Evaluation of The First Tee in Promoting Positive Youth Development: Group Comparisons and Longitudinal Trends

Maureen R. Weiss,1 Nicole D. Bolter,2 and Lindsay E. Kipp3

1University of Minnesota; 2San Francisco State University; 3Texas State University

ABSTRACT

Purpose: This manuscript represents the 3rd in a series of articles documenting our longitudinal evaluation of The First Tee, a physical activity-based youth development program that uses golf as a vehicle for teaching life skills and enhancing developmental outcomes. Previous phases of our project: (a) established initial data-based evidence of effectiveness through cross-sectional and qualitative methods (Weiss, Stuntz, Bhalla, Bolter, & Price, 2013), and (b) provided validity and reliability for a measure of life skills transfer in 3 studies using mixed methods (Weiss, Bolter, & Kipp, 2014). The purpose of the present phase was to: (a) compare youth in The First Tee to youth in other activities on life skills transfer and developmental outcomes, and (b) examine change and stability across 3 years in life skills transfer among youth in The First Tee.

Method: In Study 1, youth participating in The First Tee (N = 405) and a comparison group (N = 159) completed measures of key constructs. In Study 2, a longitudinal sample of 192 youth participating in The First Tee completed the life skills transfer measure for 3 consecutive years. Results: Study 1 revealed that youth in The First Tee compared favorably to youth in other activities on 5 of 8 life skills and 6 of 8 developmental outcomes, and Study 2 showed that scores improved or remained stable for life skills transfer over time. Conclusion: Results from both studies show that The First Tee is effective in teaching for transfer of life skills and promoting developmental outcomes.

ARTICLE HISTORY

Received 26 March 2015
Accepted 18 February 2016

KEYWORDS

Developmental; evaluation research; latent growth modeling; life skills

In October 2014, the Youth-Nex Center for Positive Youth Development at the University of Virginia held a conference for scholars, practitioners, educators, and policymakers to discuss research and share insights about afterschool programs as a context for enhancing healthy developmental outcomes (University of Virginia Curry School of Education, 2015). Afterschool contexts included academic and social programs as well as extracurricular activities such as sports and the arts. This groundbreaking meeting brought together individuals representing multiple academic disciplines and youth organizations to envision how out-of-school-time programs can utilize research to inform best practices that benefit children and adolescents. Inclusion of sport in the conversation is not surprising given the staggering numbers of youth who participate in organized physical activities and the long-held belief that sport is a context for instilling life skills and core values (Weiss, Kipp, & Bolter, 2012).

The positive youth development (PYD) theoretical framework focuses on how youth acquire competencies or assets that lead to positive psychosocial and behavioral outcomes (e.g., Larson, 2000; Lerner & Lerner, 2006). Developing social, psychological, and behavioral assets is maximized when opportunities for skill building occur within a safe and supportive climate guided by caring, competent, and compassionate adults and mentors (Eccles & Gootman, 2002; Roth & Brooks-Gunn, 2003). PYD researchers identify other important social-contextual factors such as appropriate structure (e.g., providing clear rules and expectations) and positive social norms (e.g., reinforcing desirable values and behaviors; see Eccles & Gootman, 2002; Roth & Brooks-Gunn, 2003). Collectively, these features of effective youth development programs focus on intentionally building strengths and assets intended to generalize to multiple life domains.

Given the focus on personal or life skills development, the PYD framework is appealing for conceptualizing youth development through physical activity. Considerable research has demonstrated that sport and physical activity contexts hold tremendous potential for providing the type of climate, coaching behaviors, and opportunities necessary for learning life skills and attaining positive developmental outcomes such as confidence, character, and leadership (see Gould & Carson, 2008; Petitpas, Cornelius, Van Raalte, & Jones, 2005; Weiss, 2008, for in-depth reviews). However, this research
definitively shows that positive outcomes are not an automatic consequence of sport participation. Rather, program components specified by the PYD framework are necessary to optimize acquisition of internal assets and positive developmental outcomes. Thus, youth sport researchers have recommended that the PYD framework be used to provide evidence-based best practices for coaches/teachers, health care providers, and community leaders in physical activity contexts (Weiss et al., 2012; Weiss & Wiese-Bjornstal, 2009).

Petitpas et al. (2005) were among the first scholars to integrate youth sport research with the PYD framework. They differentiated between a youth sport program and youth development program—the former focused on traditional teaching of motor and sport skills to optimize favorable performance outcomes and the latter focused on sport as a context for promoting life skills and core values through delivery of an intentional curriculum of skill-building activities by coaches trained to deliver lessons within a mastery-oriented climate. They identified three essential components of physical activity-based youth development (PA-PYD) programs: (a) a context that is intrinsically motivating and reflected by a psychologically and physically safe climate; (b) external assets that consist of supportive, caring, and competent coaches, parents, mentors, and peers; and (c) internal assets including learned skills that are effective in domains beyond the immediate activity (i.e., life skills).

PA-PYD researchers emphasize the concept of life skills development, which refers to attitudes and behaviors learned in one domain (e.g., sport) that can be generalized or transferred to multiple domains (e.g., Gould & Carson, 2008; Petitpas et al., 2005). The notion of life skills is analogous to personal development goals or assets in the broader PYD literature. Petitpas et al. (2005, p. 70) stated,

> ... effective programs should strive to teach important life skills in a systematic manner and contain clear strategies to foster generalizability of these skills to other domains ... the emphasis on the teaching and acquisition of life skills has been the core element of most youth development programs, including general ... and sport-focused ... initiatives.

Consistent with this philosophy, we believe life skills are not specific to one program or activity, but rather they are strengths and assets that can be used in many domains and allow youth to become contributing members of their community and society.

