Scale (Berscheid et al., 1972), which prompts youth to report the valence of their feelings toward different aspects of their physical appearance. Although typically assessed on a 6-point scale, with higher scores representing higher degrees of dissatisfaction with 24 different body aspects (Frederick, Hatfield, Bohrnstedt, & Berscheid, 2014), the version used in this study was reduced to 11 items. Participants were assessed on a 5-point scale (0-5), with higher scores indicating more positive feelings associated with one's own physical appearance. The 11 body aspects assessed in this study were muscular strength, biceps, buttocks, chest/breasts, hips, legs, appearance of stomach, face, weight, height, and overall body image. These items were selected from the original measure because they were identified by the research team as being most salient to adolescents and to both males and females. Sample scores ranged from 0-44 at all time points. At baseline, full sample α = .93 (girls α = .92, boys α = .95), T2 full sample α = .95 (girls α = .94, boys α = .95), and T3 α = .97 (girls α = .96, boys α = .98).

School engagement was measured by the mean score of three items included on the school records form. School administrators rated how well each student was doing with (item 1) classroom behavior, (item 2) social/peer relations, and (item 3) their overall impression of the student's school year. A 5-point scale (1-5) was used for each item, with higher scores indicating better administrator impressions of the youth's performance in each area. Scores for this measure ranged from 1-5 (α = .88; girls α = .87, boys α = .85), with higher scores indicating higher school engagement.

Unlike youth self-report measures, school records were only collected once during the course of the study (at T3; i.e., Spring 2016). Thus, this measure of school

engagement reflects administrator-report impressions for the 2015-2016 academic year overall based on their review of each student's file.

CHAPTER IV

ANALYTIC APPROACH

Overview

Demographic information, outcome variables, school records, attendance logs, and fidelity checklist data were compiled and analyzed using SPSS Statistics 23 (IBM Corp., 2013). Due to significant missingness of student self-report surveys at each time point, missing data analyses were conducted to assess for differential missingness by demographic characteristics, school attended, condition, and baseline scores on outcome variables. Descriptive statistics and bivariate correlations were conducted for participants in all conditions at T1, T2, and T3. Then, in order to assess for any baseline differences between conditions, χ^2 tests were conducted for categorical variables and Kruskal-Wallis tests for continuous variables.

Process evaluation components assessing adherence, dosage, and quality of program delivery were reported based on data collected from fidelity checklists and attendance logs. Outcome evaluation components were addressed next. For outcome variables measured at three time points via youth self-report (self-efficacy, perceived social support, pro-social skills, and body satisfaction), latent growth models were conducted using structural equation modeling techniques in Amos 23 (Arbuckle, 2014). Analysis of covariance (ANCOVA) general linear models were applied to address potential between-group differences in school engagement, the outcome variable that was only measured at one time point.

Missing Data

Item and scalar missingness for self-report measures. All survey scales included in statistical analysis had eight or more items. A scale score was computed if 20% or fewer of the items that composed the scale were missing. Because all survey scales used summed scores, the total of the items that were present was multiplied by the number of items in the scale divided by the number of non-missing items in the scale. There were no significant differences in item or scale missingness by demographic characteristics or condition. Item and scale missingness varied equally across measures and are assumed to be missing at random.

Full survey missingness for self-report measures. Of the 309 study participants originally identified by school personnel, 296 completed at least one survey. Two hundred fifty-two participants completed baseline surveys, 153 completed surveys at T2, and 191 completed surveys at T3. Ninety-nine participants completed all three surveys, 102 completed a total of two surveys, and 95 completed a total of one survey over the course of the study. Following an intention-to-treat (ITT) approach, all participants were included in analyses based on assigned condition regardless of levels of survey completion, group attendance, intervention fidelity, and missing data.

Overall, 13 students opted out of all waves of data collection, 27 opted out of T2, and 29 opted out of T3 (but may still have completed the intervention, as all were in the waitlist-control group). School administrators indicated that approximately 22 students with missing data at T2 and/or T3 moved or graduated, and two dropped out of school. The remainder of missing surveys were marked “not received” by juvenile department

staff (who collected and compiled completed surveys for distribution to the research team).

Because there was a large amount of missing self-report data (210 participants missing one or more surveys), a thorough reporting of missingness by demographic information, school, condition, and baseline scores on outcome variables is provided below. First, χ^2 tests were conducted to assess for between-group differences for participants who completed all three surveys and participants for whom at least one time point of survey data was missing. As shown in Table 2, data were not missing completely at random. Significant differences in survey missingness were observed based on school attended, school type (middle school vs. high school), and condition.

Differences in survey missingness by school and school type. All schools had some survey missingness. However, one middle school (School 1) and one high school (School 6) combined accounted for 87.9% of participants from whom all three surveys were received. The remaining 12 participants with three submitted surveys attended the other two middle schools in the study, and the two other high schools did not have any participants complete all three surveys. Overall, there was significantly more survey missingness in high schools (36 high schoolers completed all surveys, 143 were missing one or more survey) than in middle schools (63 middle schoolers completed all surveys, 67 were missing one or more survey). It is likely that survey missingness varied between schools due to differences in how each school approached data collection for the study. As part of the evaluation design, schools were responsible for their own data collection, and the research team and juvenile department did not directly oversee how and when surveys were administered. There may have been some confusion in how many surveys

Table 2

Categorical values attrition analysis of differences between participants completing surveys at all three time points (n = 99) and participants missing one or more surveys (n = 210)

		All surveys received		1 or more survey missing		χ^2 (df)
		N	%	*N*	%	
Gender						
Male		42	42.4	96	45.7	.295(1)
Female		57	57.6	114	54.3	
Ethnicity						
White		62	62.6	114	54.3	6.125(3)
Multiethnic		16	16.2	29	13.8	
Other		7	7.1	11	5.2	
Missing/refuse to ans.		14	14.1	56	26.7	
Hispanic/Latino/a						
Yes		14	14.1	29	13.8	.077(2)
No		61	61.6	127	60.5	
Missing/refuse to ans.		24	24.2	54	25.7	
Sexual Orientation						
Heterosexual/straight		76	76.8	152	72.4	5.432(2)
Bisexual		11	11.1	13	6.2	
Other		12	12.1	45	18.4	
School						
Middle Schools	School 1	51	51.5	19	9.0	145.461(5)***
	School 2	3	3.0	32	15.2	
	School 3	9	9.1	16	7.6	
High Schools	School 4	0	-	79	37.6	
	School 5	0	-	43	20.5	
	School 6	36	36.4	21	10.0	
Middle Schools		63	63.6	67	31.9	27.797(1)***
High Schools		36	36.4	143	68.1	
Condition						
Intervention		35	35.4	36	17.1	12.854(2)**
WC		23	23.2	69	32.9	
Control		41	41.4	105	50.0	
Intervention or WC		58	58.6	105	50.0	1.990(1)
Control		41	41.4	105	50.0	

Note. $^{**}p < .01$, $^{***}p < .001$. WC = Waitlist-Control.

were intended to be given over the course of the school year, unclear delegation of responsibility for distributing and collecting surveys, or other challenges related to facilitating data collection amidst other priorities in school administration over the course of the academic year.

Differences in survey missingness by condition. Significant differences in survey missingness were also observed between study conditions. As shown in Table 2, participants in the waitlist-control condition and control condition were more likely to be missing at least one survey compared to participants assigned to participate in Girls Circle or The Council in the Fall. However, no significant differences in missingness emerged between participants assigned to Girls Circle or The Council and participants assigned to services as usual; participants in the waitlist-control and intervention conditions were equally likely to be missing at least one survey as participants in the control condition.

Further examination of missingness for each condition indicated trends in missingness by time point. At T1, 11% of the control group, 45% of the waitlist-control group, and 0% of the intervention group did not complete a survey. At T2, 65% of the control group, 46% of the waitlist-control group, and 42% of the intervention group did not complete surveys. At T3, 40% of the control group, 43% of the waitlist-control group, and 28% of the intervention group did not complete surveys. These trends suggest that there was some confusion about the evaluation design specifically related to (a) how many surveys to distribute overall and (b) how survey administration related to condition. For example, that 45% of the waitlist-control group did not complete T1 surveys may indicate that some administrators believed surveys for Girls Circle and The Council

participants were only meant to be distributed immediately prior to the intervention being received (i.e., that a pre-post-test was being conducted for intervention and waitlist-control participants instead of a three-time point design regardless of condition).

