

---

# Making a Difference: Evaluating the Girls in the Game (GIG) After- School Program

Journal of Early Adolescence

33(1) 104–130

© The Author(s) 2013

Reprints and permission:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0272431612466174

jea.sagepub.com



**Amy M. Bohnert<sup>1</sup> and Amanda K. Ward<sup>1</sup>**

## **Abstract**

This randomized controlled trial evaluated the effectiveness of the Girls in the Game (GIG) after-school program (ASP) in promoting social-emotional development and reducing body mass index (BMI) and obesogenic behaviors among a group of minority, low-income girls. Data were collected from 76 third- to fifth-grade girls ( $n = 52$ , GIG;  $n = 24$ , controls) from five urban elementary schools prior to and at the conclusion of 30 weeks of programming. Multiple methods were used to assess outcomes and process variables, including questionnaires, observations, and body measurements (i.e., body mass index [BMI]). The program demonstrated high levels of program quality and was well implemented across program sites. Repeated-measures ANOVAs (analyses of variance) revealed small but significant improvements in body image, nutrition knowledge, and behavior for GIG participants. These findings were not moderated by initial levels of self-esteem or BMI. Analyses suggested that program effectiveness varied depending on process variables.

## **Keywords**

after-school programming, early adolescence, health, females, minority urban youth

---

<sup>1</sup>Department of Psychology, Loyola University, Chicago, IL, USA

## **Corresponding Author:**

Amy M. Bohnert, Psychology Department, Loyola University, 1032 W. Sheridan Road,  
241 Coffey Hall, Chicago, IL 60626, USA.

Email: abohner@luc.edu

## Introduction

Many after-school programs (ASPs) have been implemented based on the premise that participation is beneficial for youth's social-emotional growth. Through provision of adult-supervised activities, youth have the opportunities to foster connections with adults and engage in meaningful activities with peers facilitating enhanced development (Mahoney, Vandell, Simpkins, & Zarrett, 2009; National Research Council and Institute of Medicine, 2002). Another related yet but relatively unexamined potential benefit of ASPs is health promotion (e.g., Mahoney, Lord, & Carryl, 2005). Over the past two decades, many after-school-based health promotion programs have been developed and evaluated; however, very few of these interventions examine the social-emotional benefits of participation (see Beech et al., 2003; Robinson et al., 2010, Story et al., 2003, for exceptions). This study seeks to bring together and build on both the ASP and health promotion literature by conducting a randomized controlled trial to consider whether involvement in a community-based ASP (i.e., Girls in the Game [GIG]) has both social-emotional and health benefits for low-income minority girls.

A recent meta-analysis of ASPs that promote social-emotional development among youth 5 to 18 years of age found that participants exhibited improvements in self-esteem and prosocial behaviors following programming, but results varied significantly depending on program-level variables (Durlak, Weissberg, & Pachan, 2010). Another recent review examined self-esteem interventions for children and reported improved self-esteem following intervention. However, findings varied as a function of study design and intervention target (e.g., global vs. specific self-esteem) in this review as well (O'Mara, Marsh, Craven, & Debus, 2006). In addition to examining self-esteem and prosocial behaviors, the current study will expand on the literature on social-emotional outcomes by including a domain of body image. Several integrated social-emotional and health-based programs working specifically with girls have examined aspects of body image and found that program participants chose larger ideal body images (Story et al., 2003) and report less concerns about weight (Robinson et al., 2003) following intervention.

As previously mentioned, ASPs may also have a positive influence on health outcomes in youth, particularly obesity and obesogenic behaviors. One study of low-income, minority school-age children found that those enrolled in ASPs were significantly less likely to be obese 3 years later than those who were not involved (Mahoney et al., 2005). In addition, data from the 2007 National Survey of Children's Health suggest that no or minimal

involvement in structured after-school activities was associated with greater probability of being overweight or obese among children aged 10 to 17 even after adjusting for several key socioeconomic factors (Bethell, Simpson, Stumbo, Carle, & Gombojav, 2010). These findings provide preliminary evidence that ASPs, even ones that do not specifically focus on health, may have an impact on obesity.

The need for programs that can address social-emotional needs as well as obesity has never been more critical. Obesity among children and adolescents in the United States has reached unprecedented levels with low-income minority girls most affected (Ogden, Carroll, Curtin, Lamb, & Flegal, 2010). A related concern for low-income minority girls is the well-documented decrease in physical activity that occurs during early adolescence, particularly for African American girls (Kimm et al., 2002; Treuth et al., 2009). African American girls who were consistently physically active throughout early and later adolescence (ages 9-19 years) had a mean body mass index (BMI)  $3\text{kg/m}^2$  lower than girls who were consistently inactive (Kimm et al., 2005). Latino groups are the fastest-growing population within the United States (U.S. Census Bureau, 2010), and Latino youth are more likely to be overweight than both Black and White children (National Council of La Raza [NCLR], 2006). In addition, Latina girls have been found to spend less time engaging in physical activity and expend fewer calories than White girls in late childhood and early adolescence (Grieser et al., 2006), with a recent review suggesting that physical activity interventions with early adolescent Latina girls should be a health priority (Olvera, Kellam, Menefee, Lee, & Smith, 2010).

Finding ways to keep girls active and healthy has proven to be difficult during early adolescence (Stice, Shaw, & Marti, 2006), particularly for minority youth, suggesting the need for culturally tailored health-based interventions (Beech et al., 2003). Consequently, several health promotion programs have been developed and evaluated to address obesity and obesogenic behaviors among minority girls during early adolescence. Most of these interventions draw on social cognitive theory (SCT), which suggests that cognitive processes (i.e., self-efficacy) are key factors in determining health behaviors (Baranowski, Perry, & Parcel, 1997; Bandura, 1977). GIG implements these principles in its curriculum by focusing on enhancing girls' health literacy, empowering the girls to believe that they can make healthy choices as well as promoting self-control around health and life choices.

Although health promotion interventions are not ASPs in the traditional sense (with the exception of the Girl Scouts), evidence of their effectiveness offers insight into how an after-school intervention may affect health among

low-income minority girls. One such intervention was evaluated in The Trial of Activity in Adolescent Girls (TAAG) study, a multicenter, group-randomized trial testing the effectiveness of coordinated school- and community-based interventions aimed at reducing the decline in moderate-to-vigorous physical activity (MVPA) among middle-school low-income girls. Girls who participated in the TAAG program did not have decreases in percent body fat but demonstrated small yet significant improvements in MVPA during out-of-school time as assessed via accelerometers (Webber et al., 2008). Another 12-week home- and community-based health promotion program designed to address obesity among African American adolescents, Challenge!, found that overweight/obese status decreased by 5% among intervention participants while it increased by 11% among control adolescents over 2 years (Black et al., 2010). Similarly, the Girls Enrichment Multisite Studies (GEMS) program was designed specifically to address the unique needs of low-income African American girls with a focus on dance. Although initial results did not reveal statistically significant decreases in BMI  $z$  scores (a standardized score created to assess how many standard deviations one is above/below the mean BMI according to CDC growth charts) across the sites, there were positive results at several sites suggesting trends in decreases in BMI  $z$  scores ( $z$ BMI), decreases in TV viewing, and increases in physical activity (Beech et al., 2003; Robinson et al., 2003; Story et al., 2003). An adjunct health-based curriculum has also been developed to increase physical activity and healthy eating in the home among fourth- and fifth-grade girls enrolled in the Girls Scouts (Rosenkranz, Behrens, & Dzewaltowski, 2010). The majority of girls in the intervention were Caucasian (79%), and only 28% were low income. Although the intervention was not associated with improvements in  $z$ BMI, girls who participated in the program had more MVPA. Collectively, these studies of health-based after-school interventions demonstrate some success, particularly in increasing physical activity among girls in middle school residing in urban settings. This investigation of GIG builds on the current literature by examining the effectiveness of an existing community-based ASP that is (a) self-sustaining, (b) focused on promoting social-emotional development as well as health, and (c) tailored for low-income minority girls during early adolescence.