Following recommendations for adopting the PYD framework to conduct research in physical activity contexts, several descriptive and cross-sectional studies have examined youths’ perceptions of their developmental sport experiences (e.g., Fraser-Thomas & Côté, 2009), associations between the caring climate and social behaviors (Fry & Gano-Overway, 2010; Gano-Overway et al., 2009), and the relationship between perceived coaching behaviors and psychosocial outcomes (e.g., Coatsworth & Conroy, 2009). These descriptive studies provided support for conceiving PYD as a compatible framework for understanding youths’ experiences in physical activity settings. Going beyond these studies to conduct evaluation research of PA-PYD programs is necessary to determine processes and mechanisms by which targeted goals are being achieved (Gould & Carson, 2008; Petitpas et al., 2005; Weiss et al., 2012).

Few evaluation studies have assessed whether PA-PYD programs are successful in attaining the goals of teaching life skills and promoting developmental outcomes (Weiss et al., 2012). Some studies have used qualitative or pretest–posttest designs without a control or comparison group to evaluate the effectiveness of PA-PYD programs in improving self-perceptions, social skills, and respect toward others (Anderson-Butcher, Riley, Amorose, Iachini, & Wade-Mdivanian, 2014; Fuller, Percy, Bruening, & Cotrufo, 2013; Ullrich-French, McDonough, & Smith, 2012). These studies go beyond correlational designs, but without a comparison group, it is possible that positive changes observed from pretest to posttest are attributable to developmental (e.g., cognitive or physical maturity) or environmental factors (e.g., skills learned in family or school contexts). Even fewer studies have conducted follow-up assessments to immediate postprogram data (Gabriel, DeBate, High, & Racine, 2011), and we found only one study that determined whether life skills learned within the program transferred to other domains (Walsh, Ozaeta, & Wright, 2010). Research employing rigorous criteria of appropriate comparison groups and multiple assessments over time would lend more definitive evidence of program effectiveness. Such design characteristics would enable researchers to conclude whether youth participating in PA-PYD programs compare favorably to youth participating in other activities and whether life skills learned are short-lived or enduring. Assessing whether life skills learned within the program transfer to other domains is essential given that it is a signature characteristic of the PYD framework. Systematic research is needed to assess program effectiveness and translate findings to inform best practices (Weiss & Wiese-Bjornstal, 2009).

To address this gap in the literature, Weiss and colleagues engaged in a systematic research program to evaluate The First Tee, a PA-PYD program using golf as a context for teaching life skills and enhancing core values (Weiss, 2008; Weiss, Bolter, & Kipp, 2014; Weiss, Stuntz, Bhalla, Bolter, & Price, 2013). The mission of The First Tee is “to impact the lives of young people by providing
educational programs that build character, instill life-enhancing values, and promote healthy choices through the game of golf” (www.thefirsttee.org; see Weiss et al., 2013, for a description of the philosophy, curriculum, and coach training). In a series of studies, an initial qualitative investigation guided subsequent quantitative phases of research. In Year 1, focus-group and individual interviews were conducted with multiple sources—95 youth participants (ages 11–17 years, M=13.6 years), 26 coaches, and 24 parents—to assess recall of curricular content, teaching methods, life skills learned, and whether life skills learned in golf were transferred to other domains. Findings from all three sources converged to reveal strong evidence of learning and transferring interpersonal and self-management skills to other life domains. Initial evidence of program effectiveness was attributed to central aspects of The First Tee philosophy (e.g., synergy among context, program delivery, and curriculum; seamless approach of teaching golf and life skills concurrently; supporting and trusting relationships).

The next phase included three studies reflecting a comprehensive effort to develop and validate the Life Skills Transfer Survey (LSTS), a measure assessing youth-reported transfer of targeted life skills in the curriculum to other domains. Responses from youth interviews in Year 1 (Weiss et al., 2013) informed the initial pool of items, which were then presented to an expert panel and tested in two pilot studies. A resultant 50-item measure representing eight life skills was administered to 533 youth (ages 10–18 years) in Year 2. Structural validity was shown through a good-fitting model using confirmatory factor analysis, and convergent validity was achieved by finding moderate correlations between LSTS subscales and scales of another instrument that assessed general life skills learning. One year later, in Year 3, 303 youth completed the measure and results showed evidence of longitudinal factorial invariance. Over a series of studies, content, convergent, and construct validity provided evidence that the LSTS is an appropriate self-report measure of youths’ ability to generalize skills learned in a specific program to other domains such as school, home, and peer groups. Validation of the LSTS provided further evidence that The First Tee is having a positive impact in teaching for life skills transfer.

These studies set up the next phase of research—to assess the comparative and long-term impact of The First Tee in promoting life skills transfer and developmental outcomes. According to The First Tee philosophy, PYD is achieved when participants exhibit the Nine Core Values—confidence, integrity, respect, responsibility, honesty, judgment, perseverance, courtesy, and sportsmanship. Youth realize these developmental outcomes when they are engaged in the context of golf, interact with supportive and caring coaches, and learn transferable life skills. The present studies define the next logical step of our research process to evaluate the impact of The First Tee in promoting PYD. Our twofold purpose accounts for limitations in other PA-PYD evaluation research by: (a) employing an appropriate comparison group to determine whether youth in The First Tee favorably contrast to youth in other organized activities on life skills transfer and developmental outcomes, and (b) determining sustained impact by examining stability or change in life skills transfer using a longitudinal design.

**Study 1**

The purpose of this study was to evaluate the effectiveness of The First Tee by comparing youth actively participating in The First Tee with youth involved in other organized activities on life skills transfer and developmental outcomes.

