Formative Evaluation of Project TEAM (Teens Making Environment and Activity Modifications)

Jessica M. Kramer, Kristin Roemer, Kendra Liljenquist, Julia Shin, and Stacy Hart

Abstract

Research documents the negative impact of physical and social environmental barriers on engagement in school, work, and the community for youth with intellectual and/or developmental disabilities (IDD). Project TEAM (Teens making Activity and Environment Modifications) was designed to teach youth to systematically identify environmental barriers, generate modification strategies, and request accommodations. This formative evaluation used a mixed methods expansion design to investigate outcomes, activities, and experiences. Trainees had a significant increase in knowledge of environmental factors and modification strategies but no changes in applied problem-solving. 76% attained at least one goal as measured through goal attainment scaling. Intervention activities ranged in quality. Trainees enjoyed the interactive and applied aspects of Project TEAM but found some concepts and materials difficult to understand. Lessons learned from this comprehensive evaluation can inform future revisions to Project TEAM and may be equally relevant for other researchers evaluating programs targeting transition-age youth with IDD.

Key Words: self-advocacy; environmental barriers; adolescents; formative evaluation

There is a growing recognition that youth and young adults with intellectual and/or developmental disabilities (IDD) transitioning to adulthood can benefit from programs that enable them to be effective self-advocates (King, Baldwin, Currie, & Evans, 2005; Merchant & Gajar, 1997). Many programs teach youth with IDD to act as their own advocates (Agran, Wehmeyer, Cavin, & Palmer, 2008; Balcazar, Fawcett, & Seekins, 1991; Powers et al., 2001; Wehmeyer, Palmer, Agran, Mithaug, & Martin, 2000). Research suggests that transition-age youth who participate in such programs are more likely to achieve their transition goals (Shogren, Palmer, Wehmeyer, Williams-Diehm, & Little, 2012), demonstrate improved academic performance (Fowler, Konrad, Walker, Test, & Wood, 2007), and report increased self-determination (Wehmeyer, Palmer, Shogren, Williams-Diehm, & Soukup, 2013).

Yet the literature continues to document the impact of physical and social environmental barriers on youth with IDD’s engagement in age-appropriate roles and responsibilities. For example, parents of youth with disabilities report that lack of information, poorly educated professionals, and excessive bureaucracy limit participation in school and the community (Forsyth, Colver, Alvanides, Woolley, & Lowe, 2007; Hammal, Jarvis, & Colver, 2004; Law, Petrenchik, King, & Hurley, 2007; Pratt, Baker, & Gaebler-Spira, 2008; Verdonschot, De Witte, Reichrath, Buntinx, & Curfs, 2009). Youth with disabilities recognize that lack of task modifications and negative attitudes limit their inclusion in school and in the community (Kramer, Olsen, Mermelstein, Balcells, & Liljenquist, 2012). Some programs, such as the self-determined learning model of instruction (e.g., Shogren et al., 2012), direct youth to consider barriers to goal attainment. However, these programs do not explicitly ask youth to consider environmental barriers. Such curricula may assume youth have the skills to identify and resolve barriers in the social and physical environment. Yet research demonstrates that youth with disabilities are not as equipped as they could be to manage environmental barriers. One study found that youth with disabilities had less self-reported knowledge of environmental modifications and of their rights to request reasonable accommodations than knowledge regarding their health condition, transportation
safety, and community resources (Betz, Redcay, & Tan, 2003). In another study, youth with cognitive disabilities reported significantly fewer opportunities to identify and request needed accommodations than youth with other disabilities (Powers et al., 2007).

Project TEAM (Teens making Activity and Environment Modifications) was designed to teach youth with IDD to systematically identify environmental barriers and supports, generate modification strategies to address those barriers, and request reasonable accommodations. It addresses a gap in the literature as it focuses only on teaching youth to identify and resolve environmental barriers so they are better able to benefit from more general self-advocacy and self-determination curriculums. This article illustrates how a mixed-methods approach was utilized to conduct a comprehensive formative evaluation of the new Project TEAM self-advocacy program and optimize the program prior to undertaking future research.

**Project TEAM**

Project TEAM is a research-based and theoretically grounded intervention that was developed in collaboration with youth with disabilities (Kramer et al., 2013). It was designed to be accessible to youth with cognitive, physical, and sensory disabilities. The development of the intervention was informed by research that links self-awareness, personal values, goal setting, and self-monitoring with positive transition outcomes for youth with disabilities (Fowler et al., 2007; Shogren et al., 2012; Wehmeyer et al., 2013). Figure 1 depicts the Project TEAM logic model and targeted program outcomes.

Project TEAM teaches a problem-solving process referred to as the “Game Plan” (Table 1). During Project TEAM, trainees apply each step of the Game Plan to a personal activity goal to begin or increase participation in a school or community activity. The Game Plan (Goal, Plan, Do, and Check) is informed by cognitive-behavioral techniques and designed to be parallel with other self-determination and self-advocacy programs (Shogren et al., 2012). A unique feature of the Game Plan are problem-solving questions that direct trainees to attend to aspects of the environment, rather than personal impairments, that make it difficult to participate in activities. Trainees internalize these questions using self-talk (Meichenbaum, 1977) first during structured learning activities in the modules and later during experiential field trips related to their activity goals. The Game Plan Worksheet guides trainees through this problem-solving process and is designed according to universal design for learning (UDL) specifications (National Center on Universal Design for Learning, 2011).

