



Summary Report: Spring Evaluation

Girls on the Run International

by

Allison Riley, Ph.D.
Vice President of Programming & Evaluation
Girls on the Run International

Maureen R. Weiss, Ph.D.
Professor, School of Kinesiology
University of Minnesota

March 2015

Acknowledgements

Girls on the Run International would like to thank the following graduate students and undergraduate interns in the School of Kinesiology at the University of Minnesota for their assistance with data entry, verification, and analysis:

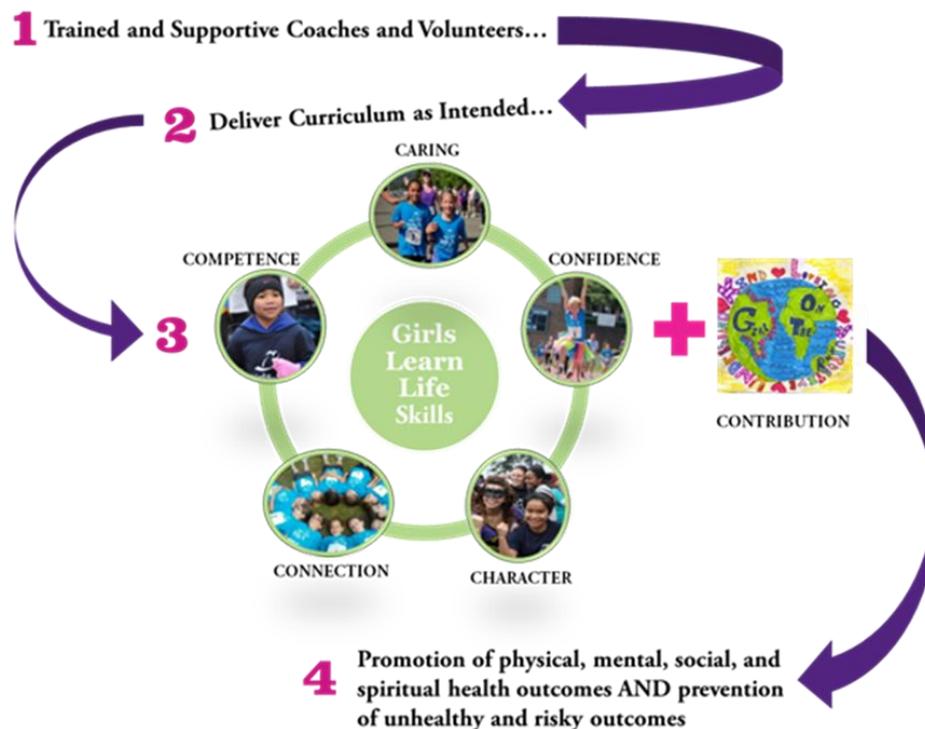
- Alison Phillips, doctoral student, project manager
- Andrea Stark, doctoral student
- Andrew White, doctoral student
- Kelly Bohren, undergraduate intern
- Torey Alberda, undergraduate intern
- Kael Jensen, undergraduate intern
- Rachel Melum, undergraduate intern
- Katie Bisch, undergraduate intern
- Megan Gietzen, master's student
- Saraswati Wilson, undergraduate intern

Overview

As a physical activity-based positive youth development (PA-PYD) program, Girls on the Run employs running and other physical activities as a medium for teaching essential life skills and core values for living a physically active and healthy lifestyle. The program contains all the social-contextual elements necessary of a positive youth development framework including, an intentional curriculum, instructor training to ensure fidelity, a caring climate and supportive relationships. Although varying conceptions of positive youth development exist, Girls on the Run's 12 week curriculum carefully and thoughtfully includes Richard Lerner's 5Cs + 1 – lessons targeting the promotion of competence, confidence, connection, character, caring and contribution to the greater good of society.¹ Specifically, lessons focus on promoting competencies and values related to self-awareness and self-care, selecting healthy social relationships, and empowering girls to have a voice in their personal and community well-being. At the end of each three month session, the girls participate in a Girls on the Run 5k event. This celebratory, non-competitive event is the culminating experience of the curriculum. Completing the 5k provides girls with a tangible understanding of the confidence that comes through accomplishment as well as a framework for setting and achieving life goals. Crossing the finish line is a defining moment when the girls realize that even the seemingly impossible is possible.

Girls on the Run Process and Impact

This diagram provides a visual of the Girls on the Run process, as well as the direct and indirect outcomes of the program.