**Participants**

Participants included 405 youth in The First Tee (301 boys, 104 girls) ranging in age from 10 to 17 years (M = 12.6 years, SD = 1.8 years) and a comparison group of 159 youth (93 boys, 66 girls) ranging in age from 10 to 17 years (M = 12.4 years, SD = 1.5 years) who were participating in other activities. The gender percentage for The First Tee sample (74% boys, 26% girls) corresponds to that of the organization at the time of data collection. The large majority of youth in both groups were aged 10 to 14 years old, or early adolescents (84% and 94% for The First Tee and comparison group, respectively). Most of the youth in the comparison group (74.2%) were involved in traditional team and individual sports. Eligible participants were at least 10 years old and had been participating in The First Tee or their specified activity for at least one season. Age criteria ensured that youth possessed the cognitive skills to differentiate information learned in multiple contexts and distinguish sources of social influence for learning life skills (Horn, 2004), while participation criteria ensured that youth had ample experience in their activity to draw on examples of life skills transfer. Participants in The First Tee had been involved in the program for an average of 2 years (M = 2.3 years, SD = 1.4 years), whereas youth in the comparison group had been involved in their specified activity for an average of 3 years (M = 3.4 years, SD = 2.6 years). Both groups were racially and ethnically diverse, with 41.5% of The First Tee youth and 45.3% of comparison-group youth reporting Non-White categories. A higher percentage of youth in the comparison.
group (39.6%) reported qualifying for free/reduced-cost lunch than did youth in The First Tee (17.5%), and a greater percentage of youth in The First Tee reported their parents had completed college or went to graduate school (66.0% vs. 34.9%).

**Measures**

**Life skills transfer**

Youth completed the LSTS (Weiss et al., 2014) to assess perceived transfer of life skills learned in The First Tee or the participant’s specified activity to other domains (e.g., school, home). Fifty items span eight life skills taught in The First Tee curriculum: meeting and greeting, managing emotions, goal setting, resolving conflicts, making healthy choices, appreciating diversity, getting help from others, and helping others. The stem for each item, “Because of participating in The First Tee (my activity) . . . ” is followed by a behavior exemplifying the life skill (e.g., “. . . I stay positive when I am frustrated with my homework”), with responses given on a 5-point scale including really not true for me, not true for me, sort of true for me, true for me, and really true for me. Because youth can learn a variety of skills and behaviors in multiple contexts (e.g., school, home), careful attention was directed to creating instructions specifying that youths’ responses should reflect the degree to which they learned competencies (e.g., managing emotions) because of their participation in The First Tee (most recent program or activity for the comparison group) and not because of what they learned from other sources. Weiss et al. (2014) demonstrated content and construct validity and internal consistency reliability for the LSTS with participants in the same age range as those in the present sample.

**Developmental outcomes**

Based on The First Tee Nine Core Values as indexes of PYD, we used age-appropriate and valid measures to assess youth on confidence, integrity, respect, responsibility, honesty, judgment, and perseverance. Two core values, courtesy and sportsmanship, were not included because they are conceptually similar to other constructs we assessed (e.g., respect, honesty, integrity).

**Confidence.** Confidence was assessed using the Academic and Social Acceptance subscales of Harter’s (1988) Self-Perception Profile for Adolescents (SPPA). These subscales tap youths’ beliefs about doing well in classwork and feeling accepted by peers. Each subscale consists of five items, and responses are given in a structured-alternative format. Youth read two statements (e.g., Some teenagers feel that they are just as smart as others their age BUT Other teenagers aren’t so sure and wonder if they are as smart), select the statement that reflects which teenager they are more like, and then choose how true the statement is for them (i.e., sort of true for me or really true for me). Responses are scored on a 4-point scale, with higher scores indicating greater perceived competence. The SPPA has shown good validity and reliability in studies of youth and adolescents in physical activity contexts (see Horn, 2004).

**Integrity.** Integrity was assessed with the Behavioral Conduct subscale of Harter’s (1988) SPPA. Five items tap youths’ perceptions of doing the right thing, avoiding getting into trouble, and acting the way they are supposed to. Youth responded to items using the structured-alternative format described earlier, with responses scored on a 4-point scale. Higher scores indicate greater perceived behavioral conduct. This subscale has shown good validity and reliability with the same age group as that in the present sample (e.g., Ebbeck & Gibbons, 2003).

**Respect, responsibility, and honesty.** These core values were measured using a rating scale created by the YMCA of the USA (2000) for assessing character development in their youth sports programs. Items show face validity and tap three attributes of character: respect, responsibility, and honesty. Respect was assessed with three items: respect for others, respect for self, and caring for others. Two items tapped responsibility toward others and personal responsibility, and two items assessed honesty about the rules and honesty about effort. Prior to rating items, positive and negative examples were provided to aid youth in responding to each of these items. Youth then rated how true each item was for them because of their participation in The First Tee or their specific activity. Responses are given on a 5-point scale, including really not true for me, not true for me, sort of true for me, true for me, and really true for me. Table 1 displays each attribute, positive and negative examples, and the actual item youth rated.

**Judgment.** This core value was measured by the Self-Regulated Learning subscale of Bandura’s (1990) Multidimensional Scales of Perceived Self-Efficacy. Self-regulated learning refers to making good judgments for finishing tasks on time, focusing on completing a task even when more interesting activities are available, and planning time to be productive (Bandura, 1990; Bandura, Barbaranelli, Caprara, & Pastorelli, 1996). Youth were presented with five items (e.g., “How well can you finish homework assignments by deadlines?”) and were asked to rate their efficacy on a 7-point scale ranging from 1 (not well at all) to 7 (very well). Bandura et al. (1996)
demonstrated good internal consistency reliability for this subscale among 11- to 14-year-old youth.

**Perseverance.** Perseverance was assessed by the Preference for Challenge subscale of the Motivational Orientation in Sport Scale (Weiss, Bredemeier, & Shewchuk, 1985). This subscale assesses the extent to which one chooses optimally challenging tasks, tries hard to accomplish challenging skills, and sustains effort in striving toward goals. Each of the five items is presented in a structured-alternative format (e.g., *Some teenagers like hard skills because they are challenging BUT Other teenagers prefer easy skills because they are sure they can do them*). Participants indicated which statement was most like them and to what degree (i.e., *sort of true for me or really true for me*). Responses were scored on a 4-point scale, with higher scores indicating greater preference for challenge. This scale has shown good validity and reliability with youth in physical activity contexts (e.g., Amorose, 2001).