The importance of examining social-emotional outcomes, particularly self-esteem, prosocial behaviors and body image among minority girls, and its relations to weight status, is well documented. Research suggests that overweight girls were more dissatisfied with their bodies than normal-weight girls (Grabe & Hyde, 2006; Stockton et al., 2009) and report significantly lower levels of self-esteem (Nishina, Ammon, Bellmore, & Graham, 2006).

Being overweight in childhood was significantly associated with lower teacher ratings of self-control and interpersonal skills in early adolescent girls (Datar & Sturm, 2006). Finally, many ASPs and health-based interventions are not designed specifically for minority girls residing in low-income urban settings who face unique challenges, such as community safety, limited access to healthy foods, economic disadvantages, and language barriers (Bandy & Moore, 2011a, 2011b). In addition, sociocultural theories of body image suggest that ethnicity is an important factor for understanding how women view their bodies (Abrams & Stormer, 2002). Research has suggested that African American women experiencing fewer body-image concerns than White women (Nishina et al., 2006). In contrast, Latina women experience levels of body-image concerns that are equivalent to White women (Grogan, 2008).

Research on ASPs has increasingly focused on features of these settings that promote positive youth development (see National Research Council and Institute of Medicine, 2002). A recent meta-analysis suggests that ASPs that utilize four practices in facilitating skill development (e.g., SAFE: sequenced, active, focused, and explicit) have the most success (Durlak et al., 2010). A review of 213 school-based social-emotional learning programs revealed that programs that utilized these four practices had positive results across all six study outcomes, whereas those that did not found significant effects in only three areas (i.e., attitudes, conduct problems, academic performance; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). GIG curriculum utilizes these practices in their effort to build participants' skills (i.e., self-efficacy, assertiveness, self-control). Specifically, each lesson follows a structured plan and builds upon previous lessons to achieve their objective (i.e., sequenced). GIG also utilizes engaging and interactive methods to help girls achieve skills, and girls are encouraged to come up with solutions (i.e., active). Finally, GIG program leaders devote a set amount of time each week (e.g., 45 min/session) to teaching these skills (i.e., focused), and girls have a clear understanding (i.e., provided with "topic of the day") about what they are expected to learn (i.e., explicit). Consequently, assessment of process evaluation measures is now incorporated into many evaluations of ASPs; however, this is still not commonly done with health promotion interventions. In the current study, the quality of programming will be assessed, as well as how these various process-related variables are associated with program outcomes.

Though much of the ASP literature seems to be guided by an assumption that participation in an activity setting is equally beneficial for all youths, the association between participation and outcomes may vary across individuals (Durlak, Mahoney, Bohnert, & Parente, 2010). Much of the research conducted on organized activity involvement with adolescents suggests that

high-risk youth benefit more from involvement (i.e., Bohnert & Garber, 2007; Mahoney, 2000; Mahoney & Cairns, 1997). One recent study, however, found that Latino youth who had greater concentration and emotion regulation skills at the outset of programming benefited from consistent participation in an academically oriented ASP (Riggs, Bohnert, Guzman, & Davidson, 2010). To our knowledge, very few health-based interventions have examined potential moderators of treatment effects. Findings from the GEMS study indicated that girls who had unmarried parents/guardians and/or who watched more television at the start of the study experienced less BMI gain over the course of the program (Robinson et al., 2010). In the current study, two potential moderators will be explored: (a) self-esteem, and (b) BMI. Activity participation has been found to be positively associated with self-esteem (Pedersen & Seidman, 2004), and a meta-analysis of self-esteem interventions found at-risk individuals (i.e., girls with lower self-esteem at the start) benefit significantly more from participation (O'Mara et al., 2006). In addition, a recent study reported that African American adolescents who were overweight/obese benefited more from participation in terms of reductions in percentage of body fat and fat mass (Black et al., 2010).

The goal of this randomized controlled trial is to examine the effectiveness of GIG ASPs in promoting social-emotional development and reducing BMI and obesogenic behaviors among a group of urban, low-income, African American and Latina girls. Process-related variables (i.e., attendance, program quality, implementation, and engagement) were assessed to determine whether the association between GIG program participation and outcomes across program sites was related to program-level characteristics. Finally, moderator analyses in the current study focused on exploring whether girls who (a) have lower levels of self-esteem, and (b) are overweight/obese, benefit more from GIG involvement.

## Method

### *Study Design and Procedure*

This randomized controlled trial took place at five public schools that were designated GIG after-school sites. All schools were located in underserved, urban low-income communities. Brief announcements about the study and GIG program were made 2 weeks prior to Time 1 data collection. Consent forms were handed out at these sessions and sent home with all female students accompanied by a cover letter from the principal investigator and an intake form for the GIG Program. A random assignment procedure within

each school was used in which girls with parental consent were assigned to the GIG or control group. Girls were not assigned to the control group if spaces in the program were still available (i.e., filling program slots took priority over balancing sample sizes between GIG and control groups). A random-number table was used to assign participants to a group.

Prior to (Time 1) and at the conclusion (Time 2) of GIG programming, data were collected from GIG and control participants by trained personnel from the research team and GIG. Girls received their group assignment at the conclusion of the Time 1 health festival. Recommendations for state-of-the-art program evaluations that included process evaluation measures were implemented (Beets, Beighle, Erwin, & Huberty, 2009). Data were collected from the girls by trained research assistants of varying ethnicities (e.g., Caucasian, African American, Latina) during health festivals held at each school. The research protocol received Institutional Review Board (IRB) approval from Loyola University and Chicago Public Schools prior to being administered.

### Participants

Participants ( $N = 76$ ) in this study were volunteers in the third to fifth grade, aged 8 to 12 ( $M = 9.13$ ,  $SD = 1.00$ ). As shown in Table 1, the participants were African American (36%), Latina (60%), and Caucasian (4%), and 100% were female. All participants attended schools located in urban communities, with low-income status ranging from 72.3% to 98.1% as indicated by city report. The average BMI for participants was 20.33, and a standardized zBMI score was also computed, with the average being 0.79. For children, the BMI can be plotted on the BMI-for-age growth charts to obtain a percentile ranking, which is used to assess the relative position of the child's BMI number among children of the same sex and age. Girls in this sample had an average BMI percentile ranking of 70.45 ( $SD = 28.70$ ), indicating normal weight status. Still, based on BMI-for-age percentiles 19.1% of girls were overweight, and 27.9% of girls were considered to be obese.

Though 133 participants ( $n = 96$  GIG,  $n = 37$  control) enrolled in the study, only 76 girls ( $n = 52$  GIG,  $n = 24$  controls) were still enrolled and/or attended both data collection sessions and were included in these analyses (see Figure 1). More girls were included in the GIG group because some of the schools did not have enough girls to allow for a control group. Similar attrition rates have been reported for studies examining community-based interventions with similar target audiences (e.g., Douyon, Chavez, Bunte, Horsburgh, & Strunin, 2010). The most common reason for participants dropping out of the program