Another potential explanation for some of the differences in survey missingness by condition may be due to who was responsible for collecting student self-report data. Intervention and waitlist-control group surveys tended to be administered by facilitators, who worked closely with small groups of Girls Circle/The Council participants over the course of 10 weeks, with a built-in meeting time to hand out surveys following the completion of the program. In contrast, survey administration for control group participants was conducted by school administrators or teachers with multiple other responsibilities and timelines.

Differences in survey missingness by baseline scores. As demonstrated in the previous sections, participant self-report data were not missing completely at random and significant differences emerged between schools, school type, and condition. In order to better assess whether data were missing at random (MAR) or missing not at random (MNAR), independent *t*-tests were run to determine if baseline scores on outcome variables predicted missingness at T2 and/or T3. MAR means that there is a systematic relationship between the likelihood of missingness and the observed data, but not the missing data. For example, that middle schoolers were more likely to complete all three surveys than high schoolers implies MAR. In contrast, MNAR means that there is a relationship between the likelihood of data to be missing and the missing values themselves. In this case, if participants with lower baseline scores on outcome variables

were then less likely to complete surveys including those variables at later time points, this would be considered MNAR.

Independent samples *t*-tests were run with T1 scores on each outcome variable entered as the dependent variable and between-group comparisons made between respondents and non-respondents at each time point. As shown in Table 3, no significant differences in missingness emerged at either T2 or T3 related to participant baseline scores on outcome variables.

Table 3

Analysis for potential differences between participant baseline scores and later survey missingness

	M	*SD*	*t*-value
Baseline self-efficacy			
T2 survey submitted	19.50	6.64	.073
T2 survey missing	19.45	5.47	
T3 survey submitted	19.33	6.27	-.473
T3 survey missing	19.71	5.64	
Baseline perceived social support			
T2 survey submitted	60.32	18.32	-.814
T2 survey missing	62.10	15.85	
T3 survey submitted	61.51	17.47	.290
T3 survey missing	60.85	16.40	
Baseline prosocial skills			
T2 survey submitted	10.82	4.46	-.121
T2 survey missing	10.89	4.96	
T3 survey submitted	10.68	4.50	-.735
T3 survey missing	11.13	5.09	
Baseline body satisfaction			
T2 survey submitted	25.53	10.82	1.70
T2 survey missing	23.34	8.98	
T3 survey submitted	23.98	10.32	-.795
T3 survey missing	25.03	9.23	

Note. No significant differences between respondents and non-respondents.

Although the possibility of MNAR cannot be ruled out completely, based on patterns of missingness discussed in this section, it is most likely that missing data were due to observed factors such as school and condition rather than unobserved factors.

Thus, for the purposes of analysis, data were considered MAR, justifying the use of Full Information Maximum Likelihood (FIML) estimation in quantitative analyses.

CHAPTER V

RESULTS

Descriptive Statistics

Descriptive statistics are provided in Table 4. Overall, levels of self-efficacy, perceived social support, prosocial skills, and body satisfaction generally increased from baseline (T1) to the end of the school year (T3). Data for each outcome variable were normally distributed at each timepoint, with skewness values between -1 to 1.

Table 4

Means, standard deviations, and ranges of outcome variables

Outcome variables	*M*	*SD*	*N*	Min	Max
Self-efficacy [0-30]					
T1	19.47	6.03	246	0	30
T2	19.61	6.17	153	0	30
T3	19.87	5.89	189	0	30
Perceived social support [12-84]					
T1	61.26	17.04	244	12	84
T2	63.46	19.12	151	12	84
T3	63.50	17.36	190	12	84
Prosocial skills [0-24]					
T1	10.86	4.73	249	0	22
T2	11.45	5.36	153	0	24
T3	11.77	5.17	188	0	24
Body satisfaction [0-44]					
T1	24.37	9.93	241	0	44
T2	25.68	10.67	151	0	44
T3	25.09	11.55	189	0	44
School engagement [1-5]					
T3[a]	3.45	1.12	192	1	5

Note. Brackets indicate possible scale range; higher scores indicate greater levels of each outcome variable. T1 = Fall 2015, T2 = Winter 2016, T3 = Spring 2016. *M* = mean. *SD* = standard deviation. *N* = number of participants completing enough items of each scale to calculate a scale score. Min/Max = minimum/maximum score within the sample at each time point. [a]School engagement was reported by school administrators at T3 only and reflects school engagement for the 2015-2016 academic year overall.

Bivariate correlation values for outcome measures at T1, T2, and T3 are provided in the Appendix. Correlations were generally significant, moderate, and in the expected positive direction.

Process Evaluation Components

At baseline, a total of 163 participants (92 girls, 71 boys) were assigned to either intervention or waitlist-control condition (i.e., assigned to participate in Girls Circle or The Council at some point during the academic year). Process elements (adherence, dosage, and quality of delivery) for Girls Circle and The Council were assessed via attendance logs and fidelity checklists completed by group facilitators. Attendance logs and fidelity checklists were received from 12 group; seven were from Girls Circle groups (four intervention condition, three waitlist-control) and five were from The Council groups (one intervention condition, four waitlist-control). The combined rosters from these 12 groups from data were received totaled 95 group participants (52 girls, 43 boys); there were no attendance or fidelity data available for 42 percent of participants assigned to participate in a group condition. In communication with juvenile department staff assisting with tracking data collection from facilitators, the research team learned that one of the Girls Circle waitlist-control groups had been cancelled, and that 10 female participants that had originally been assigned to the waitlist-control condition instead received services as usual. There was no indication that other groups had not run as intended; it is most likely that the remainder of fidelity records/attendance logs missingness is attributable to facilitator non-completion.

Overall, available group rosters reflected adherence to randomly assigned condition. Participants assigned to receive Girls Circle or The Council in either the Fall

(intervention condition) or Winter (waitlist-control condition) received the appropriate gender-specific intervention at the specified time. Exceptions were the following: (a) two boys originally assigned to the control condition instead participated in waitlist-control groups of The Council, and (b) three boys originally assigned to the intervention condition instead participated in a waitlist-control The Council group. In keeping with an ITT approach, in outcome analyses, participants were evaluated based on originally assigned condition.

Goal 1: Adherence. In order to assess the extent to which Girls Circle and The Council were delivered as designed, rosters and fidelity logs were reviewed to determine (a) whether programs were delivered with all core components to the appropriate populations and, (b) whether the correct protocols, techniques, and materials were used.

Of note, all available fidelity checklists from Girls Circle groups included data for nine (instead of 10) sessions. Attendance logs verified that the expected 10 sessions were in fact held for each of these groups. Although it is unknown why Girls Circle groups consistently were missing fidelity checklists from the tenth session, it is possible that because paper fidelity logs were distributed to facilitators, the Girls Circle versions of these logs were missing the final page.

Objective (a): Determine whether programs were delivered with all core components to the appropriate populations. Based on available data, overall program adherence seemed good for both Girls Circle and The Council. Group rosters reflected gender-specific groups appropriate to condition and at least one same-gender adult facilitator. Facilitators remained constant throughout the course of each group. Participant age ranges were appropriate to the curriculum (ages 11-19). Group size was also

observed to be consistent with what was intended (six to 10 participants); only one group fell outside of this range by having only four participants. In 92% of all sessions, group agreements were created, posted, and clearly visible and exceptions to confidentiality were reviewed. Fidelity checklist mean scores were close to the maximum possible for both interventions; Girls Circle groups *M(SD)* = 12.32 (1.35) out of a possible 13 points and The Council groups *M(SD)* = 14.04 (1.51) out of a possible 15 points.

Objective (b): Determine whether the correct protocols, techniques, and materials were used in the correct locations/contexts. Fidelity checklist items assessing completion of each step in the Girls Circle/The Council format were generally marked as present in most sessions. 96% of all sessions included an opening ritual, 97% included theme introduction, 96% included a check-in with use of the talking piece, 97% included the activity component, and 90% included a closing ritual. Girls Circle also included a sharing of the activity component (present in 92% of sessions). The Council included two additional elements: warm-up (present in 90% of sessions) and reflection (present in 86% of sessions).

Goal 2: Dosage. Dosage was assessed by review of available attendance logs. Attendance data were used to calculate both program coverage and individual dosage.

Objective (c*): *Determine the number of Girls Circle and The Council attended by participants in each condition. Of the 95 participants for whom attendance data was collected, 85% attended seven or more sessions of Girls Circle or The Council. Only nine participants in either condition completed three or fewer sessions. The average participant attended eight sessions (*SD* = 2.5).