**Procedure**

After obtaining institutional review board approval, we recruited youth participants from nine chapters in *The First Tee* that met the following inclusion criteria: (a) represented all U.S. geographical regions in *The First Tee* network; (b) offered *The First Tee* programming for 2 or more years, ensuring sustained involvement in teaching the life skills curriculum; (c) included diverse participants by age, gender, and race/ethnicity; (d) were led by directors with extensive coaching experience (i.e., affiliated with the Professional Golfers’ Association of America or Ladies Professional Golf Association); and (e) demonstrated familiarity and confidence with delivering the life skills curriculum as intended (based on information provided by *The First Tee* life skills education staff in the national home office). Because the nine chapters were a long distance from the researchers, we enlisted chapter contacts to help recruit participants in *The First Tee* as well as youth who were participating in other organized activities. Comparison-group participants were recruited from the same nine cities in which *The First Tee* chapters that participated in this study were located. We worked with chapter contacts in *The First Tee* to identify appropriate comparison groups of same-age youth based on their connections with other youth programs in their region, such as YMCA/YWCA, Boys and Girls Clubs, youth soccer/baseball/softball clubs, middle schools, etc. Recruiting included distributing flyers, sending e-mails to parents, and advertising in each chapter’s newsletter. We communicated with chapter contacts through an in-person organizational meeting, telephone follow-ups, and e-mail correspondence to monitor the recruitment process and plan for data collection.

Each participant received a parental consent form in advance of data collection and a reminder phone call to confirm the time and location of the survey. At each data collection site (e.g., golf clubhouses, community centers),
we arrived at least 1 hr in advance to coordinate with staff members and create a youth-friendly environment for administering the survey. Youth submitted their parent consent form upon arrival, gave assent to participate, and were provided with the purpose and procedures of the survey. Parents, coaches, and other individuals not involved in the study were located away from the data collection site. We provided youth with specific instructions for completing the LSTS, with emphasis on responding how true each statement was based on their experiences in The First Tee only (or most recent program or activity for the comparison group) and not because of what they may have learned at home or school or in other settings. Sufficient time was taken to walk participants through two example items and to reiterate the meaning of the phrase, “Because of participating in The First Tee (my activity) . . .” (see Weiss et al., 2014, for detailed information on instructions, sample items, and accentuation of wording on the survey to ensure comprehension). Youth finished the survey in 20–25 min.

Data analysis

We calculated scale reliabilities and correlations for all measures. Two multivariate analyses of covariance (MANCOVAs) were conducted to assess differences between participants in The First Tee group and those in the comparison group on (a) perceived life skills transfer and (b) developmental outcomes. We included parent education and free/reduced-cost lunch as covariates to account for unexpected differences that emerged between groups. If the MANCOVA was statistically significant, follow-up analyses of variance (ANOVAs; $p < .05$) revealed which dependent variables distinguished groups. Effect size was assessed using Cohen’s $d$; Values $\geq 0.20$ denote small effects, $\geq 0.50$ denote medium effects, and $\geq 0.80$ denote large effects (Cohen, 1992).

Results

Scale reliabilities and correlations among variables

Acceptable internal consistency reliability emerged for both groups on all subscales of the LSTS ($\alpha = .78 – .91$) and developmental outcomes ($\alpha = .69 – .84$). Correlations among subscales of the LSTS ranged from $r = .58$ to $r = .79$ for The First Tee and from $r = .30$ to $r = .77$ for the comparison group. Correlations among developmental outcomes ranged from $r = .04$ to $r = .80$ for youth in The First Tee and $r = .00$ to $r = .81$ for the comparison group.

Group differences on life skills transfer

The MANCOVA was statistically significant, Wilks’ $\lambda = .84, F(8, 555) = 15.9, p < .05$. The covariates (free/reduced-cost lunch and parent education) were not statistically significant. A follow-up ANOVA indicated that the two groups differed on five of eight life skills (see Table 2). Youth in The First Tee reported higher scores than the comparison group on transferring the life skills of meeting and greeting, managing emotions, resolving conflicts, appreciating diversity, and getting help from others. Effect sizes were small to medium. Youth in the two groups did not differ statistically on transferring the life skills of setting goals, making healthy choices, and helping others.

<table>
<thead>
<tr>
<th>Life skill/Developmental outcome</th>
<th>The First Tee (n = 405)</th>
<th>Comparison group (n = 159)</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Meeting and greeting*</td>
<td>3.43</td>
<td>0.79</td>
<td>2.97</td>
</tr>
<tr>
<td>Managing emotions*</td>
<td>3.19</td>
<td>0.82</td>
<td>2.94</td>
</tr>
<tr>
<td>Goal setting</td>
<td>3.23</td>
<td>0.92</td>
<td>3.27</td>
</tr>
<tr>
<td>Resolving conflicts*</td>
<td>3.03</td>
<td>0.88</td>
<td>2.77</td>
</tr>
<tr>
<td>Making healthy choices</td>
<td>2.96</td>
<td>0.80</td>
<td>3.15</td>
</tr>
<tr>
<td>Appreciating diversity*</td>
<td>3.32</td>
<td>0.86</td>
<td>3.19</td>
</tr>
<tr>
<td>Getting help*</td>
<td>3.35</td>
<td>0.81</td>
<td>3.17</td>
</tr>
<tr>
<td>Helping others</td>
<td>3.20</td>
<td>0.81</td>
<td>3.14</td>
</tr>
<tr>
<td>Perceived academic competence*</td>
<td>3.14</td>
<td>0.59</td>
<td>2.92</td>
</tr>
<tr>
<td>Perceived social acceptance</td>
<td>3.21</td>
<td>0.64</td>
<td>3.19</td>
</tr>
<tr>
<td>Perceived behavioral conduct*</td>
<td>3.13</td>
<td>0.55</td>
<td>2.87</td>
</tr>
<tr>
<td>Respect</td>
<td>3.92</td>
<td>0.78</td>
<td>3.78</td>
</tr>
<tr>
<td>Responsibility*</td>
<td>3.83</td>
<td>0.79</td>
<td>3.69</td>
</tr>
<tr>
<td>Honesty*</td>
<td>4.02</td>
<td>0.76</td>
<td>3.75</td>
</tr>
<tr>
<td>Preference for challenging skills*</td>
<td>3.09</td>
<td>0.65</td>
<td>2.79</td>
</tr>
<tr>
<td>Self-regulated learning*</td>
<td>5.42</td>
<td>0.98</td>
<td>5.13</td>
</tr>
</tbody>
</table>

Note. Means are adjusted for differences in free/reduced-cost lunch and parent education.

