FIRST® LEGO® League Jr. Season Pass Model Evaluation

Final Report

March 12, 2019

Submitted by

Melissa Collins, Rosalinda Nava, and Ardice Hartry

The Research Group
The Lawrence Hall of Science
University of California, Berkeley

Table of Contents

Introduction

The *FIRST*® LEGO® League Jr. program is designed to build an interest in science for youth aged 6-10 years old. Youth are guided by adult coaches through working with LEGO elements and moving parts to build ideas, concepts, and models and present them for review. Youth also solve a "real-world challenge" using research, critical thinking and imagination. In 2016, *FIRST* LEGO League Jr. was awarded a multi-year grant to increase the program's reach and access to underrepresented and underserved youth in the US, Mexico, Canada, South Africa, Ukraine and two other countries/areas through the creation and implementation of the *FIRST* LEGO League Jr. Season Pass Model. The *FIRST* LEGO League Jr. Season Pass Model reduces costs for sites that previously would have been unable to access the *FIRST* LEGO League Jr. program by allowing sites to work with multiple teams within the same season using a single license. Through this model, the *FIRST* LEGO League Jr. program sought to reach over 45,000 youth at 224 sites over three years.

The purpose of this project is to evaluate the implementation of the *FIRST* LEGO League Jr. Season Pass Model and measure its short-term outcomes for participating youth, focusing primarily on the United States, with additional work in Canada and Mexico.

Evaluation Questions

The primary questions we aim to address in this report are:

1. Who participated in the *FIRST* LEGO League Jr. Season Pass Model? What was the reach of the program in terms of geography and diversity? Does the program meet its stated goals to reach underserved youth in the US, Canada, and Mexico, specifically with regards to gender, race, disability status, and income?

2. How did sites implement the Season Pass Model? What challenges did coaches and site administrators face? Is the Season Pass Model implemented with fidelity across sites?
3. Is there evidence of promise in increasing youth's positive attitudes toward STEM, Emerging Activation, learning through play, and 21st century skills? Is there variation in outcomes by subgroup? That is, do participants show equal levels of learning regardless of gender, age, or other variables?
4. What are the *FIRST* LEGO League Jr. Season Pass Model's strengths and weaknesses? What are the areas for program improvement?
5. In what ways can a case study of implementation in Mexico inform dissemination of *FIRST* LEGO League Jr. to other international sites?

This report focuses primarily on data collected during the 2017-2018 implementation year (Year 2). Occasional references to or comparisons with the 2016-2017 implementation year are highlighted when compelling or informative. For a more thorough description and analysis of the 2016-2017 implementation year, see the "*FIRST*® LEGO® League Jr. Season Pass Model Evaluation Mid-Project Report" (Collins & Nava, 2017). Additionally, the main body of this report focuses on Evaluation Questions 1-4; for Evaluation Question 5, see Appendix A: International Case Site Report.

Methods

Sample

A total of 111 sites participated in Year 2 of *FIRST* LEGO League Jr. Season Pass. Most (64%) of these sites also participated in Year 1, while 40 sites (36%) were new in the 2017-2018 school year. Year 2 sites were located across 30 states, 3 Canadian provinces, and Puerto Rico,

and included schools, school districts, nonprofit organizations, and other organizations serving youth. These sites are shown in Figure 1 below.

Figure 1. *FIRST* LEGO League Jr. Season Pass Sites in the 2017-2018 Season

As shown in the map, sites were not evenly distributed across the United States, with a much higher concentration in the eastern half of the country, a smaller representation of sites from the West/Pacific Northwest, and no sites in the Mountain States/Mountain West (MT, ID, WY, NV, UT, CO). This distribution likely reflects the applications received, but does indicate potential areas for increased recruitment in the future.

Design

To investigate the evaluation questions, we amalgamated registration data, surveyed all site administrators, and collected additional information from 12 case study sites. The case study sites were selected based on geographic location, implementation model (e.g., nonprofit implementing the program in an afterschool setting, school teacher implementing the program

during school hours, etc.), and willingness to participate. The final 12 case study sites were agreed upon by the evaluation and *FIRST* teams. In addition to the site administrator survey, case study sites were also asked to participate in observations, coach surveys, parent surveys, and youth surveys.

Instruments

Registration data. Registration data were collected to document who participated in *FIRST* LEGO League Jr. through the Season Pass Model and determine the extent to which the program met its goal to reach 45,000 youth across three years, 90% of whom would be underrepresented or underserved in STEM, and 40% of whom would be female. Data collection occurred in the spring of 2018 and was coordinated by FirstPic, Inc., a project management consulting firm specializing in data management. FirstPic was tasked with designing a data submission system to collect student-level demographic data from site administrators across the *FIRST* LEGO League Jr. Season Pass network. Using a form designed by FirstPic, site administrators were asked to report each participating student's gender, race, ethnicity, free/reduced lunch eligibility, single parent household status, and special education status.

Site Administrator Survey. The site administrator survey was administered online using the *Qualtrics* platform. Initial invitations and reminders were sent to all site administrators through email by *FIRST* staff. The survey contained approximately 55-67 questions, depending on whether or not site administrators responded to follow-up questions. Topics related to sites (e.g., location, schedules), the Expo (e.g., youth attendance), youth (e.g., number participating), coaches (e.g., number participating, background experience, success), program implementation (e.g., WeDo 2.0, satisfaction, challenges), and site administrators (e.g., background, plans for

next year). Questions were posed in a variety of formats, including open-ended responses, Likert-scale ratings, and multiple choice.