¹ Lerner, R.M., & Lerner, J.V. (2006). Toward a new vision and vocabulary about adolescence: Theoretical, empirical, and applied bases of a "Positive Youth Development" perspective. In L. Balter & C.S. Tamis-LeMonda (Eds.), *Child psychology: A handbook of contemporary issues* (pp. 445-469). NY: Psychology Press.

First, we have trained coaches who provide positive relationships and create a caring and supportive climate for the girls.

Our coaches also ensure that the curriculum – which is designed to intentionally teach life skills – is implemented as intended.

Over the course of the program the girls will develop life skills – which are known as the 5Cs + 1 (i.e., confidence, competence, caring, connection, and character + contribution). Life skills are defined as competencies learned in one domain or context (e.g., Girls on the Run) that are successfully transferred to other domains and contexts (e.g., dealing with a bully at school; remaining calm after your little sibling did something to annoy you).

Each day of the curriculum focuses on one or more of these life skills. Here we provide a little more information about the 5Cs + 1 and how they relate to the Girls on the Run curriculum.

- At the end of the program we want the girls to feel greater **confidence** in who they are. For instance, the girls gain confidence in their ability to stand up for themselves and others.
- The girls also develop moral **character** as they gain awareness of their ability to intentionally choose actions that demonstrate respect and responsibility towards themselves and others.
- Several lessons teach the girls to respond to others and themselves with **care and compassion**. For example, the girls discuss friendly behaviors that show caring, and also learn strategies to stop gossip.
- The girls have the opportunity to create positive **connections** with peers and adults through lessons targeting cooperative skills and friendship qualities.
- Throughout the program the girls develop and improve physical, social, and emotional **competence** – for instance the girls are training over time for the 5k run.
- When the girls learn these life skills – they can ultimately **contribute** to community and society – which they demonstrate within the Girls on the Run program through the design and implementation of a Community Impact Project.

The 5Cs + 1 then contribute to healthy development through promotion of holistic health outcomes and prevention of unhealthy and risky outcomes. For instance, some indirect outcomes of the program may include promoting a physically active lifestyle and reducing risky behaviors such as obesity, pregnancy, substance abuse, and eating disorders.

Evaluation Spring 2014

In spring 2014 we conducted an evaluation of the Girls on the Run program to determine the impact of program participation on the 5Cs + 1 and physical activity/inactivity. Coaches administered a survey consisting of 24 items reflecting the 5Cs + 1 and physical activity/inactivity prior to the first lesson (pre-assessment) and after the 5k (post-assessment). Information includes sample characteristics, findings with respect to the 5Cs + 1 and physical activity/inactivity, comparison between scores before taking part in Girls on the Run to scores after the program, as well as % of girls who improved in the outcome areas.

Participants

A total of 5124 girls from 28 Girls on the Run councils participated in the study. Girls ranged from 7 to 13 years old, with an average age of 9.84 years. Many (58.7%) reported this was their first season in Girls on the Run, 29.9% had participated in one season, and 11.4% in two or more seasons. Girls were in the 3rd (35.4%), 4th (34.9%), 5th (27.8%), 6th (1.8%), and 7th (0.1%) grades. The majority of girls were White/Caucasian (71.7%), 7.2% Black/African American, 9.3% Hispanic/Latino, 0.6% American Indian/Alaska Native, 4.0% Asian, 5.5% Multi-racial, and 1.6% Other. Although we received surveys from 5124 girls, usable data for the 5Cs + 1 were obtained from 3205-3675 girls due to skipped items, incorrect markings, or missing pre/post data. Of the 332 sites that participated in the study, 270 sent in both pre- and post-surveys. Results are based on the maximum number of girls who completed each measure (between 3205 and 3675 girls).

Measures

Consistent with the positive youth development framework, we assessed constructs reflecting the 5Cs + 1 — confidence, competence, connection, character, caring, and contribution and physical activity/inactivity. We selected valid and developmentally appropriate measures for 8-12 year-old participants of Girls on the Run, which have been used in previous studies examining youth development through sport participation. We also gathered several demographic items, such as age, race/ethnicity, and number of past seasons in Girls on the Run from participant registration forms. We were especially mindful of a reasonable number of items and measures given time constraints by coaches prior to the first lesson.