Dosage was also examined at the population level. Program coverage is the extent to which the target population achieves the full scope of the program specified. It is calculated as the ratio of total number of program sessions received to the total number of program sessions expected. Because of the amount of missingness of attendance logs from groups that juvenile department staff indicated actually did run, target population in this instance was calculated by the number of participants listed on available attendance rosters (n = 95) plus the number of participants known to have been originally assigned to the intervention condition but who instead received services as usual (n = 10). Program exposure was expected to cover 1,050 sessions (105 participants x 10 sessions); actual coverage was 760 sessions. Program coverage was estimated to be 72% (760 sessions/1,050 sessions).

Objective (d): Determine the length of treatment provided. An additional dosage consideration is whether groups were initiated and concluded within a reasonable time frame. The interventions were designed to be conducted via weekly sessions, so range of time from beginning to end was also important to consider. Overall, most groups began and ended in approximately 10 weeks. Nine of the 12 groups began and ended within 12 weeks. Three The Council waitlist-control groups, however, each had one multi-week gap between sessions. One of these groups had had a three-week gap between sessions five and six, another had a four-week gap between sessions seven and eight, and the third group had a six-week gap between sessions two and three.

Goal 3: Quality. In order to assess the quality of program delivery, the facilitation elements of the fidelity checklist were assessed, along with facilitator impressions of Girls Circle and The Council groups.

Objective (e): Evaluate the extent to which facilitation elements were achieved. Facilitation elements included the following: (a) adequate preparation; (b) fostered space that is emotionally, culturally and physically safe; (c) engaged boys/girls in critical thinking and decision-making process; (d) utilized open-ended questions, reflections, demonstrated spirit of motivational interviewing; (e) strengths-based; (f) learning was transferred to real life circumstances; (g) effectively engaged boys/girls in managing difficult group dynamics. The Council also included an addition facilitation element: (h) normalizes boys' emotions – resists the "boy code" (act tough, be cool, don't cry). Facilitators self-reported "yes" or "no" whether they had met these fidelity components. Each "yes" was coded as 1 and each "no" as 0. For Girls Circle groups, out of a possible scale range of 0-7, with higher scores indicating higher quality facilitation, the mean score for facilitation elements was 6.46 (*SD* = 1.34). For The Council groups, out of a possible scale range of 0-8, the mean score for facilitation elements was 7.34 (*SD* = 1.41). Overall, facilitators of both Girls Circle and The Council groups tended to report that they had included most intended facilitation elements in the majority of sessions.

Objective (f): Evaluate the degree to which facilitator impressions of the Girls Circle and The Council programs were positive. As a part of the fidelity checklist completed by facilitators each session, there was a section for open-ended comments with the prompt "Describe or note any changes or relevant information about this session." Overall, facilitators tended to comment on challenges and successes that arose from group sessions, as well as techniques they used to address issues. Themes that arose from facilitator impressions for both Girls Circle and The Council included the following:

1. Students enjoyed being a part of the group.

2. Curriculum elements were presented as specified in manuals.
3. Challenging group dynamics arose and were addressed by facilitators.
4. It was hard to fit in all elements of each session during the time allotted as a part of the school day.
5. Challenges to adherence and dosage included altered school schedules (e.g. in-service days, assembly schedules, winter/spring breaks).

Other themes that arose from facilitator impressions of Girls Circle groups included:

1. The girls engaged in meaningful discussions.
2. Role plays were effective teaching elements and girls enjoyed them.
3. It was helpful to receive help from school staff to manage overly disruptive/aggressive behavior.

Additional themes specific to The Council facilitator impressions included:

1. Middle school boys seemed to struggle to maintain appropriate behavior during group discussions.
2. It was sometimes difficult to maintain emotional safety in the groups.
3. Boys "had a lot of energy" and responded well to warm-up activities that included physical movement (e.g., musical chairs).
4. Boys were initially reluctant to engage in group if it seemed "like a therapy group."
5. At the closing of the group, boys were able to share self-identified strengths.

Outcome Evaluation Components

Testing for pre-intervention differences between conditions. The intervention, waitlist-control, and control groups were compared on demographic and outcome

variables to test for pre-intervention differences using χ^2 tests for categorical variables and Kruskal-Wallis tests for continuous variables. Categorical demographic variables with more than two levels (e.g., ethnicity and sexual orientation) were dichotomized by level (e.g., percent White, percent multiethnic, etc.) before χ^2 tests were applied in order to be able to compare between three groups. No significant ($p < .05$) between-group differences were found for ethnicity, sexual orientation, or age. There were also no significant between-group differences for survey outcome values (self-efficacy, prosocial skills, perceived social support, and body satisfaction).

The groups differed significantly on gender ($\chi^2(2) = 12.81$, $p = .002$). Boys were over-represented in the waitlist-control condition and under-represented in the control condition. In order to determine if there were significant differences in gender between the control group and participants that received the intervention at any time point, the control group was then compared to the combined intervention and waitlist-control groups. There remained significant gender differences between the control and combined waitlist-control/intervention groups ($\chi^2(1) = 7.82$, $p = .005$). Female participants were mostly evenly distributed; of the 171 female participants, 93 (54.4%) were assigned to the control condition and 78 (45.6%) were assigned to participate in Girls Circle. However, male participants were more likely to be assigned to participate in The Council than to receive services as usual; Of the 138 male participants, 53 were assigned to the control condition (38.4%) and 85 (61.6%) were assigned to the participate in The Council.

Growth models assessing psychosocial assets. For growth model analyses, structural equation modeling (SEM) techniques were conducted. Advantages of SEM

approaches include flexible estimation methods, correct standard errors, ability to include constant and time-varying covariates, and effective handling of missing data (Kline, 2016). Latent growth curve models were estimated using full information maximum likelihood (FIML) estimation to reduce bias resulting from missing data, assuming those data are missing at random (Wothke, 2000). Standard measures of fit are reported, including χ^2 values, Hoelter Index values (to determine if significant chi-square tests are due to large sample size), comparative fit index (CFI), and the root mean square error of approximation (RMSEA). Non-significant χ^2 values (or Hoelter Index values greater than 200), CFI values greater than .95, SRMR values less than .08, and RMSEA values less than .08 indicate good model fit (Hu & Bentler, 2009). Of note, standardized root mean square residual (SRMR) were intended to be included in analysis, but could not be evaluated due to missing values.

Because of the nested nature of these data (student participants within schools), hierarchical growth modeling was also considered as an alternative to SEM. However, because randomization occurred at the individual level (i.e., eligible participants within each school were randomly assigned to condition) and because of the low number of schools (six), significant variance was not expected to be explained by school-level factors. For many situations, fitting growth models within multilevel modeling frameworks and SEM frameworks yields similar results (Curran, Obeidat, & Losardo, 2010). It is also commonly suggested that an Intraclass Correlation Coefficient (ICC) values smaller than 5% indicate that multilevel modeling is unnecessary (Bliese, 2000). To test the assumption that there was not significant variance explained by between-school differences, a null model was fitted and ICCs were examined. Only a very small

percentage of the total variance in the outcomes was found systematically between schools; ICCs for each outcome variable used in growth modeling reflected less than 5% of variance explained by between school differences (ICCs for self-efficacy = .023, perceived social support = .038, pro-social skills, .045, and body satisfaction = .022).

Goal 4: Overall effectiveness of offering groups in schools. In order to assess the effectiveness of participation in either Girls Circle or The Council in comparison with school services as usual, main effects growth models were conducted for the four psychosocial assets (self-efficacy, perceived social support, prosocial skills, and body satisfaction), and ANCOVAs were conducted for the school engagement outcome variable.

Objective (g): Determine whether intervention and/or waitlist-control participants demonstrated improved psychosocial assets compared to control group participants. Preliminary unconditional growth models were estimated for the four psychosocial asset outcome variables measured at three timepoints (self-efficacy, perceived social support, prosocial skills, and body satisfaction). There were statistically significant ($p < .05$) variance for the intercepts and slopes, suggesting that it was appropriate to estimate conditional models (Duncan, Duncan, & Strycker, 2006). First, as shown in Figure 2, models were run to assess for significant main effects of participation in Girls Circle/The Council on each of the four outcome variables. For the "Intervention Group" variable, participants assigned to the intervention condition were dummy coded "1" and all others were "0". For the "Waitlist-Control Group" variable, participants assigned to the waitlist-control condition were dummy coded "1" and all others were "0". Gender was also dummy coded as "1" for boys and "0" for girls. Paths were estimated to

test whether random assignment to the intervention group was a significant predictor of the slope of individual trajectories in dependent variables (Path C), as well as whether random assignment to the waitlist-control group was a significant predictor of slope (Path D). These estimates controlled for possible age differences (Path A) and gender differences (Path B). This model was run first based on the presumption that there would likely be different effects on individual slopes based on the timing of receiving the intervention (Fall term for intervention group participants vs. Winter term for waitlist-control participants).