**Table 1.** Description of Characteristics of Girls at Time 1 by Group.

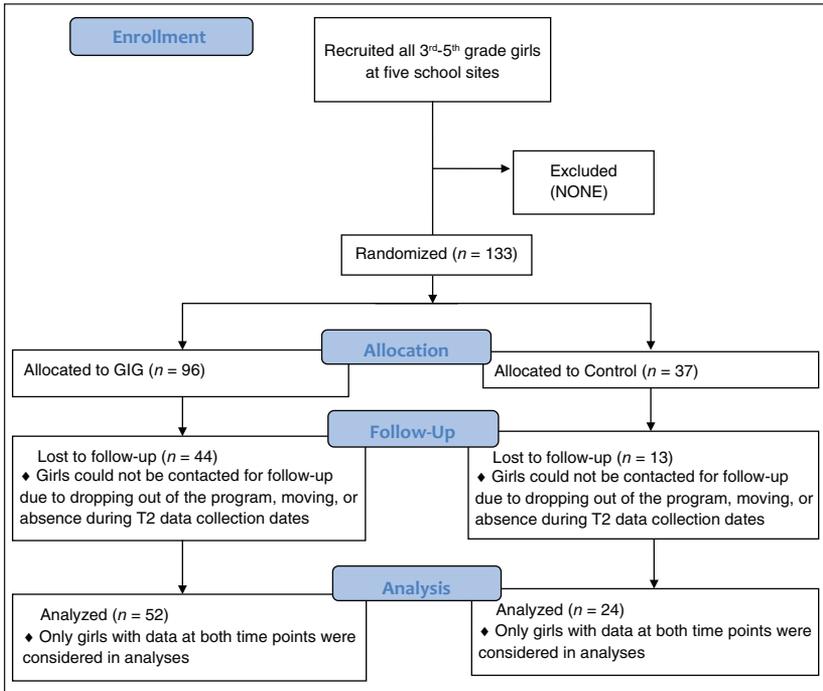
	Overall <i>M</i> ( <i>SD</i> )	GIG <i>M</i> ( <i>SD</i> )	Control <i>M</i> ( <i>SD</i> )
Age	9.13 (1.00)	9.02 (.93)	9.38 (1.13)
Ethnicity (%)			
African American	36.00	35.30	37.50
Latina	60.00	60.80	58.30
Caucasian	4.00	3.90	4.20
Height (in)	54.83 (3.92)	55.04 (3.95)	54.39 (3.91)
Weight	89.77 (31.64)	93.38 (33.08)	81.34 (26.85)
BMI	20.33 (5.03)	20.95 (5.42)	18.93 (3.77)
zBMI	0.79 (1.26)	0.92 (1.35)	0.50 (1.01)
BMI percentile (%)			
BMI = 5th to 84th %	51.5	44.7	66.7
BMI = 85th < 95th	19.1	19.1	19.0
BMI ≥ 95th	27.9	34.0	14.3

Note: BMI = body mass index; GIG = Girls in the Game. *N* = 76: GIG, *n* = 52; control, *n* = 24. No significant differences by group.

was lack of availability due to schedule changes. Other reasons included transferring schools, poor attendance record (i.e., girls were dropped from the program after three consecutive unexcused absences), and loss of interest. Participants who dropped out of the study did not differ from those who participated at both time points in terms of basic demographic characteristics including age ( $M = 9.29$ ,  $SD = 0.98$ ), race (i.e., 38.5% African American, 57.7% Latina, and 1.9% Caucasian), and SES (mean SES = 42.61,  $SD = 13.28$ ; Hollingshead, 1975), or Time 1 outcome variables. In addition, there was no differential attrition across program and control groups, or interactions between program status and Time 1 variables or demographics in predicting attrition status at the conventional ( $p < .05$ ) or trend level ( $p < .15$ ), and the GIG and control groups did not significantly differ based on demographic characteristics and Time 1 outcome variables at either the conventional or trend level of significance.

### Intervention

The GIG ASP is a 30-week curriculum that includes 10 three-week modules. Each module covered a different sport, health, and leadership topic



**Figure 1.** CONSORT diagram.

and was age-appropriate for early adolescents (see Table 2). Each session is led by trained GIG coaches, is approximately 90 min in length, and is divided into two areas of focus: 50% covers physical instruction and energetic activity through traditional and nontraditional sports and fitness activities, and 50% addresses age-appropriate health education, nutrition education, and leadership and life skills topics. (GIG) focuses on enhancing girls' health literacy, empowering the girls to believe that they can make healthy choices as well as promoting self-control around health and life choices. A "girl of the day award" is given to the girl who worked the hardest at each session, along with a small prize. A healthy snack or meal is also provided at every session, along with take-home materials for families to reinforce program messages. No specific interventions were conducted with participants in the control condition other than participating in the health festivals.

**Table 2.** Girls in the Game Program Module Activities.

Sport	Life skill	Health topic
Rhythm and movement	Agree to GIG contract	Self-care and injury prevention
Soccer	Develop a team chant	Body basics and hygiene
Flag football	Practice team chant	Body image and smart eating
Volleyball	Create a positive sign to encourage teammates	Healthy interpersonal relationships
Tennis	Acknowledge others' contributions	First aid and safety
Basketball	Listen to your inner voice	Body Image and Smart Eating II
Lacrosse	Be your own role model	Emotional health
Softball	Girl power	Drug, alcohol, and nicotine awareness
Golf	Energy in/energy out	Body Image and Smart Eating III
Track and field	Keep it going	Review session

Note: GIG = Girls in the Game.

## Measures

*Anthropometry.* Weight was measured with the participant without shoes and dressed in light clothing to the nearest 0.1 kg using a digital scale. Height was measured using a SECA stadiometer without shoes and head held in the Frankfort plane to the nearest 0.1 cm. Eight of the 76 participants did not consent to be weighed and measured, and therefore, could not be included in *z*BMI analyses. However, these participants did not differ from participants with BMI data on any basic demographic variables. These data were then used to calculate BMI according to the following formula:  $BMI = \text{kg}/\text{m}^2$ . The BMI-for-age percentile was determined using the Centers for Disease Control and Prevention (CDC) national norms using age to the nearest month and gender-specific median, *SD*, and power of the Box-Cox transformation. BMI-for-age percentile is used to interpret the BMI number because it is age and sex specific for children and adolescents. These criteria are different from those used to interpret BMI for adults, which does not take into account age or sex as the amount of body fat varies by both age and sex. BMI *z* scores were calculated based on CDC growth charts using The Children's Hospital

of Philadelphia online calculator, and these scores were used in all analyses (Kuczmarski et al., 2002). Researchers have used zBMI scores as a standardized way to assess overweight because there is no ceiling at the upper limit of percentiles where values are collapsed (Daniels, 2009).

**Nutrition.** Participants completed two multiple-choice questions about nutrition knowledge created for this study (“Which of the food groups should you eat the most of each day?” Options: *grains, protein, fruit*; and “If you want to eat more fruit, which would be the best choice?” Options: *orange soda, grape jelly, strawberry ice cream, banana*). Participants were given a score of 1 if their response was correct, and scores were summed across the two items. Participants were also asked five questions about their nutrition behaviors, taken from the CDC Youth Risk Behavior Survey (YRBS) with two questions pertaining to healthy nutritional intake ( $\alpha = .58$ ; for example, “Yesterday, did you eat any vegetables like green salad, carrots, or peas?”) and three questions pertaining to intake of unhealthy foods ( $\alpha = .76$ ; for example, “Yesterday did you eat fries, potato chips, or snack foods?”). Given the modest reliability of the healthy and unhealthy composites, individual item-level responses were also examined.

**Physical activity.** Participants responded to two multiple choice questions about physical activity knowledge created for this study (“What is the shortest time that you can exercise to be strong and healthy?” Options: *20 minutes, 45 minutes, 90 minutes*, and “To be healthy, you should exercise. . .?” Options: *most days of the week, once a week, three times a day*). Participants were given a score of 1 if their response was correct, and scores were summed across the two items. Participants were also asked two questions ( $\alpha = .68$ ) about their engagement in physical activity taken from the CDC YRBS (“During the past week, how many times did you participate in sports or exercise for at least 20 minutes that made you sweat and breathe hard?” and “During the past week, how many times did you do exercises to strengthen your muscles, such as push-ups, sit-ups, or weight lifting?”). Again, given the modest reliability of the composite, individual item-level responses were also examined.

**Social-emotional development.** The Social Skills Scale of the Social Skills Rating System (SSRS; Gresham & Elliott, 1990) is a checklist of 34 prosocial behaviors, and girls were asked to indicate using a 3-point scale, how often (*never, sometimes, very often*) they engaged in each behavior (e.g., “I tell new people my name without being asked to tell it”). The self-report version of the SSRS Social Skills Scale consists of 4 subscales including Cooperation, Assertion, Empathy, and Self-Control. There is a high internal consistency for this measure, with a reported coefficient alpha for all forms ranging from .83 to .94. For this study, the reliability coefficients for all the scales and subscales were high with Cronbach’s alpha ranging from .84 to .90.