We assessed confidence, connection, and character using the Self-Perception Profile for Children (Harter, 1985). The average of 5 items comprised scores for confidence, connection, and character (completed by girls). Scores range from 1 (low) to 4 (high) for these subscales. We assessed caring using the Caring Behaviors Scale (Gano-Overway et al., 2010) again using the average of multiple items to compute an overall score. Scores range from 1 (not at all true) to 5 (really true) for this subscale. The scales used to measure confidence ($\alpha = .77, .80$), connection ($\alpha = .69, .73$), character ($\alpha = .74, .78$), and caring ($\alpha = .76, .81$) all showed acceptable reliability. Sample items for each of these scales can be seen in Table 1.

Table 1.

5Cs+1	Sample Items
Confidence	“Some kids do very well at all kinds of physical activities” BUT “Other kids don’t feel they are very good when it comes to physical activities”
Connection	“Some kids get along with others their age” BUT “Other kids do not get along so well with others their age”
Character	“Some kids usually do the right thing” BUT “Other kids often don’t do what they know is right”
Caring	“I treat my classmates with kindness”

We assessed competence by recording whether girls completed the program capstone of the 5k run. Information collected from coaches about each team’s community impact project was used as a measure of contribution. Additionally we used items from the Youth Risk Behavior Surveillance Survey (YRBS) to measure physical activity frequency and sedentary behaviors. More specifically, girls indicated how many days they did physical activities for at least 60 minutes per day during the last full week and the past weekend. They also reported how many hours they watch TV and how many hours they do other sedentary activities such as play video games, computer games or spend time on an iPad or smartphone on a normal school day. Findings are outlined below.

Comparison between pre-season and post-season scores

First we compared girls’ scores before participating in Girls on the Run with their scores after participating in Girls on the Run. We explored this in two ways. First we looked at all the girls in the sample. Then we looked at girls who began the season with lower scores on each of the constructs (below the mean for the entire sample) because we wanted to know whether those girls most in need of a PA-PYD program benefited from participation. We ran t-tests to tell us if the changes from pre-season to post-season were statistically significant. We also calculated the effect size, which is a measure of meaningful or practical significance.

Confidence in Physical Activity

Girls’ pre-season scores were relatively high at 2.92 (2.5 is the midpoint of the scale) and remained high following the program (2.94). This means that participants were responding “Sort of true” about their confidence in doing sports and physical activities.

Next we specifically looked at girls who began the season with lower confidence scores. Results showed a statistically significant increase in confidence from pre-season to post-season. Not only was the increase statistically significant, it was also practically significant. Table 2 provides the pre-season and post-season means and effect sizes. Figure 1 shows the pre- and post-season scores visually.

Connection

Girls answered questions about how well they thought they were liked and accepted by their peers. Scores were high at pre-season (3.20 on a 4-point scale) and remained high at the end of the program (3.20). These scores mean that girls responded between “sort of true” and “really true” for feelings of connection among peers.

Girls who began the season with connection scores lower than the sample mean showed statistically significant increases over the course of the season. The increase also showed practical significance, meaning that their post-season scores were meaningfully different than those at pre-season. Table 2 provides the pre-season and post-season means and effect sizes (practical significance). Figure 1 shows the pre- and post-season scores visually.

Character

Girls responded to items reflecting the extent to which they believe they do the right thing, act the way they’re supposed to, and avoid getting in trouble. Scores were high at pre-season (3.38) and again following the end of the program (3.38). These scores mean that girls responded between “sort of true” and “really true” for perceptions of behavioral conduct.

While scores remained stable for the full sample of girls, statistically and practically significant increases were found for girls who began the season with character scores lower than the sample mean. Table 2 provides the pre-season and post-season means and effect sizes. Figure 1 shows the pre- and post-season scores visually.

Figure 1. Pre-season to post-season comparison for girls with lower pre-season scores.