*p < .05.
**Group differences on developmental outcomes**

The MANCOVA was statistically significant, Wilks’ λ = .94, F(8, 544) = 4.9, p < .05. Covariates were not statistically significant. A follow-up ANOVA revealed statistically significant group differences on six of eight developmental outcomes (see Table 2). Youth in *The First Tee* compared favorably to youth in other activities on confidence (perceived academic competence), integrity, responsibility, honesty, judgment, and perseverance. Youth in *The First Tee* scored higher on perceptions of doing well at schoolwork, doing the right thing, working well with others, being honest about rules and effort, regulating behavior to finish tasks on time, and sustaining effort to attain skills. Effect sizes were small to medium. Groups were not statistically different on respect and perceived social acceptance.

**Study 2**

The purpose of Study 2 was to evaluate the effectiveness of *The First Tee* by assessing life skills transfer across 3 years to determine whether youth were increasing or maintaining their use of life skills.

**Participants**

The longitudinal sample included 192 youth (139 boys, 53 girls) who were actively participating in *The First Tee*. At Time 1, youth ranged in age from 10 to 17 years old and 84% were 10 to 14 years old (M = 12.5 years, SD = 1.8 years). At Time 2, youth ranged in age from 10 to 18 years old and 84% were 11 to 15 years old (M = 13.4 years, SD = 1.8 years). At Time 3, youth ranged in age from 11 to 19 years old and 84% were 12 to 16 years old (M = 14.4 years, SD = 1.8 years). Thus, the large majority were early and middle adolescents. Given the range of ages in our sample, we ran a preliminary analysis to explore age as a potential factor affecting life skills transfer. Bivariate correlations for age and life skills transfer dimensions at each time point ranged from \( r = .01 \) to \( r = .17 \), suggesting no relationship. Youth represented a subset of our sample at Time 1 (N = 533; see Weiss et al., 2014) and had participated in our study for 3 consecutive years (Time 1, Time 2, Time 3). At Time 3 (N = 192), the majority of youth were active as students in the program (n = 181) and a small number were active as mentors only (n = 11). Some youth were involved as participants and mentors (n = 54). Mentors were experienced current or former participants in *The First Tee* who assisted head coaches in teaching golf and life skills to younger, less experienced participants. Youth had been involved in *The First Tee* for about 5 years (M = 4.8 years, SD = 1.6 years). Youth participants reported racial and ethnic diversity (46.9% Non-White), that a majority of parents were college educated (66.5%), and that few qualified for free/reduced-cost lunch (15.6%).

**Measures and procedure**

Youth completed the LSTS at three time points (see measures for Study 1). We called or e-mailed each participant 1 year later (Time 2) to invite them for a follow-up survey. A majority of participants (n = 303, 56.8% retention) returned to complete the survey at Time 2. One year later (Time 3), we invited the 303 youth from Time 2 to complete the same survey, and 192 (63.4% retention) agreed to participate. Youth were given a $20 gift certificate to a golf pro shop in return for continuing in the study. At Time 2 and 3, we surveyed youth who were still participating in *The First Tee* as a student and/or mentor. Once youth agreed to participate in the study, we followed the same procedures for data collection at each time point as outlined in Study 1.

**Data analysis**

We calculated scale reliabilities for variables at Time 1, Time 2, and Time 3 and examined correlations between life skills within and across time points. Latent growth modeling (LGM) was used to analyze change over time in life skills transfer (Kline, 2011; Park & Schutz, 2005). LGM is appropriate when analyzing change over time for two reasons. First, measurement error can be explicitly modeled and estimated in LGM, whereas techniques such as repeated-measures ANOVA assume errors to be equal across time points. Second, LGM can also model variation in individual- and group-level change over time instead of considering individual differences as error variance (Kline, 2011).

Figure 1 represents the hypothesized latent growth model for each life skill (Kline, 2011). The three observed variables (\( Y_1, Y_2, Y_3 \)) reflect life skill transfer measured at Time 1, Time 2, and Time 3, respectively. The observed life skill at each time point has an estimated error variance, accounted for by error terms (\( e_1, e_2, e_3 \)). Two latent factors—intercept (\( \eta_I \)) and slope (\( \eta_S \))—reflect participants’ initial level of reported life skill transfer (i.e., intercept) and rate of change over time in life skill transfer (i.e., slope). The intercept has a mean (\( \mu_I \)), representing the average of participants’ true life skill transfer (i.e., observed mean minus measurement error), and a variance (\( \Psi_I \)), reflecting individual differences in participants’ true level of life skill transfer. The slope also has a mean (\( \mu_S \)), signifying participants’ average true
change over time in life skill transfer, and a variance ($\psi_3$) that represents the variability in participants’ rate of change in life skill transfer over time. The covariance between the intercept and slope ($\psi_{IS}$) indicates the relationship between participants’ initial level of life skill transfer and change in life skill transfer over time. The model predicts positive and linear growth in participants’ life skill transfer, meaning that participants would show the same amount of growth from Time 1 to Time 2 and from Time 2 to Time 3 (i.e., one unit of increase per time point). This linear growth is indicated by the fixed pathways at each time point (i.e., 0, 1, 2) from the observed variables to the slope latent factor.