Project TEAM’s conceptualization of the environment is informed by rehabilitation frameworks (American Occupational Therapy Association, 2008; World Health Organization, 2001): 11

---

**Figure 1.** Project TEAM Logic Model.

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Activities</th>
<th>Outputs</th>
<th>Outcomes: Short Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Funding</td>
<td>• Identify personal activity goal</td>
<td>• # of persons trained to deliver Project TEAM (facilitators, research assistants)</td>
<td>• Increase knowledge of environmental factors &amp; modification strategies.</td>
</tr>
<tr>
<td>• Accessible meeting space &amp; training materials</td>
<td>• Complete 8 group modules to learn the</td>
<td>• # of youth with disabilities completing the training</td>
<td>• Increase participation in personal activity goal.</td>
</tr>
<tr>
<td>• Facilitators to implement training &amp; support youth</td>
<td>“Game Plan”</td>
<td>• # of trainings implemented</td>
<td>• Increase competence for everyday activities.</td>
</tr>
<tr>
<td></td>
<td>• Apply the “Game Plan” during experiential learning field trips</td>
<td></td>
<td><strong>Outcomes: Long Term</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Increase participation of youth with disabilities in education, employment, &amp; community life.</td>
</tr>
</tbody>
</table>

---

J. M. Kramer et al. 259
environmental categories represent physical, sensory, social, and system factors such as rules, people, entrances and exits, and light, sound, and smell. The Game Plan helps trainees systematically identify the environmental factors that help or make it harder for them to engage in an activity. Project TEAM also introduces five modification strategies that can be used to resolve barriers in the environment: (a) planning ahead, (b) teaching others about disability, (c) using things differently, (d) doing activities differently, and (e) changing spaces. The modification strategies were created in collaboration with youth with disabilities and informed by a meta-synthesis of research exploring youth’s perceptions (Kramer et al., 2012).

The Project TEAM curriculum, in conjunction with the experiential personal goal trip, was designed so trainees can learn, practice, and internalize the Game Plan problem-solving process. Project TEAM includes 8 group modules guided by a UDL framework; the organization of each module follows a similar structure. Icebreakers at the beginning of each module are designed to encourage fun, risk-free interaction between trainees. Teaching activities introduce trainees to new concepts using traditional didactic approaches such as PowerPoint presentations. Discussions provide trainees with an opportunity to share their ideas with the full group. Learning activities require trainees to apply training concepts to complete worksheets or games; examples of learning activities are provided in Table 1. For example, in the game Environment Uno, trainees match category name or picture cards to practice identifying different environment categories. After learning about three disability rights laws (The Americans with Disabilities Act (ADA), the Individuals with Disabilities Education Act (IDEA), and the Rehabilitation Act), small groups create and perform a “rhyme about rights” that explains the main focus of each law. All Project TEAM activities were designed to facilitate peer support and social learning (Joseph Rowntree Foundation, 2003).

When evaluating a new program such as Project TEAM, the first step is to conduct a formative evaluation. The aim of formative evaluation is to identify program improvements that will

<table>
<thead>
<tr>
<th>Game plan</th>
<th>Game plan self-talk question</th>
<th>Associated project TEAM module</th>
<th>Example project TEAM learning activity/game</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal</td>
<td>What activity would I like to do?</td>
<td>Module 1: Who am I and what would I like to do?</td>
<td>Strengths and Difficulties Checklist</td>
</tr>
<tr>
<td></td>
<td>What part of the environment is getting in the way?</td>
<td>Module 2: Introduction to the environment: What is it?</td>
<td>Environment Uno</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Module 3: Identifying environmental barriers and supports.</td>
<td>Barriers and Supports</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Module 4: Working around barriers: Identifying modification strategies.</td>
<td>Scavenger Hunt</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Module 5: Thinking through strategies: If…then questions.</td>
<td>Dice of Fortune: Using Strategies</td>
</tr>
<tr>
<td>Do</td>
<td>How could I change this part of the environment? What strategy can I use?</td>
<td>Module 7: Communication: Asking for change.</td>
<td>Thinking Through Strategies (Question &amp; Answer game)</td>
</tr>
<tr>
<td></td>
<td>What would happen if I changed this part of the environment or used this strategy?</td>
<td>Module 6: Your rights: Laws about access and reasonable accommodations.</td>
<td>Rhyme About rights Asking for Change Concentration</td>
</tr>
<tr>
<td>Check</td>
<td>Who do I talk to about making this change?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Am I able to do this activity now?</td>
<td>Module 8: Checking the Game Plan: Did I reach my goal?</td>
<td></td>
</tr>
</tbody>
</table>

Table 1
Project TEAM Game Plan and Associated Modules
maximize success for future implementation (Patton, 2012). A program developer may want to understand if a program achieves the anticipated outcomes and identify the most effective approach to documenting those outcomes. Program developers may also want to know if the planned activities were enacted as expected, and determine if any observed outcomes can be reasonably explained by the activities designed to operationalize the program’s proposed mechanisms of change (Kadzin, 1997). Further, a program developer may wish to understand how participants experience the program and the intended outcomes to ensure the program is relevant to their everyday lives and lived experiences (Powers et al., 2007; Priestley, 1998; Wolf, 1978). Thus, the three questions guiding the formative evaluation of Project TEAM were: (1) To what extent do trainees achieve the anticipated outcomes? (2) What is the quality of activities and do they support the attainment of anticipated outcomes? and (3) Do trainees feel that Project TEAM activities and outcomes are enjoyable and relevant to their lives?