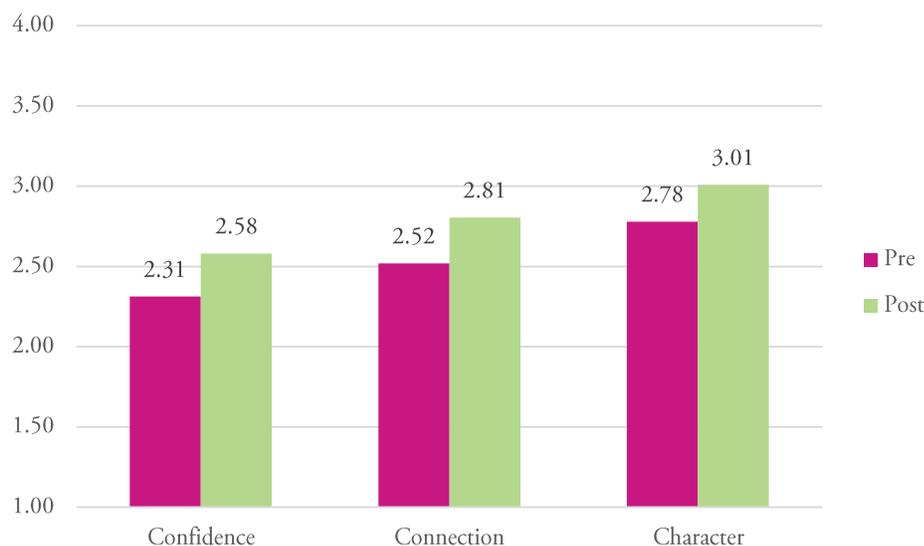


Table 2. Pre-season to post-season comparison for girls with lower pre-season scores.

	n ²	Pre-Season Mean	Post-Season Mean	Effect Size ³
Confidence	1481	2.31	2.58	0.60
Connection	1243	2.52	2.81	0.60
Character	1271	2.78	3.01	0.54
Caring	1156	4.01	4.21	0.42
PA Weekday	1993	2.09	3.29	1.29
PA Weekend	1991	0.79	1.38	1.46
Total PA	1689	2.91	4.58	1.45

²Note: Sample sizes varied due to the number of girls who scored below the sample mean on that variable.

³Note: Effect size indicates practical or meaningful significance beyond statistical significance ($p < .05$). Effect size is calculated as Mean postseason minus Mean preseason divided by the preseason standard deviation. Thus an effect size of 1.0, for example, means that on average girls scored 1 standard deviation higher at post- than at pre-season. Values of .20-.49 indicate a small effect size, .50-.79 a medium effect size, and $> .79$ a large effect size.

Caring

Caring was assessed with items reflecting respect toward others, caring about classmates, and treating others with kindness. Girls rated items very highly at pre-season—4.58 (3.0 is the midpoint of the scale)—and post-season (4.56), meaning that they responded between “pretty true” and “really true” to describe their caring behaviors. Girls who began the season with lower caring scores (below the mean for the entire sample) had statistically and practically significant improvements in caring over the course of the program. Table 2 provides the pre-season and post-season means and effect sizes.

Physical Activity and Sedentary Behavior

Physical activity/inactivity was assessed with items tapping physical activity frequency during the week and on the weekend, as well as TV time and time engaged in other sedentary activities such as on social media or playing video games on a normal school day. Physical activity frequency during the last full school week increased significantly from pre-survey (3.18 days/week) to post-survey (3.63 days/week). Additionally, there was a significant increase in physical activity frequency on the weekend (1.33 days to 1.51 days). We also calculated the total number of days by combining the weekday and weekend scores. At the pre-survey girls indicated that they did physical activity 4.50 days during the last full week, and that number increased significantly to 5.15 days at post-survey.

Statistically significant increases were also found for girls who began the season with scores lower than the sample mean. More specifically, scores increased from pre- to post-season for weekday physical activity, weekend physical activity and total physical activity. The increases were also practically significant, meaning that their post-season scores were meaningfully different than those at pre-season. Table 2 provides the pre-season and post-season means and effect sizes.

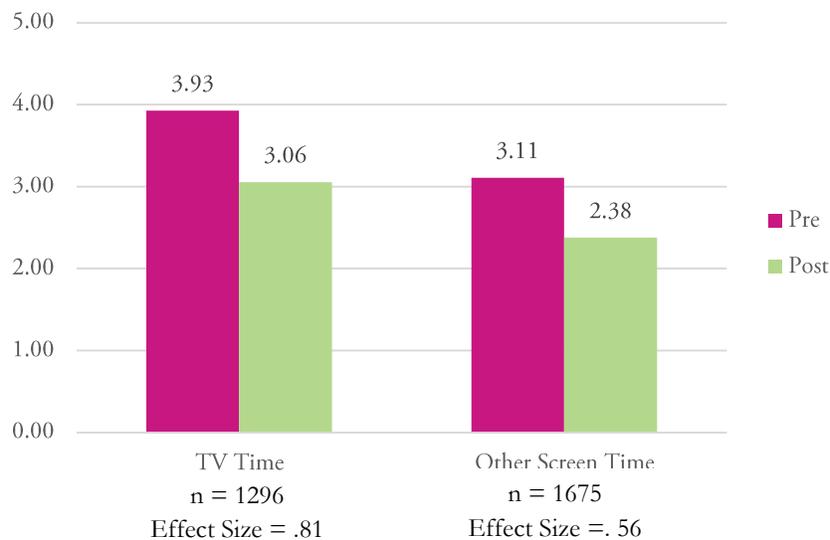
Figure 2. Pre-season to post-season comparison for girls with lower pre-season scores.