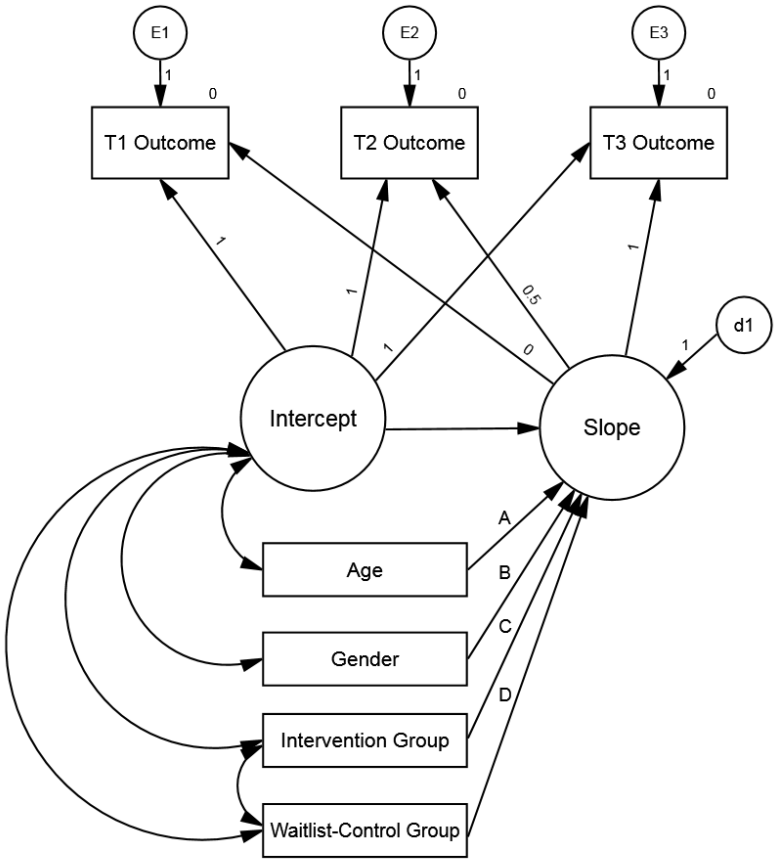

Figure 2. Main effects model predicting outcomes with intervention and waitlist-control groups considered separately.

Findings for the four main effects models run as shown in Figure 2 are summarized in Table 5. Estimated model fit was good (non-significant χ^2 or Hoelter Index > 200, CFI > .95, RMSEA < .08) for the models predicting self-efficacy, perceived social support, and body satisfaction growth. The model predicting pro-social skills had adequate fit (Hoelter Index = 248, RMSEA =.062), with a slightly lower CFI value of .922. No significant associations were found between assignment to the intervention or waitlist-control conditions and growth in any of the outcome variables. There were also no significant associations found for gender or age.

Table 5

Summary of model fit and path coefficients for preliminary main effects models

Outcome variable	Model fit indices				Paths (β values)			
	χ^2 (12)	Hoelter Index*	CFI	RMSEA	A Age	B Gender	C INT	D WC
Self-efficacy	17.279		.973	.038	.114	-.095	-.083	-.014
Perceived social support	21.469*	302	.966	.051	.091	.198	-.067	-.025
Prosocial skills	26.125*	248	.922	.062	-.336	-.028	-.237	.289
Body satisfaction[a]	21.527*	301	.959	.051	.424	-1.596	-.062	-1.646

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; INT = Intervention; WC = Waitlist-control. β = standardized regression coefficients. [a] Body satisfaction path values are unstandardized coefficients. *p < .05.

After having tested the first main effects models above, a second round of more parsimonious main effects models were run to determine if participation in Girls Circle or The Council (i.e., treatment at any time) was associated with more growth in the four outcome variables. For these analyses, participants assigned to either the intervention or waitlist-control conditions were dummy coded “1” and participants assigned to the control condition were dummy coded “0”. Figure 3 provides illustration of this second set of main effects models.

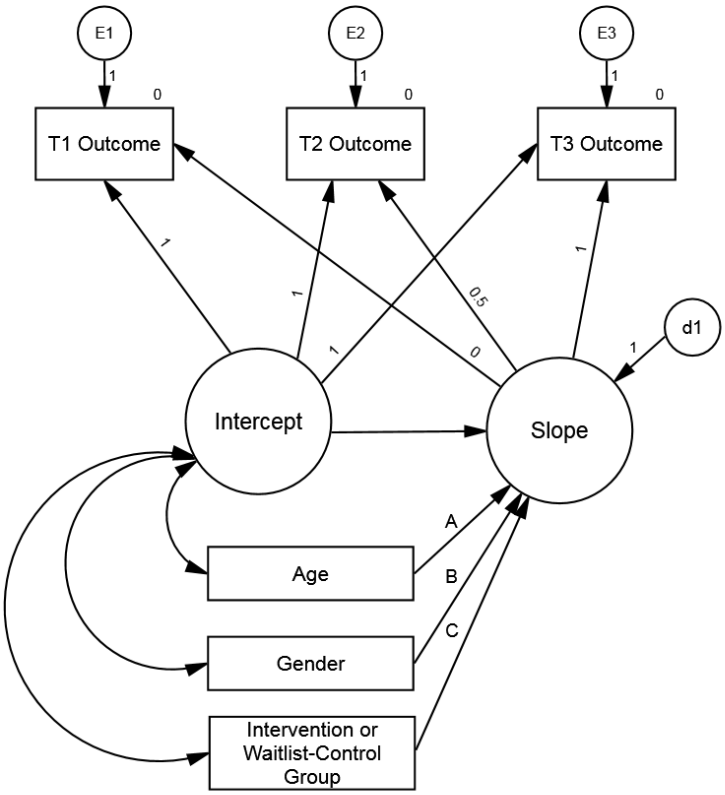

Figure 3. Main effects model predicting outcomes with intervention and waitlist-control groups considered together.

Findings for the four main effects models run as shown in Figure 3 are summarized in Table 6. As with the first round of main effects models, estimated model fit was good (non-significant χ^2 or Hoelter Index > 200, CFI > .95, RMSEA < .08) for the models predicting self-efficacy, perceived social support, and body satisfaction growth. The model predicting prosocial skills had adequate fit (Hoelter Index = 256, RMSEA =.064), with a slightly low CFI value of .917. Also similar to the previous main effects models, no significant associations were found between assignment to receive treatment at any time and growth in any of the outcome variables. There continued to be no significant associations found for gender or age.

Table 6

Summary of model fit and path coefficients for secondary main effects models

Outcome variable	Model fit indices				Paths (β values)		
	χ^2 (9)	Hoelter Index*	CFI	RMSEA	A Age	B Gender	C INT/WC
Self-efficacy	11.808		.982	.032	.116	-.092	-.055
Perceived social support	16.224		.970	.051	.096	.199	-.039
Prosocial skills	20.432*	256	.917	.064	-.322	.021	-.024
Body satisfaction[a]	14.794		.970	.046	.411	-1.685	-.714

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; INT/WC = Intervention/waitlist-control. β = standardized regression coefficients. [a] Body satisfaction path values are unstandardized coefficients. *$p < .05$.

Objective (h): Determine whether intervention and/or waitlist-control participants differed from control group participants in school engagement. Because school engagement was measured at only one time point (overall impressions from school administrators provided at the end of the academic year), growth models could not be used to assess change over time. Instead, main effects were estimated with ANCOVA models. Age was entered as a covariate and school engagement as the dependent variable. Condition (both treatment at any point and waitlist control and intervention conditions considered separately) was entered as an independent variable. No significant effects were found for condition.

Goal 5: Effectiveness of Girls Circle and The Council program models.

Although gender was included in the main effects growth models as a covariate, these models did not account for potential differences in growth trajectories based on whether a participant was assigned to The Council or Girls Circle. Thus, even though no significant intervention effects were found using the main effects analyses, there remained the possibility that intervention effects occurred, but differed by intervention type. In order to assess for differences in effectiveness by program type (i.e., intervention effects may

differ between Girls Circle and The Council), interaction terms for gender and condition were added to both growth models (psychosocial assets) and ANCOVAs (school engagement).