Methods

Design

These three formative evaluation questions asked about unique but interrelated phenomena (Greene, Caracelli, & Graham, 1989): outcomes, activities, and experiences. We took a pragmatic approach to identify the best method to answer each question. Using mixed methods in program evaluation can lead to more comprehensive findings and enhance the credibility of the inferences made based on those findings (Greene, Benjamin, & Goodyear, 2001). To identify the optimal data collection approach for each phenomenon, we considered how different data collection methods and data sources could provide unique and complimentary information about the program’s design and outcomes. A single group, repeated-measures design was used to evaluate attainment of anticipated outcomes. Retrospective qualitative observational methods were used to evaluate the quality of the intervention activities. Qualitative observations may enable us to conceptually link information about program implementation with observed outcomes (Greene et al., 1989). Finally, the youth who collaborated in the design of Project TEAM conducted an independent evaluation of trainee experiences using a survey approach; a detailed report of procedures and analysis has been reported in Kramer et al., 2013. This approach to using different methods to examine unique phenomena in has been coined an expansion design as it attempts to provide a more in-depth understanding of program processes and outcomes (Greene et al., 1989).

Participants

Research ethics clearance was secured through a large university in the northeast United States. Project TEAM was delivered to three groups: two groups were conducted in classrooms in an urban public high school, and one group was conducted at an urban after-school program for young adults with IDD. Purposeful recruitment was used to target eligible participants. Inclusion criteria were (a) 12- to 17-years old; (b) a primary diagnosis of a physical, cognitive, or sensory disability; and (c) the ability to attend to a task for 10 minutes and follow two-step directions. The classroom teacher and after school program director identified youth meeting these criteria and then sent home study materials. Research staff then obtained parent permission and youth assent.

The 21 trainees in the study (71.4% male) were ages 15–17 (M = 16.5 years, SD = .83 years) and in Grades 9–11. Thirteen trainees identified as African American and the remainder as Caucasian (n = 3), Hispanic/Latino(a) (n = 3), or mixed race (n = 2). Trainees all received special education services under the following individualized education program qualifying categories: intellectual disability (n =13); blindness (n = 2); autism (n = 2); and multiple disabilities, deafness, and speech/language impairment (each n = 1; 1 missing). All students were served in self-contained classrooms for academic and life skill courses. Eight trainees attended all modules of the Project TEAM intervention program, and 17 missed no more than 1 of the 8 modules. One trainee stopped attending school before completing module 8. Twenty trainees remained enrolled until the end of the study.

Procedure

The eight Project TEAM modules were delivered over 14 weeks (including school vacations). Intervention delivery was adjusted to fit each site’s schedule; both sites completed one module per week. At the high school, the training was provided
two times a week; each session lasted 70 minutes (140 minutes total). At the after-school program, trainees met once a week for 120 minutes. The intervention was co-led by an occupational therapist (the first author) and a youth specialist from the local Center for Independent Living who identified as an individual with a disability (the last author). Graduate students provided assistance throughout the duration of the training. All training staff attended a weekly, 1-hour meeting to review the upcoming module. All training staff completed field notes immediately after each session, and all training sessions were video and audio recorded.

Modules 1–7 were completed in weeks 1–9. After module 7, from weeks 10 and 13, trainees worked individually on personal activity goals. The self-identified activity goals covered a range of topics including (a) using community resources such as public transportation and public libraries, (b) applying for employment at a local retailer, and (c) trying new leisure activities such as dance class or flag football. Trainees had at least one coaching session with an interventionist to plan or prepare for their activity. Trainees then engaged in their personal goal activity with the support of intervention staff and when available, a peer. Module 8 was held during week 14 after all trainees finished their coaching sessions.

Assessments were administered by the primary interventionist or a trained graduate assistant. Baseline assessments were administered no more than 3 weeks prior to module 1. A progress Project TEAM Knowledge Test was administered immediately following module 7. All assessments were again administered immediately following the completion of module 8. The youth panel designed and administered a picture-based survey and a series of open-ended questions to each group after module 8. See Figure 2 for a depiction of the study design.

**Measures**

The project TEAM knowledge test. This test, developed for this study, provides standardized prompts to assess trainee’s ability to answer questions while using resources such as the Game Plan Worksheet. The test is divided into two sections: (1) fixed response questions about environment and strategy knowledge, and (2) open-ended responses to two “participation problem stories” that assess problem solving using Project TEAM concepts. The same two participation problem stories were administered at each assessment period. The fixed response question is scored as total correct. The open-ended section is scored by coding transcripts of spoken solutions to the participation problem stories; a rating of 1 is given each time trainees appropriately apply training concepts or give environmentally focused solutions.