Girls reported how many hours they engage in sedentary activities on a normal school day. Time watching TV decreased significantly from pre-season (2.10) to post-season (2.01). Girls also reported the number of hours spent playing video games, computer games or engaging in other related sedentary behaviors. Other screen time decreased significantly from pre-season (1.84) to post-season (1.74). The coding system for these items is included in the note below Figure 3.

Girls who began the season with higher pre-season TV time scores showed significant improvements. Other screen time scores also decreased significantly. Figure 3 shows the pre- and post-season scores and effect sizes.

Figure 3. Pre-season to post-season comparison for girls with higher pre-season scores.



Coding was as follows: 0 = I do not do this activity on a normal school day, 1 = Less than 1 hour per day, 2 = 1 hour per day, 3 = 2 hours per day, 4 = 3 hours per day, 5 = 4 hours per day.

% Improving from Pre-Season to Post-Season

In addition to looking at mean change in scores as shown in the previous section, we also looked at the percentage of girls who improved from pre- to post-season. We calculated these percentages for all girls, new girls (those attending Girls on the Run for the very first time), and returning girls (those who attended a previous season of Girls on the Run). Table 3 shows the percentages for these categories. In the far right column, we also calculated the percentage for girls who began the season with scores lower than the sample mean.

Table 3. Percentage of girls improving from pre-season to post-season.

	Total Sample			Girls with pre-season scores lower than the sample mean
	All Girls (n=3675)	New (n=1951)	Returning (n=1370)	All Girls
Confidence	44.3%	43.7%	44.8%	60.2% (n=1481)
Connection	41.8%	41.8%	41.0%	62.7% (n=1243)
Character	37.5%	36.7%	38.0%	58.3% (n=1271)
Caring	32.5%	33.4%	31.3%	59.7% (n=1156)
PA Weekday	45.2%	44.8%	44.8%	67.5% (n=1993)
PA Weekend	31.0%	31.3%	30.6%	55.9% (n=1991)
Total PA	50.5%	50.8%	49.9%	74.1% (n=1689)
TV Time**	31.6%	34.2%	27.7%	55.4% (n=1296)
Other Screen Time**	32.9%	32.4%	32.9%	56.4% (n=1675)

** Scores decreased from pre- to post-season

Competence

Competence is one of the 5Cs and refers to competence in physical, social, and psychological areas. Because we have measures for social (connection, caring) and psychological (confidence, character) competencies, we focused here on physical competence—specifically the culminating physical activity experience of being able to complete a 5k run at the end of the season. We requested that coaches submit documentation of how many girls completed the 5k at the end of their season. We received this information for 102 out of 332 sites, which documented that 90% (1765 of 1593 total) of girls completed the 5k event.

Contribution

All Girls on the Run participants completed a community impact project during the season. The project is designed and implemented by the girls and is intended to provide girls with important tools that they can use outside of Girls on the Run. To explore how girls contributed to their communities, coaches responded to several questions regarding their team's community impact project. Of the 332 teams involved in the spring evaluation (28 councils), contribution information was collected from 112 coaches.

Many teams completed a project that helped their schools (31%) or animals (18%). Other projects included collecting and donating shoes and raising funds to support the participation of other girls in the Girls on the Run program. Figure 4 shows the results visually. Teams created cards and crafts to support and inspire various groups (30%) such as children in the hospital and military troops overseas. Other projects included collecting shoes and clothes, collecting items for an animal shelter, conducting a food drive or a fundraiser. Figure 5 shows the results visually.

Figure 4.