Participants also completed the Rosenberg Self-Esteem Scale (RSE, Rosenberg, 1965) as a measure of global self-esteem. Participants were asked 10 questions (e.g., "Sometimes, I think that I am no good at all") with four response choices ranging from *strongly disagree* (0) to *strongly agree* (3). Language was modified to be appropriate for the study population (e.g., "I feel that I have a number of good qualities" to "I feel like there are a bunch of good things about me"). This global self-esteem measure was chosen for its good test-retest reliability ranging from .85 to .88 (Silbert & Tippett, 1965), internal consistency (Gray-Little, Williams, & Hancock, 1997), and strong validity (Rosenberg, 1979), particularly among young minority samples. Cronbach's alpha for the RSE in the current study ranged from .78 to .87.

**Body image.** Participants' perceptions of their current size and ideal shape were assessed using a visual figure rating scale consisting of eight silhouette drawings from *very thin* (0) to *very large* (7) that was adapted for the GEMS study for use with African American girls (Stunkard, Sorenson, & Schulzinger, 1983). Participants were asked two questions, first to circle the image that represented their current figure and then to circle their ideal figure. Higher ideal scores are indicative of endorsing a less emaciated ideal body image, and higher current scores designate endorsement of a larger body type. A measure of body dissatisfaction was also obtained by subtracting the ideal figure from the current figure rating, resulting in a cognitive measure of body dissatisfaction (Tiggemann, 1996). A higher score indicates that a girl prefers a smaller body size than her current shape. This measure has demonstrated high reliability and validity as demonstrated by a review of the psychometric qualities of the instrument (Thompson & Altabe, 1991). It has also been used in studies designed specifically for preadolescent minority girls (Sherwood et al., 2003).

**Process evaluation measures.** Over the course of the year-long program, GIG staff and study personnel collected data on attendance, program quality, curriculum implementation, and participant engagement from program girls only. Attendance was gathered using a sign-in sheet at each session where girls were required to check in with coaches and sign off next to their name on the roster sheet. This information was then entered into a database to track attendance rates at each site.

Program quality was assessed using the Youth Program Quality Assessment (PQA; Smith & Hohmann, 2005). A trained research assistant unobtrusively observed and recorded the program environment and activities at each site at the midpoint of programming. This assistant was trained by a certified expert observer using the PQA training manual. First, the assistant and expert had to obtain 80% agreement on a series of High Scope practice training materials.

Once this level of agreement was achieved on the training modules, both the expert and assistant rated two GIG sites together, with coder reliability being above 80% at these two sites. The trained research assistant went on to rate the three remaining sites independently. The PQA rates programs on four domains, including safe environment (e.g., “Psychological and emotional safety is promoted”), supportive environment (e.g., “Staff provide a welcoming atmosphere”), interaction (e.g., “Children have opportunities to develop a sense of belonging”), and engagement (e.g., “Children have opportunities to set goals and make plans”). Scores of 1, 3, or 5 are given based on how well the program meets the specified criteria.

Curriculum implementation and participant engagement were evaluated through coach reports. Implementation was assessed by asking each coach to rate their interactions with the group (e.g., covered planned activities related to the current sport, life skill, and health topic for the module, and coaches answered participant questions). Participant engagement was assessed by asking coaches to rate how much participants actively engaged in module activities, discussions, asked questions, and enthusiastically participated. Using a 3-point scale (0 to 2), coaches indicated either *no*, *partial*, or *full implementation/engagement*. These variables were assessed at the end of each module, resulting in 10 checks that were completed by the 2 or more coaches at each of these 5 sites. Mean scores were created for both implementation and engagement variables at each school using the averaged coaches’ ratings across all 10 modules. Therefore, each school had an average implementation score as well as an average participant engagement score that reflected the ratings of all coaches at each site.

## Results

### *Statistical Analyses*

To determine the effectiveness of the GIG program participation on outcomes, 2 (Group: GIG, control; between) X 2 (Time: preprogramming Time 1, post-programming Time 2) mixed-model repeated-measures ANOVAs were conducted.<sup>1</sup> Partial eta-squared ( $\eta_p^2$ ) is included as a measure of effect size. This measure provides an estimate of the proportion of the total variance accounted for by the factor under consideration. For example, a  $\eta_p^2$  of .1 would indicate that the factor accounts for 10% of the overall variance. Simple-effects tests were conducted to determine the nature of all significant Group X Time interactions to examine whether GIG participants differently benefited from participation. Given the clustered nature of the data, dummy

variables were created to represent school membership and were used as covariates in all analyses. Interactions between program status and school membership were also tested. No main or interaction effects of school membership were significant. A mean imputation procedure was used in which values were imputed if participants responded to at least 80% of items on a scale. These imputed values were based on participants' responses to the completed items. Table 3 presents the differences between GIG program girls and controls girls on the primary outcome variables at Time 1 and Time 2.

### Process Evaluation

Across all sites, on average, girls who participated in the program throughout the year attended 73.6% of GIG sessions. Ratings of program quality assessed via the PQA were high at all program sites compared to normative data (PQA; Smith & Hohmann, 2005), particularly for safe environment ( $M = 4.78$ ,  $SD = 0.23$ ), supportive environment ( $M = 3.84$ ,  $SD = 0.24$ ), and interaction ( $M = 3.93$ ,  $SD = 0.36$ ) domains, which were all above normative score distributions in validity studies. The fourth domain, engagement, was relatively lower ( $M = 2.64$ ,  $SD = 0.28$ ), but still at the higher end of the distribution for Youth PQA Scales. Implementation data suggest that curriculum was implemented very well across the five school sites ( $M = 1.85$ ,  $SD = 0.12$ ) and participant engagement was high ( $M = 1.81$ ,  $SD = 0.16$ ) on average.

### Social-Emotional and Body-Image Outcomes

There were no significant effects for girls' social-emotional development outcomes as assessed using the SSRS. In addition, levels of self-esteem remained stable over the course of the year for both groups. There was, however, a significant Group X Time effect for ideal body image ( $p = .01$ ). Simple-effects tests (i.e., one-way within-subjects ANOVAs) revealed that GIG girls chose a significantly less emaciated ideal body at Time 2 ( $M = 2.00$ ) than at Time 1 ( $M = 1.59$ ),  $F(1, 48) = 8.94$ ,  $p = .004$ , while control girls remained the same at Time 1 ( $M = 1.83$ ) and Time 2 ( $M = 1.61$ ),  $F(1, 22) = 1.09$ , *ns*. GIG participants also differed from control girls in how dissatisfied they were with their body over the course of the year ( $p = .002$ ). Simple-effects tests revealed that GIG girls were significantly less dissatisfied with their body at Time 2 ( $M = 0.98$ ) than they were at Time 1 ( $M = 1.59$ ),  $F(1, 48) = 8.04$ ,  $p = .007$ , but control girls actually exhibited a trend toward becoming more dissatisfied from Time 1 ( $M = 0.87$ ) to Time 2 ( $M = 1.43$ ),  $F(1, 22) = 3.88$ ,  $p = .06$ .