**Goal Attainment Scaling (GAS).** GAS was used to measure achievement of individualized goals in a self-selected activity (personal activity goal). Three additional knowledge application goals addressed the trainees’ ability to (1) identify environmental factors in the activity, (2) identify and use modification strategies during the activity, and (3) communicate about the activity goal and any needed environmental modifications. The knowledge application goals could be scaled together because they all applied to the trainees’ personal activity goal. Activity and knowledge application goals were identified in collaboration with trainees during the baseline interviews. Trainees selected an activity of their choice and had the option of viewing pictures of activity ideas if they could not spontaneously identify an activity of interest. Trainees also indicated if they would like to learn about parts of the environment, strategies to resolve environmental barriers, and asking for changes in their environment. Goal

---

**Figure 2.** Project TEAM Study Design^a^.

^a^Child Occupational Self-Assessment = COSA; Goal Attainment Scaling = GAS.

©AAIDD
DOI: 10.1352/1934-9556-52.4.258
attainment levels for the activity and three knowledge application goals were written prior to module 1 by the primary interventionist and a graduate research assistant who was not involved in implementing the training. Expected outcomes were estimated from baseline performance on the Project TEAM Knowledge Test. A 6-point scale was used to classify goal attainment levels and reduce positive outcome bias (Sakewski, Boyd, & Ziviani, 2007; Steenbeek, Ketelaar, Galama, & Gorter, 2007): 0 indicated expected goal attainment; values above 0 indicated higher than expected goal attainment; values below 0 indicated less than expected goal attainment, no change, or decline. After the intervention, goal levels for each trainee were scored by the training staff who conducted the coaching session, and reviewed by a second training staff member who had extensive knowledge of the trainee’s performance. Field notes and assessment results were used to justify final goal scoring decisions; consensus between the staff ratings was achieved for all final GAS scores.

The Child Occupational Self-Assessment (COSA). The COSA includes 25 questions about activities that youth engage in at home, school, and in their communities. Youth indicate their level of competence for each activity using a 4-point difficulty scale. The COSA has good content, structural, and substantive validity for youth with disabilities, as given by acceptable item and child fit and unidimensionality evaluation (Kramer, Kielhofner, & Smith Jr, 2010; Kramer, Smith, & Kielhofner, 2009).

Activity quality. The research team conducted structured, qualitative observations of video and audio recordings of all Project TEAM activities across the three implementation groups. One observer completed detailed written descriptions of positive and negative trainee reactions to each Project TEAM activity for each implementation group. The qualitative observations informed Project TEAM activity quality ratings. An activity quality rating scale was developed specifically for the formative evaluation of Project TEAM. Scales described three facets of activity characteristics: (1) engagement (extent to which activity is enjoyable, interesting, relevant, and fun to trainees); (2) understanding of activity rules (extent to which trainees understand and follow rules and steps of activity); and (3) understanding of training concepts (extent to which the activity helps trainees learn, understand, and/or apply concepts introduced in Project TEAM). Each facet is rated on scale from 1 (most to all students were not engaged or did not understand) to 4 (most to all students were engaged or demonstrated understanding). The observer (a graduate student) who completed the written qualitative observations for an activity also rated the quality of each activity. Then, the detailed written observations for each activity were reviewed by a second team member and the first author. The reviewers then triangulated the quality rating. As needed, quality ratings were revised to reflect consensus.

Trainee experiences. The youth panel and the first author developed an accessible survey that included questions about curriculum materials and delivery approaches such as worksheets and the use of peer support during activities. The survey also included questions about the impact of Project TEAM on trainees’ everyday lives. The content of this survey has been described in Kramer et al., 2013.

Analysis
Change on the Project TEAM Knowledge Test was examined using a repeated measures ANOVAs for all trainees with three completed assessments (n = 20); separate tests were conducted for the sum correct and open-ended scores. An adjusted F was obtained if the assumption of sphericity was violated.

GAS scores for each trainee’s personal activity and three knowledge application goals were transformed to a T-score using the following formula: 

$$T = 50 + C(\Sigma \chi_i)$$

where \(\chi_i\) is the level of goal attainment for each goal and C is a constant that depends on the number of goals contributing to the T-score (for four goals, C = 4.56; Kiresuk, Smith, & Cardillo, 1994). This formula has been used with the 6-point GAS scale used in this study (Steenbeek et al., 2007; Turner-Stokes & Williams, 2010). A mean T-score of 50 indicates that the sample achieved the expected level of change. All goals were scaled for all trainees using all available data, including the trainee lost to attrition.

Baseline and outcome COSA responses were entered into an existing database of over 500 COSA responses. A Rasch Rating Scale Model (RSM) analysis (Wright & Masters, 1982) was then performed to convert the ordinal scale responses to interval-level measures. Paired samples t-tests were used to examine changes from baseline to outcome.