Objective (i): Evaluate whether group type (Girls Circle vs. The Council) is a significant predictor for improved psychosocial assets. Interaction models were constructed similarly to main effects growth models, with the addition of a (gender x condition) interaction term(s). Effects coding was used in place of dummy coding for categorical variables following best practices for measuring interaction effects (Daly, Dekker, & Hess, 2016). For all interaction models, "Gender (EF)" was coded as "1" for boys and "-1" for girls. For the "Intervention Group (EF)" variable, participants assigned to the intervention condition were dummy coded "1" and all others were "-1". For the "Waitlist-Control Group (EF)" variable, participants assigned to the waitlist-control condition were dummy coded "1" and all others were "-1". As shown in Figure 4, for the first series of interaction models run, paths were estimated to test for interaction effects between gender and group assignment at T1 (Path E) and T2 (Path F).

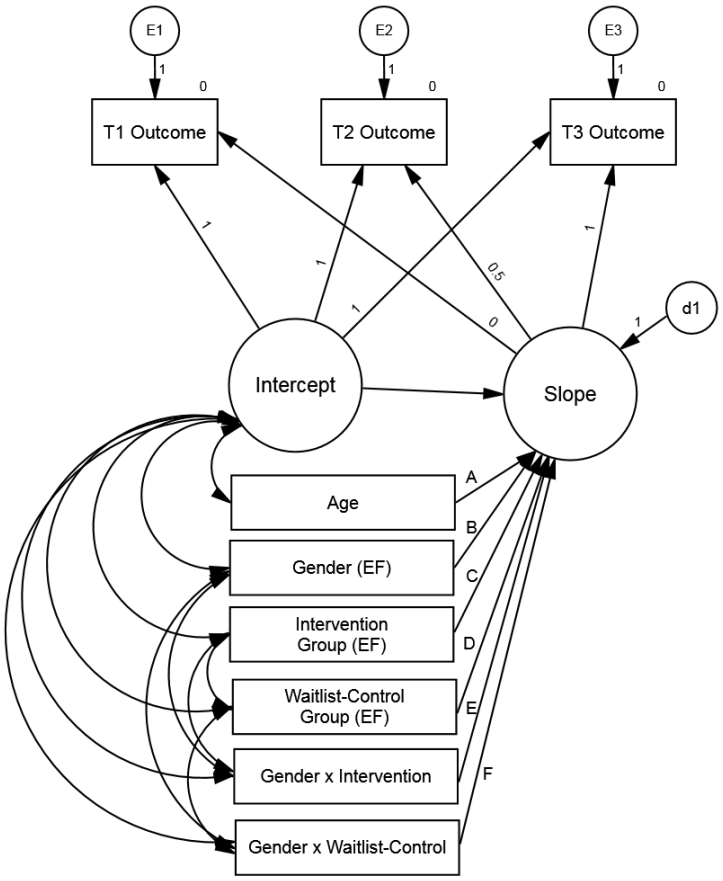

Figure 4. Interaction effects model predicting outcomes with intervention and waitlist-control groups considered separately. EF notation denotes effects coding.

Findings for the four main effects models run as shown in Figure 4 are summarized in Table 7. Estimated model fit was good (non-significant χ^2 or Hoelter Index > 200, CFI > .95, RMSEA < .08) for all models. There was a significant interaction between gender and waitlist-control assignment on self-efficacy change over time (β = .328, p = .016). This implies that the effect of waitlist-control group assignment on growth in self-efficacy differs based on youth gender. As shown in Figure 5, boys in the waitlist-control condition evidenced faster rates of growth in self-efficacy between T2 and T3 than participants in other conditions (ANOVA testing revealed no significant mean differences between groups for self-efficacy at T1 and T2). Because all boys in this

condition were assigned to participate in a The Council group between T2 and T3, this provides some evidence for The Council increasing rates of growth in self-efficacy. There is also the possibility that unobserved individual differences between facilitators may have contributed to this effect; The Council groups taking place in the Fall (intervention condition) were conducted by three facilitators who were different from the three facilitators conducting groups in the Winter/Spring (waitlist-control condition). Four of the five Girls Circle facilitators conducted groups in both intervention and waitlist-control conditions.

Table 7

Summary of model fit and path coefficients for preliminary interaction effects models

Outcome variable	Model fit indices				Paths (β values)					
	χ^2 (19)	Hoelter Index*	CFI	RMSEA	A Age	B Gender (EF)	C INT (EF)	D WC (EF)	E Gender x INT	F Gender x WC
Self-efficacy	29.808		.974	.043	.068	.075	-.062	.027	.038	.328*
Perceived social support	31.013*	300	.975	.045	.065	.291	-.080	-.025	-.155	.301
Prosocial skills	38.893**	239	.950	.058	-.352	-.233	-.301	.343	-.138	-.246
Body satisfaction[a]	37.830**	246	.958	.057	.388	-.451	.263	-.572	.883	.181

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; INT = Intervention; WC = Waitlist-control. β = standardized regression coefficients. [a] Body satisfaction path values are unstandardized coefficients. *$p < .05$; **$p < .01$.

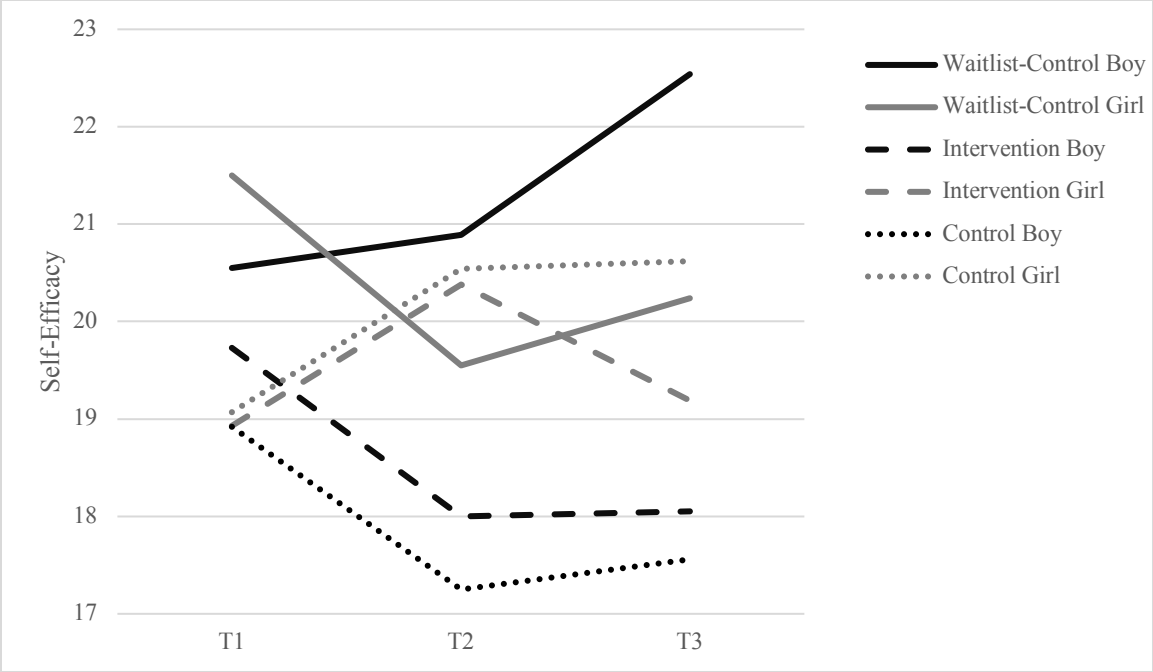

Figure 5. Self-efficacy mean scores by condition over three time points.

As with the main effects models, it was also of interest to determine if assignment to participate in Girls Circle or The Council at any point in the school year had an impact on outcome variable slopes. As shown in Figure 6, for the final set of growth models, an interaction term (gender x treatment at any time) was used to assess whether between-group differences emerged for participation in Girls Circle or The Council regardless of whether the intervention occurred in the Fall or Winter. Table 8 reflects good fit for each model, but no significant paths for predictor variables, including the interaction term.