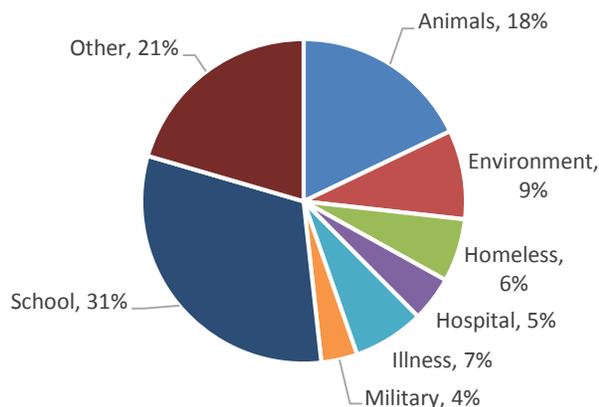
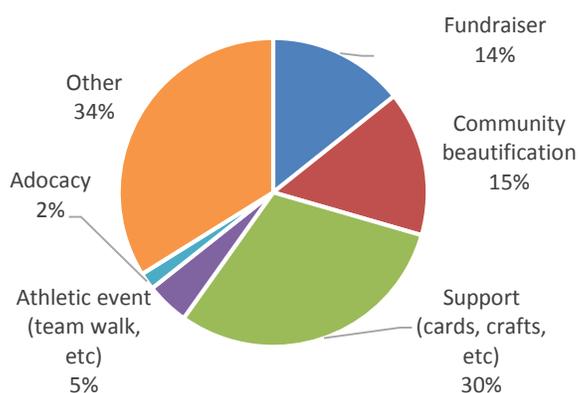


Figure 5.



Coaches were also asked to describe their team's project and how girls took ownership to design and implement the project. A few examples are provided in Table 3.

Table 3.

Example Projects and Girl Ownership
No Littering Campaign using visual arts (posters) and the girls made morning announcements on the loud speaker for the school.
Each Girls on the Run participant was paired with 1 or 2 first grade students who struggle with reading. The girls created a bookmark for the younger student and chose a book to read.
Girls made healthy snack bags for a local homeless organization.
The girls came up with the entire project! We asked them about some issue that was near and dear to their hearts and they came up with the idea.

Coaches also highlighted how girls would extend their contribution beyond the community impact project. One coach said "A few of the girls had so much fun washing windows they were going to go home and ask to wash their house windows!" Finally, coaches highlighted the learning that occurred during the project. One coach put it this way "Having the girls complete a community impact project is essential because it creates a space and time for developing an empathy for others, as well as an avenue to demonstrate gratitude."

Summary of Findings

- Positive change in physical activity and sedentary behavior was observed for the total sample of participants.
 - Girls' physical activity levels increased on weekday, weekend, and total assessments, and TV time and other screen time decreased. Given that Girls on the Run is a physical activity-based youth development program, these findings are encouraging.
- Collectively, findings show that participants in Girls on the Run scored relatively high at pre-program assessment on confidence, connection, character, and caring and their scores remained high at post-program assessment. Stability in scores from pre- to post-program assessment is a positive sign—girls see themselves as healthy in physical, social, and psychological competencies before and after the program.
- Girls who began with lower pre-season scores on confidence, connection, character, caring, and physical activity frequency, and higher scores on sedentary behaviors showed statistically significant improvements from pre-season to post-season. Improvements suggest that girls who need the program the most are likely to improve.
- Many girls showed improvements from pre- to post-season on life skills and physical activity/inactivity. This was especially true for girls who began with lower pre-season scores.

Conclusion

For the overall sample, confidence, connection, character, and caring did not significantly increase from pre- to post-season. Below are some possible reasons for these non-changes, and Girls on the Run International is already addressing these issues as a result of this evaluation.

- Because coaches administered pre- and post-surveys, girls' ratings may have been elevated due to a social desirability effect. We are encouraging Councils to use "arms-length" adults to administer surveys going forward, such as volunteers, board members, and staff.
- Variations in pre- to post-survey scores may be due to variability in program delivery among Councils. Such variability in program delivery became evident during data analysis. Some sites within councils showed increases over time, others showed no change, and still others showed decreases over time. Girls on the Run International has taken steps to standardize processes for coach training to ensure that the intentional curriculum is delivered with integrity.
- The measures were carefully selected and valid and appropriate for 8 to 11 year-olds. However, it could be that alternative or additional measures might better capture impact of curricular content.

An encouraging note, however, is that girls who began with lower pre-season scores on the 5Cs + 1 and physical activity behavior showed statistical and meaningful differences from pre- to post-season. In other words, the girls who are perhaps in most need of a positive program experience to improve psychosocial and physical qualities benefited the most from their participation in Girls on the Run. These results show that evaluation data should be analyzed with consideration of the total sample as well as sub-samples, such as those scoring lower at pre-season, new participants, and returning participants.