Effect sizes for the Project TEAM Knowledge Test and COSA were calculated using the following formula: 

$$[\text{Mean score at outcome} - \text{Mean score at baseline}] / \text{Standard deviation}$$

J. M. Kramer et al. 263
Descriptive statistics were used to examine intervention activity quality ratings by activity category. In each activity category, the activity that received the highest total quality rating was selected for more in-depth review. We further examined the qualitative observations from these activities to better understand trainee reactions and identify aspects of the activity that effectively supported anticipated outcomes.

For the trainee experiences survey, the youth panel entered data into Excel and created histograms. The group then compared graphs to identify patterns in the trainees’ responses (Kramer et al., 2013).

Results

Knowledge of Environmental Factors and Modification Strategies
There was a significant increase in the sum correct scores over time, $F(2, 20) = 35.40, p = 0.0$ (Table 2). Post-hoc comparisons using Bonferroni $t$-tests indicated significant ($p = .00$) differences between baseline ($M = 9.2, SD = 3.62$) and progress assessment ($M = 20.1, SD = 9.6$), as well as baseline and outcome assessment ($M = 21.85, SD = 10.57$). Thirteen of 20 trainees had greater than 100% change in their sum correct scores between baseline and outcome assessment. The standard error of the mean indicated the sum correct score had the capacity to capture underlying change (Lexell & Downham, 2005; $SE = 1.99$). There was a small, nonsignificant increase in the open-ended knowledge application score over time, $F(2, 20) = 3.11, p > .05$ (Table 2). Baseline scores were the lowest ($M = 1.00, SD = 1.81$) and progress scores were the highest ($M = 2.44, SD = 3.48$) and then decreased slightly at outcome assessment ($M = 1.72, SD = 1.7$). The sum correct and knowledge application effect sizes for the test were large (3.49) and moderate (0.40), respectively. No ceiling or floor effects were observed for any assessment period.

Goal Attainment
$T$-scores indicate trainees achieved only partial goal attainment; across all four goals, the mean $T$-score was 42.66 ($SD = 7.15$), and $T$-scores ranged from 27.17–55.45. Sixteen (76.2%) trainees met the expected attainment level (rating of 0 or higher) for at least one goal, 9 (42.9%) and 4 (19%) trainees met the expected attainment level for at least two and three goals, respectively. No trainees met the expected attainment level for all four goals. Goals related to personal goal activities had the highest frequency of goal attainment (Table 3).

Self-Reported Competence
There was no significant difference, $t(19) = 0.15, p > .05$ between self-reported competence on the COSA at baseline, $M = 1.26, SD = 1.32$ and outcome, $M = 1.03, SD = 1.46$. The effect size of the COSA competence scale was small ($0.17$), and no trainee exceeded 10% change between baseline and outcome, $M = -0.12, SD = 2.07$.

Activity Quality
The mean quality ratings indicate that trainees demonstrated the highest level of “engagement” during teaching activities and learning activities and games (Table 4). The mean quality ratings also suggest that trainees demonstrated the best

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of squares</th>
<th>Df</th>
<th>Within subjects variance ($M^2$)</th>
<th>$F$ ratio</th>
<th>$p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum correct score (repeated measure)</td>
<td>1879.30</td>
<td>2</td>
<td>1504.06</td>
<td>35.40</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>1008.70</td>
<td>38</td>
<td>42.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open ended knowledge application score (repeated measure)</td>
<td>18.78</td>
<td>2</td>
<td>12.76</td>
<td>3.11</td>
<td>.08</td>
</tr>
<tr>
<td>Error (time)</td>
<td>102.56</td>
<td>38</td>
<td>4.12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2
Project TEAM Knowledge Test Changes Over Time

Formative Evaluation Project TEAM
understanding of activity rules” during teaching activities, which primarily consisted of answering questions while viewing PowerPoint presentations. All activity categories had a mean quality rating for “understanding of training concepts” of less than 2.5. Discussions consistently received the lowest average rating on all three facets of activity quality.

Analysis of the qualitative observations from the highest rated activities identified two common activity qualities. First, many top-rated learning activities and games such as Asking for Change Concentration involved teamwork and trainees were observed to enjoy working with or competing against peers in a low-risk learning environment. Second, top-rated activities included supports that helped trainees complete the activity successfully. Examples of supports included (1) a cheat sheet of potential answers (as in the learning activity Dice of Fortune), (2) rules adapted from familiar games (such as the learning activity Environment Uno),

Table 3
Goal Attainment by Goal Area (N = 21)

<table>
<thead>
<tr>
<th>Goal area</th>
<th>Goal attainment level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>−2 (Below baseline)</td>
</tr>
<tr>
<td>Participation in self-selected activity</td>
<td>9.5%</td>
</tr>
<tr>
<td>Knowledge and application of environmental factors</td>
<td>14.3%</td>
</tr>
<tr>
<td>Knowledge and application of modification strategies</td>
<td>14.3%</td>
</tr>
<tr>
<td>Asking for change in the environment (n = 20)</td>
<td>5%</td>
</tr>
</tbody>
</table>

*aOne trainee chose not to set a goal in this area.