**Table 3.** Means and Standard Deviations of Outcome Measures.

Construct	Time 1		Time 2		ES
	GIG M (SD)	Control M (SD)	GIG M (SD)	Control M (SD)	
zBMI	0.92 (1.35)	0.50 (1.01)	1.03 (1.07)	0.55 (0.98)	0.00
Nutrition					
Knowledge	0.92 (0.55)	0.96 (0.46)	1.23 (0.55)	0.83 (0.38)	0.09**
Healthy Intake	2.25 (1.30)	2.39 (1.20)	2.06 (1.26)	2.22 (1.35)	0.00
Fruits	1.25 (0.76)	1.46 (0.66)	1.29 (0.80)	1.33 (0.82)	0.01
Veggies	1.00 (0.82)	0.96 (0.83)	.77 (0.78)	0.87 (0.76)	0.00
Unhealthy intake	2.73 (1.82)	2.75 (1.51)	2.58 (1.91)	3.25 (1.48)	0.03
Fries and chips	1.00 (0.79)	0.71 (0.75)	.92 (0.68)	1.08 (0.65)	0.06*
Cookie and candy	0.75 (0.79)	0.96 (0.69)	.75 (0.79)	0.96 (0.81)	0.00
Soda	0.98 (0.80)	1.08 (0.72)	.90 (0.80)	1.21 (0.83)	0.01
Physical activity					
Knowledge	0.94 (0.70)	0.88 (0.68)	1.12 (0.78)	0.92 (0.78)	0.00
Behavior	2.98 (2.18)	3.42 (2.22)	3.79 (2.06)	3.13 (1.96)	0.04
Sports	1.63 (1.24)	1.92 (1.38)	2.23 (1.29)	2.00 (1.25)	0.02
Strengthen	1.35 (1.25)	1.50 (1.22)	1.56 (1.26)	1.13 (0.95)	0.04
Body image/weight perception					
Ideal figure	1.59 (1.21)	1.83 (1.15)	2.00 (1.06)	0.61 (1.20)	0.09*
Current ideal	1.59 (1.57)	.87 (1.63)	.99 (1.60)	1.43 (1.12)	0.13**
Self-report social-emotional development					
Self-esteem	1.85 (0.64)	1.86 (0.45)	1.97 (0.68)	1.94 (0.45)	0.00
SSRS:Total	1.49 (0.28)	1.52 (0.24)	1.47 (0.29)	1.44 (0.18)	0.01
SSRS:Cooperation	1.54 (0.33)	1.59 (0.23)	1.54 (0.38)	1.53 (0.24)	0.01
SSRS:Assertion	1.45 (0.30)	1.45 (0.27)	1.43 (0.32)	1.39 (0.20)	0.01
SSRS:Empathy	1.63 (0.27)	1.67 (0.30)	1.61 (0.28)	1.60 (0.27)	0.00
SSRS:Self-control	1.35 (0.38)	1.36 (0.34)	1.27 (0.38)	1.24 (0.31)	0.00

Note: BMI = body mass index; GIG = Girls in the Game; ES = effect size =  $\eta_p^2$ ; SSRS = Social Skills Rating System. ES is represented by  $\eta_p^2$  for all outcomes. This measure of effect is calculated by  $\eta_p^2 = SS_{\text{effect}} / (SS_{\text{effect}} + SS_{\text{error}})$ , and is the method preferred in SPSS.

\* $p < .05$ . \*\* $p < .01$ .

### Health-Related Outcomes

As shown in Table 3, analyses indicated that there were no significant main intervention effects on girls BMI z scores. In terms of nutrition knowledge, there was a significant Group X Time effect ( $p = .007$ ). Simple-effects tests

revealed that GIG participants' scores at Time 2 ( $M = 1.23$ ) were higher than at Time 1 ( $M = 0.92$ ),  $F(1, 51) = 10.88, p = .002$  suggesting improved knowledge about nutrition for GIG participants. Control participants' scores, however, did not differ from Time 1 ( $M = 0.96$ ) to Time 2 ( $M = 0.83$ ),  $F(1, 23) = 1.30, ns$ . Although there were no significant findings for consumption of healthy foods, a significant Group X Time effect emerged for consumption of one particular unhealthy food category (i.e., fries and chips;  $p = .031$ ). Simple-effects tests revealed that GIG girls reported eating approximately the same amount of these unhealthy foods at Time 2 ( $M = 0.92$ ) as they did at Time 1 ( $M = 1.00$ ),  $F(1, 51) = 0.50, ns$ , but control girls reported eating significantly more fries and chips at Time 2 ( $M = 1.08$ ) than Time 1 ( $M = 0.71$ ),  $F(1, 23) = 3.96, p = .059$ . No significant findings emerged for self-reported physical activity knowledge or behavior.

### *Associations Between Program Attendance, Quality, and Implementation Measures With Youth Outcomes*

Analyses were conducted using process evaluation measures taken at each site. For attendance, participants were split into two categories: low (0-74%,  $n = 22$ ) or high (75%-100%,  $n = 30$ ) attendance due to initial evidence suggesting a dose-response effect of high thresholds of attendance levels (Beets et al., 2009). For the other process evaluation measures (quality, implementation, and engagement), each site was categorized as low or high for each process variable based on the relative standing (i.e., being in the top half or the bottom half of the five sites). With only five scores (i.e., one for each site), this relative standing procedure allowed for an examination of differences between sites with higher versus lower ratings of process measures. Results suggest that some of the treatment effects varied by process variables.

Specifically, the association between GIG participation and nutrition knowledge varied by curriculum implementation ( $p < .001$ ). Participants at sites with relatively low implementation showed no significant differences from Time 1 ( $M = 1.20$ ) to Time 2 ( $M = 1.20$ ),  $F(1, 24) = 0.00, ns$ , while those at sites with higher implementation scores significantly improved over time ( $M = 0.67$  and  $M = 1.26$ , respectively),  $F(1, 26) = 19.69, p < .001$ . Similarly, nutrition knowledge varied by participant engagement ( $p = .008$ ). Sites that rated participants as relatively less engaged did not improve in nutrition knowledge from Time 1 ( $M = 1.15$ ) to Time 2 ( $M = 1.14$ ),  $F(1, 19) = 0.00, ns$ , while sites with more participant engagement did improve ( $M = 0.78$  and  $M = 1.28$ ),  $F(1, 31) = 17.71, p < .001$ . Consumption of unhealthy foods varied by attendance ( $p = .012$ ). Participants with relatively lower attendance rates

showed no significant differences from Time 1 ( $M = 2.82$ ) to Time 2 ( $M = 3.27$ ),  $F(1, 21) = 2.85$ , *ns*, while participants with higher attendance ate significantly less unhealthy foods at Time 2 ( $M = 2.67$ ) compared to Time 1 ( $M = 2.01$ ),  $F(1, 29) = 4.40$ ,  $p = .045$ .

Participants' physical activity knowledge differed as a function of program quality ( $p = .009$ ), with participants at relatively lower-quality sites showing no significant differences at Time 2 ( $M = 0.93$ ) compared to Time 1 ( $M = 1.07$ ),  $F(1, 27) = 0.56$ , *ns*, while participants at higher-quality sites experienced significant increases in their knowledge from Time 1 ( $M = 0.79$ ) to Time 2 ( $M = 1.33$ ),  $F(1, 23) = 11.60$ ,  $p = .002$ . Physical activity knowledge was also different depending on the site's rating of participant engagement ( $p = .05$ ). However, sites with participants who were relatively less engaged actually improved in physical activity knowledge from Time 1 ( $M = 0.90$ ) to Time 2 ( $M = 1.40$ ),  $F(1, 19) = 7.31$ ,  $p = .014$ , and sites with more engaged participants did not differ across the two time points ( $M = 0.97$  and  $M = 0.94$ ),  $F(1, 31) = 0.03$ , *ns*. This was also the case for self-reported physical activity ( $p = .035$ ), with less engaged sites reporting more physical activity at Time 2 ( $M = 4.35$ ) than at Time 1 ( $M = 2.65$ ),  $F(1, 19) = 8.70$ ,  $p = .008$ , and sites with more engaged participants reporting no differences at Time 2 ( $M = 3.44$ ) and Time 1 ( $M = 3.19$ ),  $F(1, 31) = 0.41$ , *ns*.

### Moderation Analyses

Analyses were conducted to determine whether GIG girls who (a) demonstrate lower ratings of self-esteem, or (b) are overweight/obese, may benefit more from program involvement. Self-esteem was analyzed as a continuous variable. Weight categories were created based on CDC BMI percentile recommendations, and two groups (e.g., underweight/normal weight vs. overweight/obese) were defined similar to Black et al. (2010). Results indicated that the program outcomes did not differ as a function of initial levels of global self-esteem or BMI.