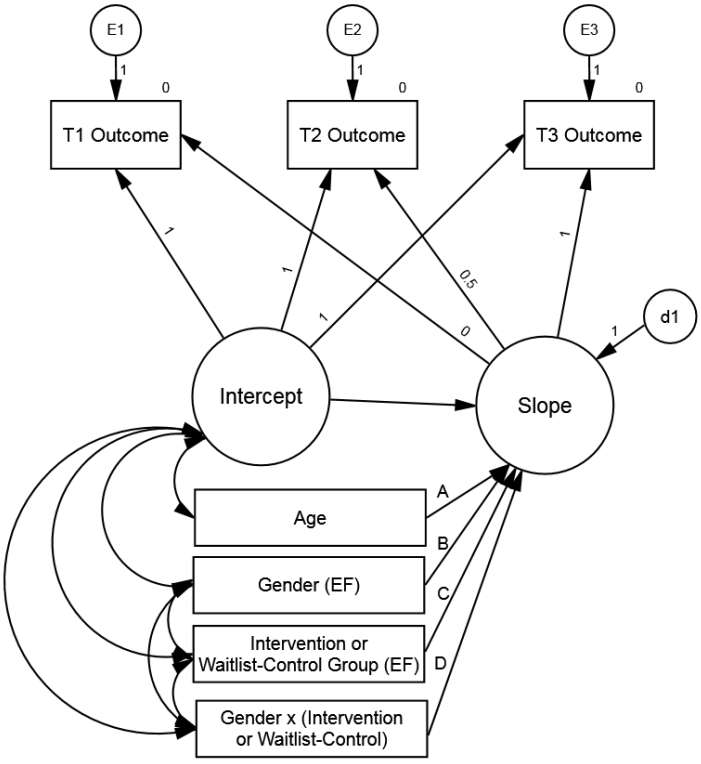

Figure 6. Interaction effects model predicting outcomes with intervention and waitlist-control groups considered together.

Table 8

Summary of model fit and path coefficients for secondary interaction effects models

Outcome variable	Model fit indices			Paths (β values)			
	χ^2 (10)	CFI	RMSEA	A Age	B Gender (EF)	C INT/WC (EF)	D Gender x INT/WC
Self-efficacy	8.536	.999	.001	.086	-.120	-.002	.173
Perceived social support	13.039	.988	.031	.089	.180	-.016	.034
Pro-social skills	17.329	.949	.049	-.331	.050	-.060	-.105
Body satisfaction[a]	18.266	.960	.052	.363	-.971	-.148	.531

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; INT = Intervention; WC = Waitlist-control. β = standardized regression coefficients. [a] Body satisfaction path values are unstandardized coefficients. No significant values ($p < .05$).

Objective (j): Evaluate whether group type (Girls Circle vs. The Council) is a significant predictor for school engagement. As with prior ANCOVA models, age was

entered as a covariate and school engagement as the dependent variable. Gender and condition (both treatment at any point and waitlist control and intervention conditions considered separately), as well as a condition x gender interaction term were entered as independent variables. No significant main or interaction effects were found. Significant effects were found for gender as a predictor of differences in school engagement ($F(1) = 18.997$, $p < .001$; $\eta_{partial}^2 = .103$). Girls were more likely to have higher school engagement scores (*M(SE)* = 3.81(.106)) than boys (*M*(*SE*) = 3.00(.120)), regardless of condition.

Goal 6: Dosage effects. The models described above used rigorous ITT approaches that evaluated outcomes based on assigned condition regardless of actual attendance in the Girls Circle or The Council programs. Because previous studies of these interventions have found significant dosage effects when no main effects were observed (Gies et al., 2015; Mankowski, 2011), post-hoc analyses were then conducted to assess for any significant dosage effects in the present study. Of note is that dosage measures (i.e., group attendance logs) were missing for 43 percent of participants assigned to participate in a Girls Circle or The Council group; this significantly limited any conclusions related to potential dosage effects.

Objective (k): Incorporate Girls Circle and The Council session attendance data into outcome analyses. First, in order to maintain the inclusion of control group participants in analysis, interaction terms (dosage x condition) were incorporated into growth models evaluating self-efficacy, prosocial skills, perceived social support, and body satisfaction. No significant main or moderation effects were found for dosage. The one significant interaction effect (boys in the waitlist-control condition showing higher

rate of increased self-efficacy) found in ITT analysis remained when dosage was added to the model ($\beta = .321$, $p = .017$), although overall model fit decreased: ($\chi^2(19) = 54.086$, $p < .001$, Hoelter Index = 172, CFI = .918, RMSEA = .077).

Next, in order to assess whether dosage predicted differences in outcomes for Girls Circle and The Council participants, control group participants were excluded from analysis and dosage was entered into growth models as a continuous independent variable with age as a covariate. Again, no significant effects were found.

Finally, a similar analytic progression was conducted for the ANCOVA models evaluating school engagement as a dependent variable. No significant moderating or main effects were found for dosage in these analyses.

CHAPTER VI

DISCUSSION

Summary of Findings

This study adds to the existing articles and program evaluations of Girls Circle and The Council by applying a randomized experimental design to the interventions' implementation in school-based settings. Previous studies evaluating these programs have solely focused on Girls Circle or The Council. This study is unique in that it not only considers both intervention types separately, but also evaluates their combined implementation in co-ed schools. Because co-ed schools are likely striving to provide equitable services across genders, it is important to consider the aggregate effect of offering gender-specific programming for both boys and girls.

Process evaluation components of this study provide justification for the feasibility of implementing Girls Circle and The Council in public school settings. Overall, both interventions were delivered with fidelity. Adherence, individual participant dosage, and quality of program delivery mean scores were typically close to the maximum possible scale values, as indicated by attendance logs and fidelity checklists completed by facilitators. Maintaining program fidelity within the context of diverse school settings throughout a full calendar year is a good indication that Girls Circle and The Council are appropriate to implement in school-based settings. It should also be noted that the evaluation context included support from the local juvenile department and investment from school administrators, which are likely key elements to implementation success. Schools considering adopting these interventions into their existing services will likely benefit from having a designated staff member or group facilitator serve as a skills

group coordinator (a role fulfilled by the juvenile department in this study) who is able to oversee facilitator training and scheduling, as well as assist schools in trouble-shooting potential challenges to implementation.

Facilitator comments indicated largely positive impressions of both Girls Circle and The Council. Group participants were able to engage in meaningful discussions and seemed to engage with the curriculum and each other. Overall, students tended to enjoy being a part of the group and facilitators self-reported being able to address challenging group dynamics while maintaining program adherence and quality.

In contrast to the generally positive findings related to process evaluation components, the outcome evaluation portion of this study yielded mostly null effects for both Girls Circle and The Council. Results did not provide additional validation for intervention effects found in previous evaluations of these programs. Girls Circle participation has been associated with increased body image, perceived social support, and self-efficacy (Irvine, 2005; Irvine & Roa, 2010; Steese et al., 2006). A previous evaluation of The Council showed improved school engagement following group participation (Gray et al., 2008). However, these prior significant results were found using pre-post designs without inclusion of a comparison or control group. In the present study, mean scores for the entire sample tended to increase over the course of the school year (see Table 4). It is possible that significant intervention effects found in previous studies without control or comparison groups could be attributed to maturation rather than a specific intervention; i.e., self-efficacy, prosocial skills, perceived social support, and body satisfaction may tend to increase as a student progresses through the school year regardless of whether they attend a Girls Circle or The Council group. The null

effects found in the present study generally replicate findings of other studies that used a randomized design and/or control group (Gies et al., 2015; Mankowski et al., 2011).

Because Girls Circle and The Council are widely implemented programs in a variety of settings, it will likely be helpful to diverse stakeholders (e.g., intervention designers, funders, school administrators, students, facilitators, teachers, etc.) to better understand the potential feasibility and impact of this program in school-based settings. Study strengths include a randomized design, successful randomization of eligible participants, baseline equivalence between groups, large sample size, and use of latent growth curve modeling to capture change in survey outcomes over three time points. Additionally, an ITT approach and investment of community partners in this collaborative research effort mean that findings are likely to be high in internal validity, with the potential to have significant impact on program design and implementation in the future.

Girls Circle and The Council are gender-specific programs that are based on the premise that traditional constructs of masculinity and femininity can increase risk behaviors during adolescence that have short- and long-term impact on development of psychosocial assets such as those measured in this study. Although latent growth modeling revealed mostly null effects for participation in Girls Circle or The Council, there was some evidence that participation in The Council improved self-efficacy in boys who were assigned to participate in groups mid-way through the school year. Of note is that mostly null effects for changes in outcome variables do not necessarily mean that intervention effects did not occur for participants in Girls Circle or The Council groups. The ITT approach employed in this study provides a conservative test of intervention

effects, as it is used to evaluate outcomes based on original group assignment regardless of whether participants received intended services or anticipated dosage. Post-hoc analyses of potential dosage effects were also limited by missing data; it is possible that dosage effects did occur, but were not measurable in this particular study.

Additionally, because participants in all conditions were only followed over the course of one academic year, it is possible that interventions had more long-term impacts on youth trajectories, and that significant differences between groups might emerge if analyses had continued beyond the nine-month span during which data were collected. Other possible explanations of null results could be that the outcome measures used were not sensitive enough to capture change across the available timeframe for data collection, or that the interventions had positive impacts on outcomes that were not assessed in this evaluation.