Table 4
Activity Quality Ratings

<table>
<thead>
<tr>
<th>Activity categoryb</th>
<th>Engagement M (range)</th>
<th>Understanding of activity rules M (range)</th>
<th>Understanding of training concepts M (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Icebreakers (n = 7)</td>
<td>2.86 (2–4)</td>
<td>2.86 (2–4)</td>
<td>N/A</td>
</tr>
<tr>
<td>Teaching Activities (n = 5)</td>
<td>3.00 (3)</td>
<td>3.20 (3–4)</td>
<td>2.20 (1–3)</td>
</tr>
<tr>
<td>Discussions (n = 5)</td>
<td>2.40 (2–3)</td>
<td>2.80 (2–3)</td>
<td>1.80 (1–2)</td>
</tr>
<tr>
<td>Learning Activities &amp; Games (n = 20)</td>
<td>3.00 (1–4)</td>
<td>2.95 (1–4)</td>
<td>2.00 (1–3)</td>
</tr>
</tbody>
</table>

*Icebreakers (fun, risk free interaction); Teaching activities (didactic approaches to introduce new concepts); Discussions (youth share ideas and reflect as a full group); Learning activities & games (youth use and apply concepts to complete worksheets and games).
and (3) pictures that helped depict new concepts (as in the teaching activity Environment Slide Show). However, even these top-rated activities were not fully effective in facilitating application of Project TEAM concepts. Trainees had a difficult time using new and unfamiliar Project TEAM vocabulary during activities. In addition, the trainees struggled to understand and solve the problems presented in the picture or text-based stories used in the Project TEAM activities. Finally, qualitative observations suggest that some games did not require trainees to apply concepts. Instead, trainees used the supporting resources to match symbols or the number of letters to guess answers rather than problem solve to apply concepts and to derive an answer.

**Trainee Experiences**

Results have been reported in full in Kramer et al., (2013), but we will highlight key findings that inform our formative evaluation. Results suggest that trainees found the applied and interactive components of Project TEAM enjoyable. For example, 93.7% and 100% of trainees reported that the Project TEAM games and activities and field trips, respectively, were “good/really good.” Conversely, responses also suggest problems with the design of Project TEAM materials and activities. For example, 25% of trainees reported that materials such as the Game Plan Worksheet and Asking for Change Script were “bad/really bad” and 62.5% reported that the concepts were “sometimes” or “always” difficult to understand. Best practices such as UDL and social learning utilized during implementation appear to be effective: 93.7% of trainees reported that the symbols used in the training “sometimes” or “always” helped them understand things they were learning, and 75% and 78% reported helping or being helped by a peer respectively during Project TEAM. Finally, trainees reported that Project TEAM had an impact on their everyday lives: 87.4% identified supports, 93.8% identified environmental barriers, 43.8% used a Project TEAM strategy to change their environment, and 81.2% asked for a change they needed in the environment.

**Discussion**

The use of a mixed-methods expansion design examining three phenomena (outcomes, activities, and experiences) provided a comprehensive understanding of the strengths and limitations of Project TEAM. Although the findings from this evaluation can inform the future development of Project TEAM, other researchers may find a similar approach to examining outcomes, activities, and experiences useful when evaluating new programs for transition-age youth with IDD.

This formative evaluation was guided by three research questions. In response to the first question, “To what extent do trainees achieve anticipated outcomes?” results suggest that trainees are not attaining outcomes at the expected level. Positive findings from the fixed response section of the Project TEAM Knowledge Test and the trainee experiences survey suggest that the universally designed resources developed for Project TEAM helped trainees recall Project TEAM concepts during assessment. Project TEAM also increased participation in a self-selected activity for just over half of the trainees. However, objective assessment data from the open-ended problem-solving section of the Project TEAM Knowledge Test and GAS suggest that trainees had a difficult time applying this new knowledge to problem solve solutions to environmental barriers. These findings are in contrast with trainee self-reports that Project TEAM enabled them to identify environmental factors and request changes in their environments during their everyday lives.

In response to the second question, “What is the quality of Project TEAM activities and do they support anticipated outcomes?” the answer also points to the need for significant revisions. Qualitative observations reveal high variability in the quality of Project TEAM activities. Further, although trainees indicated that they enjoyed Project TEAM games, they reported the most dissatisfaction with the primary resources developed to support the problem-solving approach taught in Project TEAM including the Game Plan Worksheet and the Asking for Change Script.

In response to the third question, “To what extent do trainees feel that Project TEAM is enjoyable and relevant to their lives?” findings suggest Project TEAM has the potential to make a meaningful impact on transition-age youth. Trainees enjoyed the individualized and experiential aspects of the program, and 75% reported that the things they learned in Project TEAM helped them in their daily lives (Kramer et al., 2013).

The discrepant findings of this formative evaluation make it difficult to draw conclusions about the efficacy of Project TEAM. However, the
The aim of formative evaluation is not to establish efficacy but to better understand how a program could be improved. Other scholars have cautioned against drawing conclusions prematurely on the basis of formative evaluation results (Patton, 2012). This discussion will illustrate how findings across the three phenomena examined in this formative evaluation provided not only a better understanding of the potential problems in the current Project TEAM curriculum, but also allowed the team to identify the optimal way to address those problems. The results do suggest that three program components require revision before further efficacy testing: the measurement of expected outcomes, the quality of activities and materials, and the delivery protocol.