### Discussion

This study evaluates the effectiveness of a community-based ASP, that is, GIG, in promoting social-emotional and health outcomes among a sample of low-income urban minority girls. Process evaluation results suggest that the GIG program was well implemented, promoted participant engagement, and is of high quality. Although after-school interventions focused on health have assessed attendance and implementation to demonstrate feasibility, none to

our knowledge have assessed quality using observational measures. In addition, very few prior studies have considered the full “participation equation” by assessing engagement as well as attendance (Weiss, Little, & Bouffard, 2005). The results from this study also demonstrate the complexity of assessing process evaluation variables and the importance of relying on multiple methods to assess these variables. Although implementation and participant engagement were measured through coaches’ (i.e., program leaders) reports, quality was assessed via an outside observer. Notably, program quality variables were negatively correlated with curriculum implementation and participant engagement. It seems that these unexpected findings may be due in part to method variance. Quality observations were completed only once during the study, and the session that was observed may not have been an accurate reflection of typical programming. Furthermore, coaches may be likely to rate curriculum implementation and participant engagement higher given that they are one important indicator of their job performance.

As suggested by Durlak and Dupre (2008), implementation indicates how well a curriculum or program is put into practice. Despite the fact that assessing implementation is essential to all program evaluations (Durlak et al., 2010), there has been little discussion in the literature about the best way to assess it as well as participant engagement. Some studies have relied on observational reports to assess implementation (see Rosenkranz et al., 2010, for an example) while the current study had program leaders complete implementation checks. Future research should focus on developing a “best practices” framework for collecting these process variables.

In their review, Durlak et al. (2010) suggests that studies of ASPs should not only assess attendance but also determine whether it affects treatment outcomes. We believe that it is also important to consider whether other process variables such as implementation, engagement, and quality are associated with program effectiveness. In the current study, three of the process variables (i.e., attendance, implementation, and engagement) were related to participant knowledge as well as behavioral outcomes. In general, better outcomes were achieved when program sites were rated as higher on a particular process variable. For instance, higher implementation and engagement ratings were associated with improvements in nutrition knowledge, and better attendance was associated with less unhealthy eating. Two exceptions though emerged for engagement. In both cases, at sites that were rated lower on engagement, participants demonstrated more knowledge about appropriate levels of physical activity and reported higher levels of physical activity. It is important to note that the high versus low designation for all process variables reflected each program’s standing relative to other sites. All program

sites were well rated in terms of process variables. In future studies, researchers should also pay careful attention to the issue of selection effects given that participants who are more engaged in a program may also differ in their motivations to make changes that could affect the findings.

Findings provide preliminary evidence that the GIG program may be effective in increasing nutrition knowledge and decreasing consumption of unhealthy foods. Although these changes did not translate into reduced BMI, these results suggest girls who attend GIG appear to know more about how to engage in healthier eating practices and report limiting intake of energy-dense foods, which is an essential first step to modifying one of the most important risk factors for obesity (Ello-Martin, Ledikwe, & Rolls, 2005). In addition, contrary to expectations, these results did not differ as a function of initial levels of global self-esteem or BMI. One possible explanation is that these risk factors may not be as relevant for outcomes among girls of this age group. This possibility should be explored in future work.

Similar to prior findings (Robinson et al., 2010, 2003), the results from this study suggest that although this program is designed to address obesity, it does not appear to put girls at higher risk for body-image concerns. Notably, girls in the GIG group endorsed less emaciated ideal body images at the conclusion of the program. In contrast, the control girls reported higher levels of body dissatisfaction at the conclusion of programming. This is particularly important as high levels of body dissatisfaction are a risk factor for disordered eating (Striegel-Moore & Franko, 2002). In addition, earlier onset of body dissatisfaction has been linked to more problems, such as decreased self-esteem in adolescence and young adulthood (Ohring, Graber, & Brooks-Gunn, 2002), highlighting the importance of developing and maintaining positive body image.

Although there were improvements in body image, participating in the GIG program was also not associated with changes in social-emotional development outcomes, including self-esteem and social skills. One recent meta-analysis of ASPs designed to evaluate social-emotional outcomes (e.g., Durlak et al., 2010) found significant improvements across studies, but only for programs that implemented SAFE (e.g., sequenced, active, focused, and explicit) practices. Despite implementing these four practices, there were no significant changes following intervention. There are several potential reasons for these null findings. First, Durlak et al. noted the importance of assessing alternative activity participation of control groups. The alternative programming received by the control group in this study was not carefully assessed, which may have influenced the estimate of the impact of the intervention. The authors also highlight the importance of critically examining

attendance and engagement as a combined variable to gain a more accurate assessment of “full participation.” Although this study made an attempt at assessing “full participation” noted above (i.e., engagement and attendance), engagement was measured at the school level in this study, making it difficult to account for individual participation. Last, the global measure of self-esteem that was utilized (e.g., Rosenberg Self-Esteem Scale) may not have been sensitive enough to detect intervention effects (O’Mara et al., 2006). Future efforts should assess the effects of the GIG intervention in relation to particular self-concept domains that are relevant (i.e., athletic self-concept) rather than just assessing global self-esteem (O’Mara et al., 2006).

As mentioned previously, the GIG program was also not effective in changing BMI. In fact, many obesity prevention trials find no effect on BMI (see Kanekar & Sharma, 2009), particularly with a small sample size and with girls this age (Doak, Heitmann, Summerbell, & Lissner, 2009). In their meta-analysis, Stice et al. (2006) suggest that interventions designed to reduce mean BMI were least effective among 9- to 11-year-old participants, an age at which fat begins to accumulate and continues through age 14 (Kimm et al., 2001). Other researchers have suggested that intervening with girls well before puberty may lead to more success, which is consistent with one recent study that reported larger effect sizes for girls who were less than 9 years of age at enrollment (Klesges et al., 2010). In addition, other studies examining program effectiveness among African American girls have also reported no effect on obesity, and researchers surmise that the difficulty of avoiding weight gain in high-risk populations may be greater because of the surrounding social and environmental contexts that may not support adoption of healthier choices (see Klesges et al., 2010). Indeed, one recent study that was successful in decreasing obesity status among African American adolescents consisted of both a home- and community-based intervention (Black et al., 2010). This issue may be particularly relevant to this intervention where very little attention is directed toward changing the youth’s environments. Thus, although the GIG program provides take home-materials for families to reinforce program messages about health, these findings suggest that interventions may need to place greater emphasis on targeting multiple contexts of influence beyond the individual.

The most notable limitations of this study are the small sample size and attrition. Similar to other community-based programs with similar populations, attrition was high in this study despite extensive efforts to retrieve follow-up data (i.e., at least three alternate data collections dates per school). Still, mean levels of attendance were quite high in our study (73.6%) among program completers. The issue of sample size is also complicated by the partially

nested nature of the data. Although analyses were run to take into account school membership as control variables for all analyses and tested for interactions with program versus control status, this does not address random variance associated with effect sizes across sites. An additional methodological limitation was the involvement of GIG staff in collecting questionnaire data. All these staff members were supervised by the research team, but their involvement does introduce potential bias into some aspects of the data collection. Furthermore, training sessions in standardized administration were required of all personnel who were involved in data collection in which they were instructed simply to read the questionnaire out loud to small groups of girls (3 to 4) and answer any questions that arose.

Variability in ASPs, particularly among the control group as mentioned earlier, was also a limitation. ASP is typically only one component of how youth spend after-school time (e.g., Mahoney, Parente, & Lord, 2007), making it difficult to isolate the specific role of ASPs on various outcomes (Durlak et al., 2010). The majority of studies of ASPs have been vague about what types of arrangements act as a point of comparison (Durlak et al., 2010), and thus it is not clear whether program effects may be due in part to participation in any program or the program of interest. Future work also needs to more carefully consider the full picture of how youth spend their discretionary time to better delineate the role of a particular after-school intervention.

In addition, researchers are encouraged to conduct well-controlled studies that allow for the identification of the components of ASPs that promote physical activity and other outcomes (Pate & O'Neill, 2009). Increasingly, it is important to identify mechanisms of change, not only between group differences on outcomes (Mun, Bates, & Vaschillo, 2010). Comparing programs that systematically vary what is provided as part of the ASP will lead to better understanding of the relative effectiveness of various program components (Durlak et al., 2010). In addition, examining the maintenance of program effects using postintervention assessment is also an important agenda item for future research in order to demonstrate that changes are enduring.