Limitations

One of the major limitations of this study was the large amount of missing data from all sources (youth self-report surveys, facilitator-reports of fidelity and attendance, and school records). Although FIML was used in growth models to reduce bias from missing survey data, findings would be more conclusive if a greater percentage of participants in all conditions had completed self-report measures at the intended time points. Additionally, missingness in school records (available for only 191 of the original 309 participants) and facilitator-report measures (provided for 95 out of 163 participants in intervention and waitlist-control conditions) limit the interpretation of evaluation components measuring dosage and fidelity. Not only do these missing records from schools and facilitators limit statistical power, but they also may bias the process

evaluation results. For instance, it may be that facilitators who ran groups with lower fidelity or less consistent attendance were less likely to submit documentation. Similarly, available school records may overly represent students who were already more engaged in school and thus more likely to score higher on the school engagement outcome variable regardless of condition.

Other significant limitations relate to outcome variables and process evaluation components relying on single-rater measures. For one, reliance on youth self-report measures for all but one of the outcome variables is a methodological limitation. Findings related self-report data could be enhanced in future studies by including multiple raters (e.g., parents, teachers) and methods (e.g., behavioral observations) assessing youth outcomes. Similarly, process evaluation conclusions (which relied solely on facilitator-report measures) would be strengthened by outside-rater observations, participant surveys, and interviews/focus groups with facilitators and participants.

A further limitation of this study is one common in the field of program evaluations; namely, the mechanisms of change behind potential intervention effects remain unresolved. This is often referred to as the "black box problem" in evaluation studies, and refers to how interventions are viewed primarily in terms of effects, often with limited understanding of how these effects are produced (Astbury & Leeuw, 2010). Although Girls Circle and The Council are based in resiliency practices, motivational interviewing, and relational-cultural theory, it is unclear how theoretical elements of the interventions directly map onto activities and associated outcomes in each session. Even if strong and significant intervention effects had been found, it also might be that improvements in student outcomes would be due to the therapeutic process of engaging

in a supportive group with skilled adult facilitators, building positive relationships with same-gender peers, or periodic access to a same-gender educational setting instead of continued coeducational contexts. Future research evaluating these interventions and other gender-specific programs might consider including a control group condition where participants engaged in same-gender groups without specific curriculum components in order to address whether differences between conditions were due to specific program components or due to receiving same-gender services as a part of coeducational school experience.

Conclusion

This study adds to existing literature evaluating gender-specific school-based programs, as well as to the body of evidence on evaluations of Girls Circle and The Council. It is the hope of the research team that diverse stakeholders better understand the potential impact and feasibility of these programs as implemented in coeducational public middle schools. Although missing data is addressed as a limiting factor, it is worth noting that without the data collection and administrative support of our community partners in the juvenile department and schools, the research team would have no surveys at all. We were also only able to apply a randomized design because of schools' willingness to completely change (for the whole academic year) how they normally scheduled students to participate in Girls Circle and The Council. These interventions are much beloved by the facilitators and staff we collaborated with in conducting this study, and it speaks highly of program designers and implementation teams that they were so open and willing to apply a methodologically rigorous design to assessing their work.

Overall, results from this study do not provide conclusive evidence that participation in Girls Circle or The Council significantly impacts self-efficacy, prosocial skills, perceived social support, body satisfaction, or school engagement as measured over the course of an academic year above and beyond effects seen in a no-services control condition. However, participants, facilitators, and school administrative staff generally seem to like these programs, no iatrogenic effects were found, and mean level scores for boys and girls generally increased over time on all measures. As noted in previous sections, it is possible that these interventions have statistically significant effects on youth resiliency factors that were either not measured in this study or not observed due to challenges with missing data. It is recommended that schools considering implementation of Girls Circle and/or The Council as a part of the school day (a) allot a minimum of 60 minutes per session to allow enough time for core components; (b) provide sufficient facilitator training, especially to manage challenging group dynamics; and (c) continue evaluation of program effectiveness for outcome variables that are meaningful within their schools' context.

APPENDIX

BIVARIATE CORRELATIONS BETWEEN OUTCOME MEASURES AT T1, T2, AND T3

	1	2	3	4	5	6	7	8	9	10	11	12
1. Self-Eff (T1)	-											
	.668***											
2. Self-Eff (T2)	115	-										
	.557***	.670***										
3. Self-Eff (T3)	154	126	-									
	.358***	.414***	.256**									
4. Pro-Soc (T1)	244	114	153	-								
	.342***	.434***	.257**	.621***								
5. Pro-Soc (T2)	151	153	126	114	-							
	.565***	.500***	.452***	.512***	.621***							
6. Pro-Soc (T3)	241	127	186	150	127	-						
	.512***	.412***	.359***	.158*	.129	.168*						
7. Soc Supp (T1)	241	115	153	241	115	150	-					
	.512***	.610***	.430***	.170	.342***	.408***	.719***					
8. Soc Supp (T2)	113	151	124	.103	151	125	113	-				
	.401***	.477***	.477***	.187**	.203*	.298***	.675***	.755***				
9. Soc Supp (T3)	153	128	188	239	128	187	152	126	-			
	.471***	.404***	.298***	.207*	.259**	.197*	.433***	.435***	.388***			
10. Body Sat (T1)	115	113	152	114	113	149	238	111	151	-		
	.381***	.424***	.318***	.121	.200*	.187*	.395***	.386**	.397***	.646***		
11. Body Sat (T2)	115	151	126	153	151	127	115	149	128	113	-	
	.404***	.326***	.361***	.121	.132	.193**	.390***	.337***	.421***	.663***	.630***	
12. Body Sat (T3)	154	126	188	153	126	186	153	124	188	152	126	-
	.185*	.242**	.175*	.123	.278**	.143	.237**	.263**	.235**	.179*	.168*	.034
13. Sch Eng (T3)	149	146	153	149	146	153	150	144	154	146	145	153

Note. Self-Eff = Self-Efficacy, Pro-Soc = Prosocial Skills, Soc Supp = Perceived Social Support, Body Sat = Body Satisfaction, Sch Eng = School Engagement. Values for each cell include correlation and *n*. *$p < .05$, *$p < .01$, *** $p < .001$.

REFERENCES CITED

American Psychological Association. (2018). *The road to resilience*. Online resource: http://www.apa.org/helpcenter/road-resilience.aspx

Arbuckle, J. L. (2014). Amos 23.0 User's Guide. Chicago: IBM SPSS.

Benard, B. (2004). *Resiliency: What we have learned.* San Francisco, CA: WestEd.

Berscheid, E., Hatfield [Walster], E., & Bohrnstedt, G. (1972). Body image - A Psychology Today Questionnaire. *Psychology Today, 6,* 57-67. doi: 10.1037/e400462009-006

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 349–381). San Francisco: Jossey-Bass.

Bosworth, K. & Espelage, D. (2005). Ethnic identity-teen conflict survey. In L. L. Dahlberg, S. B. Toal, M. Swahn, & C. B. Behrens (Eds.), *Measuring violence-related attitudes, behaviors, and influences among youths: A compendium of assessment tools, 2nd Edition.* Atlanta, GA: Centers for Disease Control and Prevention. doi: 10.1037/e554402012-001

Cannon, K. B., Hammer, T. R., Reicherzer, S., & Gilliam, B. J. (2012). Relational-cultural theory: A framework for relational competencies and movement in group work with female adolescents. *Journal of Creativity in Mental Health, 7*, 2–16. doi:10.1 080/15401383.2012.660118

Center for Disease Control (2006). *Morbidity and Mortality Weekly Report, 55*(20), 553-556. doi: 10.1037/e514412006-001

Colorado Foundation for Families and Children (2006). *Quantifying school engagement: Research report.* Online resource: http://www.peecworks.org/peec/peec_inst/017962E8-001D0211.0/Finlay%202006%20Quantifying%20School%20Engagement.pdf.