First, findings indicate that the outcomes articulated in the Project TEAM logic model and measures used to operationalize those outcomes require revision. During the Project TEAM Knowledge Test, research staff observed problems with the written “participation problem stories.” Trainees could not remember aspects of the story or did not understand the direction “What advice would you give to the teen in the story so he/she could do this activity?” Developing video versions of these stories may be more accessible, and asking trainees to problem-solve an actual personal, rather than hypothetical, participation problem may result in a more valid assessment of knowledge application. In addition, the GAS procedures should be revised, as low levels of goal attainment may be a product of poorly written goals, poor clinical judgment regarding trainee potential, or unrealistic expectations for training outcomes (Kiresuk et al., 1994). The experiences from this formative evaluation suggest that the ability to apply Project TEAM concepts to identify, resolve, and communicate about environmental barriers may follow a linear progression. Using this hypothesis, we could create goal attainment continuums for each knowledge application area that meet GAS quality standards (Kiresuk et al., 1994). Points of entry for each knowledge application continuum could be selected based on initial performance on the Project TEAM Knowledge Test. This may result in more appropriately scaled goals that could be standardized across trainees for future large-scale studies.

A secondary expected outcome of Project TEAM was an increase in competence, yet the COSA found no significant change after the completion of the program. The trainee experiences survey suggests youth felt more confident in their ability to respond to environmental barriers upon completion of Project TEAM. Thus, revising the Project TEAM logic model and evaluation design to include the outcomes of self-determination and self-efficacy may capture these changes in future studies.

A second program component requiring revision is the quality of Project TEAM activities and materials. Findings across all three phenomena point to the types of revisions that may be more likely to support anticipated outcomes. For example, observations revealed that discussion activities were not an effective pedagogical tool. In contrast, trainees were observed to respond positively to facilitator-led teaching activities. The pedagogical teaching methods used in these activities, such as PowerPoint and a question/answer format, were highly structured and presented concepts in a concrete way with definitions and pictures. Trainees with IDD may benefit if open-ended discussion questions are asked after providing concrete examples and stories. Observations demonstrated that trainees utilized the universally designed supports during learning activities but were not required to articulate the reasoning for their answers; this led to a reliance on guessing. Game rules and supports should be revised to ensure trainees actively problem solve. Observation data suggest that trainees who worked in pairs or small groups were more likely to problem-solve together, and incorporating more group activities into the training may further facilitate active problem solving.

Resources such as the Game Plan Worksheet could be improved by incorporating aspects of UDL that were observed to support trainees’ successful completion of learning activities. For example, activities provided multiple options for engagement and completion by having picture-based lists available for reference, providing example responses that trainees could choose before progressing to the next step of an activity, or giving the opportunity to create new examples that were more relevant to an individual trainee. In a similar way, trainees may better navigate the Game Plan problem-solving process if the Game Plan Worksheet was revised to reduce the amount of text, to provide more white space for writing and individualization, and to give specific checkbox examples of things to consider during each step. Finally, the strategy categories used in the Game Plan require...
The two strategy categories recalled by the most trainees were “Plan ahead” and “Change Spaces” (9 of 20 trainees recalled these strategies at outcome assessment). These two categories describe a concrete action that can resolve environmental barriers. In contrast, only two to four trainees could recall any of the other strategy categories at outcome assessment. Creating strategy category names that describe concrete actions may reduce confusion and enhance application to real-life problems.

A third program component that requires revisions is the Project TEAM implementation protocol. Trainees with IDD may need additional opportunities to practice skills across contexts of school, home, and the community to internalize the Game Plan self-talk questions and increase their capacity to generate and request modifications. Involving parents and other professionals across various contexts may support more successful outcomes by providing additional opportunities to practice the application of Project TEAM concepts. New procedures are also needed to identify youth who are most appropriate for Project TEAM. Observations suggest several trainees did not understand how to categorize objects, which is an essential prerequisite to using categories of the environment and strategies taught in the Game Plan. Some trainees were observed to require extensive support for reading and spelling, which interfered with completion of learning activities. Future studies will need to develop more specific inclusion criteria and should explore if reading level, cognitive impairment, or adaptive behavior are associated with success in Project TEAM.

This study demonstrates the insights that can be gained when researchers go beyond the measurement of primary outcomes and use multiple methods to evaluate new programs for transition age youth with IDD. In this study, we purposefully selected complimentary methods to better understand the relationship between the current program design and anticipated outcomes (Greene et al., 1989). For example, systematic qualitative observations provided a better understanding of both the mechanisms that detracted from and supported the attainment of anticipated program outcomes. When these findings were complemented by youth self-reports, findings suggested that both the design of the activities and the articulated secondary outcomes required revision. Without this information, researchers may have reviewed the outcome data and concluded prematurely that the Game Plan was not an effective problem-solving approach. Further, researchers may have continued to allocate valuable resources to the measurement of inappropriate outcomes or to numerous rounds of curriculum revisions.