This study utilized a best-practices approach by including multiple methods and process evaluation measures (Beets et al., 2009; Pate & O'Neill, 2009). The inclusion of these practices may have facilitated the moderate effect sizes demonstrated in this study, which is particularly encouraging considering many ASPs produce only small effects when significant findings emerge (Durlak et al., 2010). In addition, the current study assessed multiple domains of adjustment as outcomes and assessed preintervention equivalence between treatment and control groups to rule out selection bias (see Durlak et al., 2010). In the current study, the two groups were equivalent on all key indicators eliminating

the necessity to control for differences. However, when Time 1 levels of variables were included as covariates, the findings remained unchanged.

To summarize, these findings suggest that ASPs have the potential to be one place that early adolescents can establish positive health-related behaviors. Getting girls in early adolescence to engage in health-related programming, however, has been particularly challenging despite their high-risk status for obesity (Stice et al., 2006). Although a more rigorous test of the intervention in a larger controlled trial is needed, these results suggest that GIG warrants further investigation as a promising ASP for urban, minority girls.

### **Acknowledgments**

We would like to thank all GIG staff particularly executive director Amy Skeen, Liz Pearson, Miriam Merrill, Elizabeth Okey, and Kristi Skala, as well as GIG coaches for their tireless efforts on behalf of Chicago girls. We would also especially like to acknowledge and thank all the girls and their families who participated in the project and took the time to share their perspectives with us.

### **Declaration of Conflicting Interests**

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### **Funding**

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a seed grant from the Chicago Consortium to Lower Obesity in Chicago Children (CLOCC: AU 508485). None of the authors have any financial involvement with this organization.

### **Note**

1. To address regression to the mean as a threat to internal validity when using the simple difference approach of repeated measures ANOVAs (analyses of variance), an ANCOVA (analysis of covariance) procedure was also conducted and findings were identical.

### **References**

- Abrams, L. S., & Stormer, C. C. (2002). Sociocultural variations in the body image perception of urban adolescent females. *Journal of Youth and Adolescence, 31*(6), 443-450.
- Bandy, T., & Moore, K. (2011a). *What works for African American children and adolescents: Lessons from experimental evaluations of programs and interventions*

- (Child Trends Fact Sheet). Retrieved from [http://www.childtrends.org/files/Child\\_Trends-2011\\_02\\_01\\_RB\\_WW4AAchildren.pdf](http://www.childtrends.org/files/Child_Trends-2011_02_01_RB_WW4AAchildren.pdf)
- Bandy, T., & Moore, K. (2011b). *What works for Latino/Hispanic children and adolescents: Lessons from experimental evaluations of programs and interventions* (Child Trends Fact Sheet). Retrieved from [http://www.childtrends.org/files/Child\\_Trends-2011\\_02\\_01\\_RB\\_WW4LatinoChildren.pdf](http://www.childtrends.org/files/Child_Trends-2011_02_01_RB_WW4LatinoChildren.pdf)
- Bandura, A. (1977). *Social learning theory*. Upper Saddle River, NJ: Prentice Hall.
- Baranowski, T., Perry, C., & Parcel, G. (1997). *How individuals, environments, and health behavior interact: Social cognitive theory*. In K. Glanz, F. Lewis, & B. Rimer (Eds.), *Health behavior and health education: Theory, research, and practice*. San Francisco, CA: Jossey-Bass.
- Beech, B., Klesges, R., Kumanyika, S., Murray, D., Klesges, L., McClanahan, B., Slawson, D., & Pree-Cary, J. (2003). Child-and-parent-targeted interventions: The Memphis GEMS pilot study. *Ethnicity and Disease, 13*, S1-40.
- Beets, M. W., Beighle, A., Erwin, H. E., & Huberty, J. L. (2009). After-school program impact on physical activity and fitness: A meta-analysis. *American Journal of Preventive Medicine, 36*(6), 527-537.
- Bethell, C., Simpson, L., Stumbo, S., Carle, A. C., & Gombojav, N. (2010). National, state, and local disparities in childhood obesity. *Health Affairs, 29*(3), 347-356.
- Black, M., Hager, E., Le, K., Anliker, J., Arteaga, S., DiClemente, C., & Wang, Y. (2010). Challenge! Health promotion/obesity prevention mentorship model among urban, black adolescents. *Pediatrics, 126*(2), 280-288.
- Bohnert, A. M., & Garber, J. (2007). Prospective relations between organized activity participation and psychopathology during adolescence. *Journal of Abnormal Child Psychology, 35*(6), 1021-1033.
- Daniels, S. (2009). The use of BMI in the clinical setting. *Pediatrics, 124*, 35-41.
- Datar, A., & Sturm, R. (2006). Childhood overweight and elementary school outcomes. *International Journal of Obesity, 30*, 1449-1460.
- Doak, C., Heitmann, B. L., Summerbell, C., & Lissner, L. (2009). Prevention of childhood obesity—What type of evidence should we consider relevant? *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity, 10*(3), 350-356.
- Douyon, M., Chavez, M., Bunte, D., Horsburgh, R., & Strunin, L. (2010). The GirlStars program: Challenges to recruitment and retention in a physical activity and health education program for adolescent girls living in public housing. *Preventing Chronic Diseases: Public Health Research, Practice, and Policy, 7*(2), 1-6.
- Durlak, J. A., & DuPre, E. P. (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.

- Durlak, J. A., Mahoney, J. L., Bohnert, A. M., & Parente, M. E. (2010). Developing and improving after-school programs to enhance youth's personal growth and adjustment: A special issue of *AJCP. American Journal of Community Psychology*, *45*(3-4), 285-293.
- Durlak, J., Weissberg, R., Dymnicki, A., Taylor, R., & Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405-432.
- Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology*, *45*(3-4), 294-309.
- Ello-Martin, J. A., Ledikwe, J. H., & Rolls, B. J. (2005). The influence of food portion size and energy density on energy intake: Implications for weight management. *American Journal of Clinical Nutrition*, *82*, 236-241.
- Grabe, S., & Hyde, J. S. (2006). Ethnicity and body dissatisfaction among women in the United States: A meta-analysis. *Psychological Bulletin*, *132*(4), 622-640.
- Gray-Little, B., Williams, V., & Hancock, T. (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, *23*(5), 443-451.
- Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system manual*. Circle Pines, MN: American Guidance Service.
- Grieser, M., Vu, M., Bedin-Rung, A., Neumark-Stainzer, D., Moody, J., Young, D., & Moe, S. (2006). Physical activity attitudes, preferences, and practices in African American, Hispanic, and Caucasian girls. *Healthy Education and Research*, *33*, 40-51.
- Grogan, S. (2008). *Body image: Understanding body dissatisfaction in men, women and children* (2nd ed.). New York, NY: Routledge.
- Hollingshead, A. B. (1975). *Four factor index of social status*. Yale University. New Haven, CT.: Yale University.
- Kanekar, A., & Sharma, M. (2009). Meta-analysis of school-based childhood obesity interventions in the U.K. and U.S. *International Quarterly of Community Health Education*, *29*(3), 241-256.
- Kimm, S. Y., Barton, B. A., Obarzanek, E., McMahon, R. P., Sabry, Z. I., Waclawiw, M. A., & Daniels, S. R. (2001). Racial divergence in adiposity during adolescence: The NHLBI growth and health study. *Pediatrics*, *107*(3), E34.
- Kimm, S. Y., Glynn, N. W., Kriska, A. M., Barton, B. A., Kronsberg, S. S., Daniels, S. R., & Liu, K. (2002). Decline in physical activity in black girls and white girls during adolescence. *New England Journal of Medicine*, *347*(10), 709-715.
- Kimm, S. Y., Glynn, N. W., Obarzanek, E., Kriska, A. M., Daniels, S. R., Barton, B. A., & Liu, K. (2005). Relation between the changes in physical activity and body-mass index during adolescence: A multicentre longitudinal study. *Lancet*, *366*(9482), 301-307.