The model in Figure 1 was specified to examine change across three time points for each measured construct (eight models for eight life skills). Several indexes were used to assess the fit of each model to the data: chi-square ($\chi^2$), Non-Normed Fit Index (NNFI), Comparative Fit Index (CFI), and root mean square error of approximation (RMSEA). A nonsignificant $\chi^2$ suggests a good-fitting model but is sensitive to sample size (Park & Schutz, 2005). A model is considered an acceptable fit when the NNFI and CFI are $\geq .90$ and the RMSEA is $\leq .08$, whereas a model is deemed a good fit when the NNFI and CFI are $\geq .95$ and the RMSEA is $\leq .05$ (Hu & Bentler, 1999; Tabachnick & Fidell, 2007). We examined parameter estimates for significance ($t > 1.96$) and were particularly interested in statistical significance of the slope mean, which represents linear change per year.

Results

Scale reliabilities and descriptive analyses for life skills transfer

Table 3 shows reliabilities ($\alpha$), means, and standard deviations for the LSTS subscales at each time point. Alpha values showed good internal consistency reliability for all subscales at each time point ($\alpha = .76 - .92$). Mean scores increased over time for perceived life skills transfer of meeting and greeting, appreciating diversity, getting help from others, and helping others, and scores slightly declined for managing emotions and making healthy choices. Goal-setting scores increased from Time 1 to Time 2 and then decreased at Time 3, while scores for resolving conflicts remained stable from Time 1 to Time 2 and decreased at Time 3. Bivariate correlations showed low to moderately high relationships between life skills within (average $r = .61$) and across time (average $r = .37$).

Latent growth models for life skills transfer

Model fit indexes and parameter estimates are shown for all life skills in Table 4. The nonsignificant chi-square, high values for NNFI and CFI, and low values for

Table 3. Descriptive statistics for life skills transfer at three time points ($N = 192$).

<table>
<thead>
<tr>
<th>Life skill</th>
<th>Time 1 Mean (M)</th>
<th>SD</th>
<th>$\alpha$</th>
<th>Time 2 Mean (M)</th>
<th>SD</th>
<th>$\alpha$</th>
<th>Time 3 Mean (M)</th>
<th>SD</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meeting and greeting</td>
<td>3.57</td>
<td>.78</td>
<td>.84</td>
<td>3.72</td>
<td>.79</td>
<td>.86</td>
<td>3.91</td>
<td>.70</td>
<td>.83</td>
</tr>
<tr>
<td>Managing emotions</td>
<td>3.34</td>
<td>.81</td>
<td>.91</td>
<td>3.32</td>
<td>.81</td>
<td>.92</td>
<td>3.27</td>
<td>.75</td>
<td>.92</td>
</tr>
<tr>
<td>Goal setting</td>
<td>3.41</td>
<td>.88</td>
<td>.89</td>
<td>3.46</td>
<td>.89</td>
<td>.92</td>
<td>3.42</td>
<td>.84</td>
<td>.91</td>
</tr>
<tr>
<td>Resolving conflicts</td>
<td>3.17</td>
<td>.88</td>
<td>.86</td>
<td>3.17</td>
<td>.86</td>
<td>.89</td>
<td>3.14</td>
<td>.83</td>
<td>.89</td>
</tr>
<tr>
<td>Making healthy choices</td>
<td>3.04</td>
<td>.78</td>
<td>.76</td>
<td>3.01</td>
<td>.77</td>
<td>.79</td>
<td>2.91</td>
<td>.83</td>
<td>.82</td>
</tr>
<tr>
<td>Appreciating diversity</td>
<td>3.44</td>
<td>.89</td>
<td>.87</td>
<td>3.51</td>
<td>.86</td>
<td>.88</td>
<td>3.68</td>
<td>.86</td>
<td>.90</td>
</tr>
<tr>
<td>Getting help from others</td>
<td>3.45</td>
<td>.81</td>
<td>.82</td>
<td>3.53</td>
<td>.76</td>
<td>.82</td>
<td>3.58</td>
<td>.72</td>
<td>.83</td>
</tr>
<tr>
<td>Helping others</td>
<td>3.31</td>
<td>.84</td>
<td>.84</td>
<td>3.35</td>
<td>.79</td>
<td>.84</td>
<td>3.36</td>
<td>.72</td>
<td>.81</td>
</tr>
</tbody>
</table>
RMSEA suggested the growth model fit the data well for each life skill. The means for the latent intercepts were statistically significant, indicating participants scored significantly different from 0 in their level of each life skill at each time point. The means of the latent slopes revealed a statistically significant increase over time in youths’ perceived transfer of three life skills—meeting and greeting, appreciating diversity, and getting help from others. These results suggest that participants improved in their perceived ability over time to introduce themselves to new students, relate to kids with different backgrounds, and go to people for help in solving a problem. For the other five life skills (managing emotions, goal setting, resolving conflicts, healthy choices, and helping others), the mean slope was not statistically significant, indicating that youths’ perceptions of transferring these life skills remained stable over time. Youth maintained the abilities to calm themselves down after receiving a bad grade, set goals to get better grades in school, work out a disagreement with a friend, avoid unhealthy behaviors, and comfort a friend when they are upset. Although making healthy choices remained statistically stable over time, this life skill started with the lowest mean at Time 1 (3.04) and showed the greatest decline over time (3.01 and 2.91 at Time 2 and 3, respectively). For all life skills, moderate to moderately high negative correlations emerged between the intercept and slope ($r = -.56$ to $-.64$). Participants with lower scores at Time 1 made greater gains over time in life skills transfer compared with participants with higher scores at Time 1.

**Discussion**

As the third in a series of studies, we determined the effectiveness of The First Tee by (a) comparing participants to youth in other organized activities, and (b) examining change or stability in perceived life skills transfer over time. Our purposes extended past evaluation research of PA-PYD programs (e.g., Anderson-Butcher et al., 2014; Ullrich-French et al., 2012) by including an appropriate comparison group and longitudinal design to provide a more definitive test of program effectiveness. Results for both purposes showed that The First Tee is effective in promoting PYD by teaching for life skills transfer and achieving favorable core values among participants. In the following paragraphs, we systematically summarize and explain the findings, offer theoretical and practical implications, and provide future directions for evaluating whether programs are successful in promoting PYD through physical activity.