Cook, T. D., & Campbell, D. J. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Geneva, IL: Houghton Mifflin

Crick, N. R. & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development, 66*(3), 710-722. doi: 10.2307/1131945

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society*, *11*(2), 121–136. doi:10.1080/15248371003699969

Cushing, C., Jensen, C., Miller, M., Leffingwell, T., & Nezu, Arthur M. (2014). Meta-Analysis of Motivational Interviewing for Adolescent Health Behavior: Efficacy Beyond Substance Use. *Journal of Consulting and Clinical Psychology, 82*(6), 1212-1218. doi: 10.1037/a0036912

Daly, Dekker, & Hess. (2016). Dummy coding vs effects coding for categorical variables: Clarifications and extensions. *Journal of Choice Modelling, 21*, 36-41. doi: 10.1016/j.jocm.2016.09.005

Dean, Britt, Bell, Stanley, & Collings. (2016). Motivational interviewing to enhance adolescent mental health treatment engagement: A randomized clinical trial. *46*(9), 1961-1969. doi: 10.1017/S0033291716000568

Deutsch, N., Reitz-Krueger, C., Henneberger, A., Futch Ehrlich, V., & Lawrence, E. (2017). "It Gave Me Ways to Solve Problems and Ways to Talk to People": Outcomes from a combined group and one-on-one mentoring program for early adolescent girls. *Journal of Adolescent Research, 32*(3), 291-322. doi: 10.1177/0743558416630813

Donnon, T.(2010). Understanding how resiliency development influences adolescent bullying and victimization. *Canadian Journal of School Psychology, 25*(1), 101-113. doi: 10.1177/0829573509345481

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues and applications*. Mahwah, NJ: Erlbaum.

Durlak, J., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3-4), 327-350. doi: 10.1007/s10464-008-9165-0

Frederick, D., Bohrnstedt, G. W., Hatfield, E., & Berscheid, E. (2014). Factor structure and validity of the Body Parts Satisfaction Scale: Results from the 1972 Psychology Today survey. *Psychological Topics, 23*(2), 223–242.

Gayes, L., & Steele, R.G. (2014). A meta-analysis of motivational interviewing interventions for pediatric health behavior change. *Journal of Consulting and Clinical Psychology, 82*(3): 521-35.

Gies, S. V., Cohen, M. I., Edberg, M., Bobnis, A., Spinney, E., & Berger, E. (2015). *The Girls Circle: An evaluation of a structured support group program for girls.* Final Report. Development Services Group Inc., MD.

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586. doi: 10.1111/j.1469-7610.1997.tb01545.x

Gray, M., Braun, M., Galvez, G., Leach, A., Finney, B., Garcia, T., Boal, A., & Mankowski, E. (2008). *An evaluation of Boys Council.* Final Report. Portland, OR: Portland State University Department of Psychology. doi: 10.1037/e628592012-339

Gresham, F. M. (2014). Measuring and analyzing treatment integrity data in research. In L. M. H. Sanetti & T. R. Kratochwill (Eds.), *Treatment integrity: Conceptual, methodological, and applied considerations for practitioners* (pp. 109–130). Washington, DC: American Psychological Association.

Hollander, M., & Wolfe, D. A. (1999). *Nonparametric statistical methods* (2nd ed.). Hoboken, NJ: Wiley. doi: 10.1002/9781119196037

Holloman, H., & Yates, P. H. (2013). Cloudy with a chance of sarcasm or sunny with high expectations: Using best practice language to strengthen positive behavior intervention and support efforts. *Journal of Positive Behavior Interventions*, *15*(2), 124-127. doi: 0.1177/1098300712459905

Hossfeld, B., Gibraltarik, R., Bowers, M., & Taormina, G. (2008). Boys Council Facilitator Manual: A strengths-based group approach to promote boys' & young men's safe & healthy development. Cotati, CA: Self-published.

Hu, L., & Bentler, P. M. (2009). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1-55.* doi: 10.1080/10705519909540118

IBM Corp. (2013). IBM SPSS Statistics, Version 22.0. Armonk, NY: IBM Corp.

Irvine, A., & Roa, J. (2010). *Title II Evaluation: Gender specific program in Sonoma and Santa Cruz counties 2007-2010.* Santa Cruz, CA: Ceres Policy Research.

Irvine, A. (2005). *Girls Circle:* Su*mmary of outcomes for girls in the juvenile justice system.* Santa Cruz, CA: Ceres Policy Research.

Jordan, J. (2009). *Relational-cultural therapy.* Washington, DC: American Psychological Association.

Kline, R. (2016). *Principles and practice of structural equation modeling* (Fourth ed., Methodology in the social sciences). New York: The Guilford Press.

Lenz, A. S., Speciale, M., & Aguilar, J. V. (2012). Relational-cultural therapy intervention with incarcerated adolescents: A single-case effectiveness design. *Counseling Outcome Research and Evaluation, 3,* 17–29. doi:10.1177/2150137811435233

Levant, R. F., Richmond, K., Sellers, A., Majors, R. G., Inclan, J. E., Rosello, J. M., . . . Lisak, David. (2003). A multicultural investigation of masculinity ideology and alexithymia. *Psychology of Men & Masculinity, 4*(2), 91-99. doi: 10.1037/1524-9220.4.2.91

Malecki, C.K., & Demaray, M.K. (2002). Measuring perceived social support: Development of the Child and Adolescent Social Support Scale (CASSS). *Psychology in the Schools, 39*(1), 1-18.

Mankowski, E., Gray, M., Viola, W., Vance, K., Lenseigne, L., Wheeler, J., Miles, R., & Jones, C. (2011). *The Council for Boys and Young Men: An assessment of effectiveness in the Ohio Department of Youth Services Final Phase II Report*. Portland, OR: Portland State University Department of Psychology.

Marcotte, D., Fortin, L., Potvin, P., & Papillon, M. (2002). Gender differences in depressive symptoms during adolescence. *Journal of Emotional and Behavioral Disorders, 10*(1), 29-42. doi: 10.1177/106342660201000104

McCabe, M.P. & Ricciardelli, L.A. (2003). Sociocultural influences on body image and body changes among adolescent boys and girls. *The Journal of Social Psychology, 143*, 5-26. doi: 10.1080/00224540309598428

Miller, J. B. (1986). *Toward a new psychology of women* (2nd ed.). Boston, MA: Beacon Press.

Miller, J. B. (1991). The development of women's sense of self. In J. Jordan et al. *Women's Growth in Connection: Writings from the Stone Center.* New York, NY: Guilford Press.

Miller, W. R., & Rollnick, S. (1991). *Motivational Interviewing: Preparing People to Change Addictive Behavior.* New York: Guilford Press. doi: 10.1097/01445442-200305000-00013

Millstein, K. (1993). Limitsetting, coping and adaptation: A theoretical context for clinicians and caregivers. *Child and Adolescent Social Work Journal, 10*(4), 289-300. doi: 10.1007/bf00758260

Morray, E. B., & Liang, B. (2005). Peace talk: A relational approach to group negotiation among Arab and Israeli youths. *International Journal of Group Psychotherapy, 55,* 481–506. doi:10.1521/ ijgp.2005.55.4.481

One Circle Foundation. (2012). Retrieved April 2017, from One Circle Foundation: https://onecirclefoundation.org.

Orpinas, P. (1993). Skills training and social influences for violence prevention in middle schools: A curriculum evaluation. Unpublished doctoral dissertation. Houston, TX: University of Texas Health Science Center. (modified by Bosworth & Espelage, 1993).

Park, R., Goodyer, I., & Teasdale, J. (2005). Self-devaluative dysphoric experience and the prediction of persistent first-episode major depressive disorder in adolescents. *Psychological Medicine, 35*(4), 539-548. doi: 10.1017/s003329170400371x

Patton, M. Q. (2012). Essentials of utilization-focused evaluation. Thousand Oaks, CA: Sage.

Power, T., Blom-Hoffman, J., Clarke, A., Riley-Tillman, T., Kelleher, C., Manz, P., & Chafouleas, S. M. (2005). Reconceptualizing intervention integrity: A partnership-based framework for linking research with practice. *Psychology in the Schools, 42*(5), 495-507.

Roa, J., Irvine, A., & Cervantez, K. (2007). *Girls Circle national research project.* Santa Cruz, CA: Ceres Policy Research.

Rossi, P. H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A Systematic Approach (7th edition)*. Thousand Oaks, CA: Sage.

Sanetti, L., Hagermoser, M., Gritter, K. L., & Dobey, L. M. (2011). Treatment integrity of interventions with children in the school psychology literature from 1995 to 2008. *School Psychology Review,40*(1), 72-84.

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston, *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-NELSON.

Steese, S., Dollette, M., Phillips, W., Hossfeld, E., Matthews, G., & Taormina, G. (2006). Understanding Girls' Circle as an intervention on perceived social support, body image, self-efficacy, locus of control, and self-esteem. *Adolescence*, *41*(161), 55–74.

Tucker, C., Smith-Adcock, S., & Trepal, H. C. (2011). Relational-cultural theory for middle school counselors. *Professional School Counseling, 14,* 310–316. doi:10.5330/PSC.n.2011-14.310

Wothke, W. (2000). Longitudinal and multigroup modeling with missing data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 219-240, 269-281). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment, 52*(1), 30-41. doi: 10.1207/s15327752jpa520