- Klesges, R. C., Obarzanek, E., Kumanyika, S., Murray, D. M., Klesges, L. M., Relyea, G. E., & Slawson, D. L. (2010). The Memphis Girls' Health Enrichment Multi-Site Studies (GEMS): An evaluation of the efficacy of a 2-year obesity prevention program in African American girls. *Archives of Pediatrics and Adolescent Medicine, 164*(11), 1007-1014.
- Kuczumski, R. J., Ogden, C. L., Guo, S. S., Grummer-Strawn, L. M., Flegal, K. M., Mei, Z., & Johnson, C. L. (2002). 2000 CDC growth charts for the United States: Methods and development. *Vital and Health Statistics Series 11, Data From the National Health Survey, 246*, 1-190.
- Mahoney, J. L. (2000). School extracurricular activity participation as a moderator in the development of antisocial patterns. *Child Development, 71*(2), 502-516.
- Mahoney, J. L., & Cairns, R. B. (1997). Do extracurricular activities protect against early school dropout? *Developmental Psychology, 33*(2), 241-253.
- Mahoney, J., Lord, H., & Carryl, E. (2005). An ecological analysis of after-school program participation and the development of academic performance and motivational attributes for disadvantaged children. *Child Development, 76*, 811-825.
- Mahoney, J. L., Parente, M. E., & Lord, H. (2007). After-school program engagement: Links to child competence and program quality and content. *The Elementary School Journal, 107*(4), 385-404.
- Mahoney, J. L., Vandell, D. L., Simpkins, S., & Zarrett, N. (2009). Adolescent out-of-school activities. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (3rd ed., pp. 228-269). New York, NY: John Wiley.
- Mun, E., Bates, M., & Vaschillo, E. (2010). Closing the gap between person-oriented theory and methods. *Development and Psychopathology, 22*, 261-271.
- National Council of La Raza. (2006). *Key facts about childhood obesity in the Latino community* (Fact Sheet). Retrieved from [http://www.viveligero.org:8081/documents/file\\_FS\\_childObesity\\_FNL.pdf](http://www.viveligero.org:8081/documents/file_FS_childObesity_FNL.pdf)
- National Research Council and Institute of Medicine. (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.
- Nishina, A., Ammon, N. Y., Bellmore, A. D., & Graham, S. (2006). Body dissatisfaction and physical development among ethnic minority adolescents. *Journal of Youth and Adolescence, 35*(2), 179-191.
- Ogden, C. L., Carroll, M. D., Curtin, L. R., Lamb, M. M., & Flegal, K. M. (2010). Prevalence of high body mass index in US children and adolescents, 2007-2008. *JAMA: The Journal of the American Medical Association, 303*(3), 242-249.
- Ohring, R., Graber, J. A., & Brooks-Gunn, J. (2002). Girls' recurrent and concurrent body dissatisfaction: Correlates and consequences over 8 years. *The International Journal of Eating Disorders, 31*(4), 404-415.

- Olvera, N., Kellam, S., Menefee, K., Lee, J., & Smith, D. (2010). Physical activity in Latino children: Research and its implications. *Journal of Applied Research on Children: Informing Policy for Children at Risk, 1*, 1-26.
- O'Mara, A., Marsh, H., Craven, R., & Debus, R. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41*, 181-206.
- Pate, R. R., & O'Neill, J. R. (2009). After-school interventions to increase physical activity among youth. *British Journal of Sports Medicine, 43*, 14-18.
- Pedersen, S., & Seidman, E. (2004). Team sports achievement and self-esteem development among urban adolescent girls. *Psychology of Women Quarterly, 28*(4), 412-422.
- Riggs, N. R., Bohnert, A., Guzman, M., & Davidson, D. (2010). After school programs for rural and urban Latina youth: For which children do they work best? *American Journal of Community Psychology, 45*, 417-429.
- Robinson, T. N., Killen, J. D., Kraemer, H. C., Wilson, D. M., Matheson, D. M., Haskell, W. L., & Varady, A. (2003). Dance and reducing television viewing to prevent weight gain in African American girls: The Stanford GEMS pilot study. *Ethnicity & Disease, 13*, S65-77.
- Robinson, T. N., Matheson, D. M., Kraemer, H. C., Wilson, D. M., Obarzanek, E., Thompson, N. S., & Killen, J. D. (2010). A randomized controlled trial of culturally tailored dance and reducing screen time to prevent weight gain in low-income African American girls: Stanford GEMS. *Archives of Pediatrics and Adolescent Medicine, 164*(11), 995-1004.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rosenberg, M. (1979). *Conceiving the self*. New York, NY: Basic Book.
- Rosenkranz, R. R., Behrens, T. K., & Dziewaltowski, D. A. (2010). A group-randomized controlled trial for health promotion in girl scouts: Healthier troops in a SNAP (Scouting Nutrition & Activity Program). *BMC Public Health, 10*, 81.
- Sherwood, N., Story, M., Beech, B., Klesges, L., Mellin, A., Neumark-Sztainer, D., & Davis, M. (2003). Body image perceptions and dieting among African American pre-adolescent girls and parents/caregivers. *Ethnicity and Disease, 13*, 200-207.
- Silbert, E., & Tippet, J. (1965). Clinical assessment and measurement validation. *Psychological Reports, 16*, 1017-1071.
- Smith, C., & Hohmann, C. (2005). *Youth program quality assessment validation study: Findings for instrument validation*. Ypsilanti, MI: High/Scope Press.
- Stice, E., Shaw, H., & Marti, C. N. (2006). A meta-analytic review of obesity prevention programs for children and adolescents: The skinny on interventions that work. *Psychological Bulletin, 132*(5), 667-691.

- Stockton, M. B., Lancot, J. Q., McClanahan, B. S., Klesges, L. M., Klesges, R. C., Kumanyika, S., & Sherrill-Mittleman, D. (2009). Self-perception and body image associations with body mass index among 8-10-year-old African American girls. *Journal of Pediatric Psychology, 34*(10), 1144-1154.
- Story, M., Sherwood, N. E., Himes, J. H., Davis, M., Jacobs, D. R., Cartwright, Y., Smyth, M., & Rochon, J. (2003). An after-school obesity prevention program for African American girls: The Minnesota GEMS pilot study. *Ethnicity & Disease, 13*, S54-64.
- Striegel-Moore, R., & Franko, D. (2002). Body image issues among girls and women. In T. F. Cash & T. Pruzinsky (Eds.), *Body image: A handbook of theory, research, and clinical practice* (pp. 183-191). New York, NY: Guilford.
- Stunkard, A. J., Sorenson, T., & Schulsinger, F. (1983). Use of the Danish adoption register for the study of obesity and thinness. In S. Kety (Ed.), *The genetics of neurological and psychiatric disturbances*. (pp. 115-120). New York, NY: Raven.
- Thompson, K., & Altabe, M. (1991). Psychometric qualities of the figure rating scale. *International Journal of Eating Disorders, 10*(5), 615-619.
- Tiggemann, M. (1996). "Thinking" versus "feeling" fat: Correlates of two indices of body image dissatisfaction. *Australian Journal of Psychology, 48*, 21-25.
- Truth, M. S., Baggett, C. D., Pratt, C. A., Going, S. B., Elder, J. P., Charneco, E. Y., & Webber, L. S. (2009). A longitudinal study of sedentary behavior and overweight in adolescent girls. *Obesity, 17*(5), 1003-1008.
- U.S. Census Bureau. (2010). *Population profile of the United States*. Retrieved from <http://www.census.gov/population/www/pop-profile/natproj.html>
- Webber, L. S., Catellier, D. J., Lytle, L. A., Murray, D. M., Pratt, C. A., & Young, D. R. (2008). Promoting physical activity in middle school girls: Trial of activity for adolescent girls. *American Journal of Preventive Medicine, 34*(3), 173-184.
- Weiss, H. B., Little, P. M. D., & Bouffard, S. (2005). More than just being there: Balancing the participation equation. *New Directions for Youth Development, 105*, 15-31.

## Bios

**Amy M. Bohnert**, PhD, is an associate professor of psychology at Loyola University Chicago. Her research examines psychological and health-related outcomes of involvement in various after-school contexts, especially organized extracurricular activities.

**Amanda K. Ward**, MA, is a doctoral graduate student at Loyola University Chicago. Her research focuses on organized activity involvement and well-being, particularly among urban youth.