First, youth in The First Tee compared favorably to youth in other activities on transferring five of eight life skills to school, home, and other everyday domains. They scored higher on transferring the life skills of meeting and greeting, self-management, conflict resolution, appreciating diversity, and seeking others’ help. Youth in The First Tee also reported higher levels of core values than did youth in the comparison group—confidence in doing schoolwork, doing the right thing, being honest, showing self-responsibility and other responsibility, making good judgments to regulate learning, and sustaining effort over time. Group differences were even stronger considering that youth in other programs had on average 1 more year of experience in their activity than did youth in The First Tee. These results reinforce that a youth development program that employs an intentional curriculum of skill-building activities delivered by trained coaches (and not simply youth sport) can lead to acquired assets and developmental outcomes that generalize to other life domains (Petitpas et al., 2005; Weiss & Wiese-Bjornstal, 2009).

No statistically significant differences emerged for transferring the skills of goal setting, making healthy choices, and helping others or for the core values of respect and perceived social acceptance. Other contexts, such as school and home, also emphasize these skills and values: Teachers and parents, for example, encourage goal setting in preparing for standardized tests. In fact, 84% of our sample in The First Tee reported participating in youth sports where these skills could be learned. Thus, it is conceivable that youth are learning how to set goals, including...
eat a balanced diet, and help friends as well as show respect and feel accepted by peers in other settings and activities. In sum, youth in The First Tee compared favorably to youth in other activities on the large majority of indices reflecting life skills transfer and developmental outcomes.

Second, findings from the longitudinal analysis revealed improving or sustaining life skills transfer across 3 years among youth in The First Tee. Participants reported increasing ability in transferring the skills of meeting and getting to know new people, appreciating diversity such as getting along with others of different backgrounds, and getting help such as finding good role models and trusted individuals for advice. Stable scores emerged over time for the other five life skills. Youth reported maintaining their abilities to set goals to improve grades, show patience when learning a difficult school subject, look for a beneficial solution with friends, choose healthy foods, and give good advice to friends. Importantly, LGM revealed that youth who scored lower on life skills transfer at Time 1 showed greater rates of positive change over time than did those who started with higher scores. This finding indicates that youth who had the most to gain from a PA-PYD program benefited accordingly from participation. Anderson-Butcher et al. (2014) similarly found that youth who began their PYD sports camp with lower perceived social competence scores showed the greatest gains during the course of the program. These findings are also reminiscent of those found by Smoll, Smith, Barnett, and Everett (1993), in that youth baseball players who were lower in self-esteem in the preseason made the greatest gains from preseason to postseason when playing for coaches who were trained to use effective feedback and instructional strategies. Longitudinal findings provide strong support that The First Tee is effective in promoting PYD by having a sustained impact on life skills learning.

Several processes and mechanisms explain the favorable comparison of participants in The First Tee to same-age youth in other organized activities as well as sustained or improving scores on life skills dimensions over time. First, as implied by The First Tee’s mission statement, philosophy, and core values, golf is the context in which assets are taught by coaches trained to deliver an intentional curriculum of life skills. The curriculum and The First Tee Coach Philosophy were created to promote internal assets reflecting personal and interpersonal life skills. Second, the teaching strategies and instructional behaviors endorsed in the four building blocks of The First Tee Coach Philosophy—activity-based, mastery-driven, empower youth, and continuous learning—provide a related explanation. Coaches frequently use memorable teaching cues and effective aspects of observational learning and instructional feedback. Finally, the organization’s emphasis on caring and supportive relationships with coaches is a common thread pervading curricular lessons and interactions on and off the golf course.

According to Patton (2012), evaluation research addresses program effectiveness such as goals achieved, benefits for participants, and to what extent the program was implemented as intended. In addition to empirical questions of program impact, evaluation research is also useful in a practical sense as in, “What, so what, and now what?” (Patton, 2012, p. 3). That is, what experiences does the program offer and what changes in attitudes, skills, and behaviors occur in participants? So what do findings mean in terms of positive and negative outcomes and the degree to which the program is considered a success? Now what recommendations are made based on the findings for improving program impact? We apply Patton’s approach to our three project phases (Weiss et al., 2013, 2014, and present article). First, as described in the previous paragraph, youths’ successful recall and transfer of life skills learning was attributed to the synergy among program components—intentional curriculum, teaching strategies, and supportive relationships. Second, implications of the present studies’ findings can be summarized as follows: (a) favorable comparison of The First Tee to other youth activities in perceived life skills transfer and developmental outcomes, (b) improvement or maintenance of life skills transfer over time, and (c) individuals scoring lower on initial life skills transfer showing a greater rate of positive change over time. These findings are meaningful in that generalizing life skills learned in one context to other salient domains is a signature of PYD programs. Validating a measure of perceived life skills transfer as part of our evaluation research process allowed us—and will allow others—to assess whether this essential outcome has been achieved.

Based on our findings, we made constructive recommendations to The First Tee administrators for improving program impact (i.e., now what). Borrowing from one of the program’s coaching strategies, we shared our findings in the form of, “What was good, what could be better, and how?” Program personnel were appreciative of our efforts and information for “what was good.” They were also receptive to our recommendations for “what could be better and how.” We honed in on results for making healthy choices and how teaching this life skill could be improved. Results for making healthy choices were not significantly different from those in the comparison group (Study 1), and the mean was lowest and declined from Time 1 to Time 3 (Study 2). Given that The First Tee is a PA-PYD program and items for making healthy choices were selected based on the curriculum
(e.g., drink lots of water during the day, choose healthy foods, do sports or exercise every day), these findings were unexpected.