Other evaluators may experience similar benefits by gathering information about these three phenomena using multiple methods; questions, such as those posed in Table 5, may help evaluators identify the optimal data-collection approach for their program outcomes, activities, and experiences. For example, questions about the purpose, structure, and organization of program activities may help evaluators determine the specific activity, data-collection approach, and source that provide the best understanding of how or why an outcome of interest was or was not obtained. These questions do not value one epistemological approach to theory building (inductive vs. deductive), stance (subjective vs. objective), or inference (context dependent vs. generalizable) over another. Rather, questions about the program’s mechanism of change and that consider a range of potential data collection methods and sources may help evaluators operationalize a third, pragmatic approach that values abduction (moving between inductive and deductive reasoning), intersubjectivity, and transferability (Morgan, 2007). Evaluators looking to optimize the integration of quantitative and qualitative methodologies are also encouraged to utilize mixed-method typologies to ensure the legitimacy, or validity, of multiple methods in evaluation (Onwuegbuzie, Johnson, & Collins, 2011).

Implementing this comprehensive approach to program evaluation has unique challenges. Systematic qualitative analysis, such as the analysis of activity quality, often requires a larger team and more time than measuring and analyzing quantitative outcome data such as test scores. Funding and productivity constraints can make it difficult to delay the dissemination of primary outcomes to wait for the analysis of complimentary data about program processes and experiences. However, the potential benefits for program implementation and, ultimately, for participants in the program, may outweigh the challenges.

This formative evaluation had several limitations that are inherent to single group designs. Without a comparison group, it is difficult to determine if change reflected increase in skills or other artifact such as deviation to the mean or developmental maturation. Changes in the Project
<table>
<thead>
<tr>
<th>Phenomena</th>
<th>Questions about the mechanisms of change</th>
<th>Questions about data collection methods</th>
<th>Questions about data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcomes</td>
<td>• Are anticipated outcomes expected to occur before or after transition?</td>
<td>• How can outcomes be demonstrated in real-life contexts vs. contrived contexts (i.e., testing)?</td>
<td>• What stakeholders may be able to report on outcomes (i.e., transition-age youth, family, other professionals, employers)?</td>
</tr>
<tr>
<td></td>
<td>• Are anticipated outcomes expected to occur in one context (school, work, community) or across contexts?</td>
<td>• How can outcomes be measured objectively and subjectively?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• How do program activities operationalize the hypothesized mechanisms of change?</td>
<td>• How do transition-age youth engage in the activities (individual, small group, or large group)?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Are different activities expected to support the attainment of specific outcomes?</td>
<td>• How do transition-age youth demonstrate their engagement in the activities (verbally, written, experientially)?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• What is the hypothesized relationship of the activity context (school, work, community) on participant outcomes?</td>
<td>• Will stakeholders feel more comfortable sharing their experiences alone (privately) or with other stakeholders (in a group)? With outsiders or known individuals involved in the program? With individuals in positions of authority or peers?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• How do stakeholders experience the unique or novel approaches used in the program to meet the needs of transition-age youth?</td>
<td>• Will stakeholders feel more comfortable expressing their perspectives verbally, in writing, or other expressive methods?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• What is the value of anticipated outcomes to transition-age youth and other stakeholders?</td>
<td>• Will stakeholders more effectively share their perspectives during their engagement in program activities (concurrently) or by reflecting on program experiences (retrospectively)?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• How or why do transition-age youth and other stakeholders think the program supported change?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiences</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TEAM Knowledge Test scores could be a practice effect; however, this is not likely as no significant difference was found between progress and outcome assessment and no ceiling effects were observed. Other limitations include a small sample size and limited power, unmasked assessors, a convenience sample that limits generalizability, and unequal exposure time across groups. Video data informing the Activity Quality Ratings were only viewed by one member of the research team, so it is possible that some trainee responses were not documented. Future evaluation of Project TEAM should incorporate a comparison group, use masked assessors, and monitor fidelity to training activities.

**Conclusion**

This mixed-methods formative evaluation that gathered information about outcomes, activity quality, and participant experiences provided valuable information about the program elements that may and may not facilitate the outcomes targeted by a new intervention, Project TEAM. Lessons learned from this comprehensive evaluation can inform future revisions to Project TEAM and may be equally relevant for other programs targeting transition-age youth with IDD.

**References**


Received 11/12/2012, accepted 4/03/2014.

This study was funded by Deborah Munroe Noonan Memorial Research Fund, Bank of America, N.A., Trustee. PI: Kramer. This study received ethical clearance from the Boston University Institutional Review Board.

Thank you graduate students Alicia Heinz, Sarah Kreditor, Christine Lin, Eleanor Mockler, Jenny Pianucci, and Allie Taylor.

Authors:
Jessica M. Kramer, Boston University, MA; Kristin Roemer, Arlington, VA; Kendra Liljenquist and Julia Shin, Boston University, MA; Stacy Hart, Boston Center for Independent Living, MA.

Address correspondence concerning this article to Jessica M. Kramer, Boston University, Department of Occupational Therapy & ScD Program in Rehabilitation Sciences, 635 Commonwealth Ave., Boston, MA 02215 (e-mail: kramerj@bu.edu).