The mean response for transferring healthy behaviors started out at sort of true for me and slipped to between not true for me and sort of true for me. We considered this a negative result and, as such, we suggested how informational content during coach training sessions might include strategies for changing this pattern. Recommendations included providing coaches with the knowledge and skills to: (a) model desirable behaviors to youth participants (e.g., drinking water throughout a lesson); (b) discuss the importance of an active lifestyle with all age groups; (c) encourage golf facilities to eliminate soda options; (d) monitor participants in keeping logs of daily physical activity and eating habits; and (e) create scenarios during lessons to prompt youth to make healthy choices (e.g., peer pressure to engage in risky behaviors). As a result, The First Tee increased its focus on “wellness and physical activity” in the curriculum and coach training. The program also initiated The First Tee Nine Healthy Habits to complement The First Tee Nine Core Values. The nine healthy habits include: energy, play, and safety (physical health); vision, mind, and family (emotional health); and, friends, school, and community (social health). According to The First Tee Web site, “ … The healthy habits are a list of nine health and wellness topics … formatted for easy understanding and learning by elementary-age students and chapter participants.” Thus, in addition to evaluation research serving a theoretical and empirical purpose, it contributes an important practical purpose—in this case The First Tee’s response to placing additional emphasis on physical, social, and emotional health and wellness in its curriculum and coach training.

Along with supportive evidence that The First Tee is an effective PA-PYD program, we note some limitations of our studies and suggest future research directions. First, inclusion criteria resulted in assessing youth from nine chapters of The First Tee that had delivered the curriculum as intended for 2 years or more. It is possible that newer chapters may not have shown similar results. Second, free/reduced-cost lunch and parent education were unexpectedly different for participants in The First Tee and those in the comparison group. The First Tee network includes youth varying in socioeconomic status, but participants in the chapters included in our study happened to be on the higher end. We statistically controlled for these differences, but future studies might strive to assess life skills learning in a primarily low socioeconomic-status population. Third, the longitudinal sample included youth aged 11 to 19 years old, suggesting that youth reflected different developmental periods (e.g., some went from age 11 to 13 years old and others went from age 15 to 17 years old during the 3-year period). It is possible that maturational differences may have affected life skills learning. Future research might explore whether life skills learning in PA-PYD programs is differentially effective for youth who vary in cognitive, social, emotional, or physical developmental level. Finally, because the LSTS is a self-report measure, other modes of assessment such as observations, parent report, and focus groups with various stakeholders might be employed to cross-validate the sources of youths’ life skills learning and whether this learning was effectively transferred to other domains.

In Petitpas et al.’s (2005) framework for planning youth development programs in physical activity contexts, an overlooked component is “ … positive psychosocial growth is most likely to occur when young people … benefit from the findings of a comprehensive system of evaluation and research …” (p. 67, italics ours). In advocating that programs need to be evaluated to determine whether they are accomplishing the goals they purport to foster, Petitpas et al. emphasized that evaluation research should be: (a) multidimensional by assessing internal assets with valid and reliable measures so that comparisons can be made between programs; (b) longitudinal in design to assess outcomes beyond the conclusion of the program; and (c) inclusive of assessing processes and mechanisms explaining change in youth outcomes. We believe our program of research addressed these goals. Along with Petitpas et al., other scholars in youth development (e.g., Arnold & Cater, 2011) and physical activity science (e.g., Armour, Sandford, & Duncombe, 2013; Weiss, 2013) echo the importance of design, methodology, and measurement issues as well as program implementation quality to accurately assess the effectiveness of programs in promoting PYD. The Harvard Family Research Project (Harris, 2011) provides a valuable resource for developing an evaluation strategy for out-of-school-time programs. A toolkit includes a step-by-step approach that includes assessing program’s capacity for evaluation, choosing the focus of evaluation, selecting the evaluation design, deciding on sampling methods and evaluation tools, analyzing data, and using evaluation data.

In conclusion, the present study extends the knowledge base of evaluation research on PA-PYD programs by employing a rigorous design including an appropriate comparison group and multiple assessments over time on life skills transfer. Coupled with our previous two project phases—focus group and individual interviews with youth, parents, and coaches (Weiss et al., 2013) and validation of a measure of life skills transfer (Weiss et al., 2014)—the evidence provides strong support that The First Tee is effective in promoting PYD through its
intentional curriculum of life skills, training of coaches to effectively deliver the curriculum, and supportive and caring relationships with coaches, mentors, and peers. Continued evaluation research is necessary to identify successful PA-PYD programs and the processes they employ to create optimal contexts for physical activity engagement and achieve holistic developmental outcomes (Weiss & Wiese-Bjornstal, 2009).

**What does this article add?**

The PYD framework has flourished in recent years as a compatible way for out-of-school-time physical activity programs to create mission statements, identify core values, and envision processes and mechanisms to make a difference in young people’s lives. Evaluating the impact of PA-PYD programs on intended outcomes must employ rigorous designs such as appropriate comparison groups and multiple assessments over time to determine whether such programs are achieving the goals of teaching life skills and enhancing developmental outcomes. Our two studies significantly contribute to the existing knowledge base on PYD through physical activity by going beyond qualitative, cross-sectional, and descriptive evaluation research to essential design features (comparison group, multiple assessments over time) and statistical analysis (LGM) that provide a more definitive assessment of program impact. These studies also lend further validity to our measure of life skills transfer, the LSTS, which can be used with traditional youth sport as well as youth development programs to assess the degree to which these programs are having a positive impact on life skills transfer. Collectively, our systematic line of research conveyed through three interrelated articles provides a template for conducting evaluation research and translating findings to inform best practices in PA-PYD programs.

**Acknowledgments**

The authors wish to thank Jennifer Bhalla, Ellen Markowitz, Melissa Price, and Cheryl Stuntz for their valuable assistance with data collection. We are grateful to The First Tee life skills education staff for allowing us access to chapters at multiple regional sites that enabled us to carry out a research study of this magnitude. We also thank Beth Brown, director of curriculum and research at The First Tee, for her constant support throughout the longitudinal study, and chapter directors, youth participants, coaches, and parents who graciously facilitated our research efforts through investment of time and resources.

**Funding**

The studies reported in this article were supported by a grant awarded to the first author by Philip Morris USA Youth Smoking Prevention Programs.

**References**