Coach Survey. The coach survey was administered online using the *Qualtrics* platform. Initial invitations and reminders were emailed by the evaluation team to site administrators to distribute to coaches at case study sites. The survey was administered once, near the end of the program. The survey contained 61-64 questions, depending on whether or not coaches responded to follow-up questions. Topics related to teams (e.g., number of youth on teams, number of teams coached), team meetings (e.g., schedule, preparation), the Expo (e.g., youth participation, parent attendance, outcomes), youth outcomes (e.g., development in STEM skills, teamwork), experience as a coach (e.g., challenges, goals), overall reflections (e.g., satisfaction, quality), and coach background. Questions were posed in a variety of formats, including open-ended responses, Likert-scale ratings, and multiple choice.

Parent Survey. Parents had the option to complete the parent survey either online using the *Qualtrics* platform or on paper. The evaluation team communicated with each case study site to discuss which option would work best for their audience. Then, the evaluation team sent sites either index cards including a link to the electronic version of the survey, full paper copies of the survey, or both, depending on which process the site preferred. The survey was administered once, near the end of the program. The survey contained 29-43 questions, depending on whether or not parents were aware of what their child did in *FIRST* LEGO League Jr. and/or responded to follow-up questions. Topics related to the Expo (e.g., satisfaction, attendance), parent involvement in program, and family background (e.g., family views on science, demographics). Parents were also asked to talk with their child about their experience in the *FIRST* LEGO

League Jr. program at the end of the survey. Questions were posed in a variety of formats, including open-ended responses, Likert-scale ratings, and multiple choice.

Youth Survey. The youth survey was administered on paper by *FIRST* LEGO League Jr. coaches, with the exception of one site where a member of the evaluation team administered the survey. The survey was administered once, near the end of the program. The survey contained 6 demographic questions, 14 retrospective-pre/post questions, and 2 open-ended questions. These questions were drawn from the *Emerging STEM Learning Activation* survey, developed by the Activation Lab (www.activationlab.org). The *Emerging STEM Learning Activation* survey (for use with 6-10 year-olds) is based on the *STEM Learning Activation Survey* (intended for 10-14 year-olds) with some modifications in order to be more developmentally appropriate. First, the retrospective-pre questions and post questions were grouped together in separate sections. That is, participants first reflect on their perceptions *before* they started on their team in one section, and then on their perceptions *now* in a separate section. Second, their responses were across a five-point scale using smiley faces as answer choices to indicate level of agreement. Coaches were provided with a survey implementation script to support them in administering the survey with their youth. The script provided language to help guide youth through switching between reflecting on themselves prior to participating in *FIRST* LEGO League Jr., and after having participated in the program.

Observation Protocol. Observations were conducted by the evaluation team once at each case study site at a day and time scheduled with local site administrators and coaches. At sites where the program was delivered in a variety of settings (e.g., afterschool, during school hours), the evaluation team attempted to observe each setting. The observation protocol addressed logistics (e.g., location type, number of youth per group, room layout, activity observed), activity

implementation (e.g., quality of instruction, fidelity of implementation), materials, student outcomes, balance between student-led and instructor-led activities, and salient examples of constructs of interest. Questions were posed in a variety of formats, including open-ended responses and Likert-scale ratings. Evaluators also took detailed running notes of all activities and interactions observed.

Results - Evaluation Question 1: Participants

The present section focuses on Evaluation Question 1: *Who participates in the FIRST LEGO League Jr. Season Pass Model? What is the reach of the program in terms of geography and diversity? Does the program meet its stated goals to reach underserved youth in the US and Canada, specifically with regards to gender, race, disability status, and income?* To address these questions, we drew from registration data and site administrator surveys.

Registration Data Analyses

In the registration data, there were missing data at the site, variable, and child level, prohibiting an exact count of participating youth during the 2017-2018 program season of *FIRST* LEGO League Jr. Season Pass. Data were received from 69 of the 111 sites. However, two of these organizations entered data for fewer than 5 youth, so these organizations were dropped from analyses, leaving 9,050 youth across 67 organizations. These sites had an average of 135.1 youth per site. Extrapolating from these 67 sites (i.e., using the registration data's mean of 131.2 youth per site), it is estimated that 14,996 youth participated across the 111 sites in Year 2. Combining data from 2016-2017 (7,340 youth) and 2017-2018, approximately 22,336 youth participated across the two years, unadjusted for youth who participated both years (i.e., unique participants may be lower). This is approximately 49% of the stated goal to reach 45,000 youth across three years, and does not include participants from international sites.

Missing Data

As previously stated, in total, the initial registration database included 9,055 youth across 69 organizations, and two organizations were dropped after it was determined they had entered data for fewer than 5 youth each, leaving 9,050 youth across 67 organizations. Further, two organizations were missing data for all youth for all variables (i.e., they entered each student into the database but did not provide any data on that student). These two organizations were included for overall count analyses (above), but deleted from the dataset for all further analyses looking at demographics, resulting in a final database with 8,833 youth from 65 organizations.

The registration data fields included each student's race, ethnicity, gender, free/reduced lunch eligibility, single parent household status, and special education status. However, there were concerns with the validity of the responses for single parent household status and special education status, which were reported as "0" for every student. Therefore, these two variables were dropped from analyses.

Within the remaining variables (gender, race, ethnicity, and eligibility for free/reduced lunch), there were missing data for each variable, presumably because not every site administrator had access to this information for each student in their program. In cases where more than 50% of data for a variable was missing from an organization, we dropped the organization from analysis of that particular variable because we could not be confident in how to interpret the missing cases (e.g., why was race available for some youth but not others?). Cases where site administrator selected "prefer not to say," "not specified," or "not sure" were treated as missing data because we did not have enough information to categorize them more precisely.

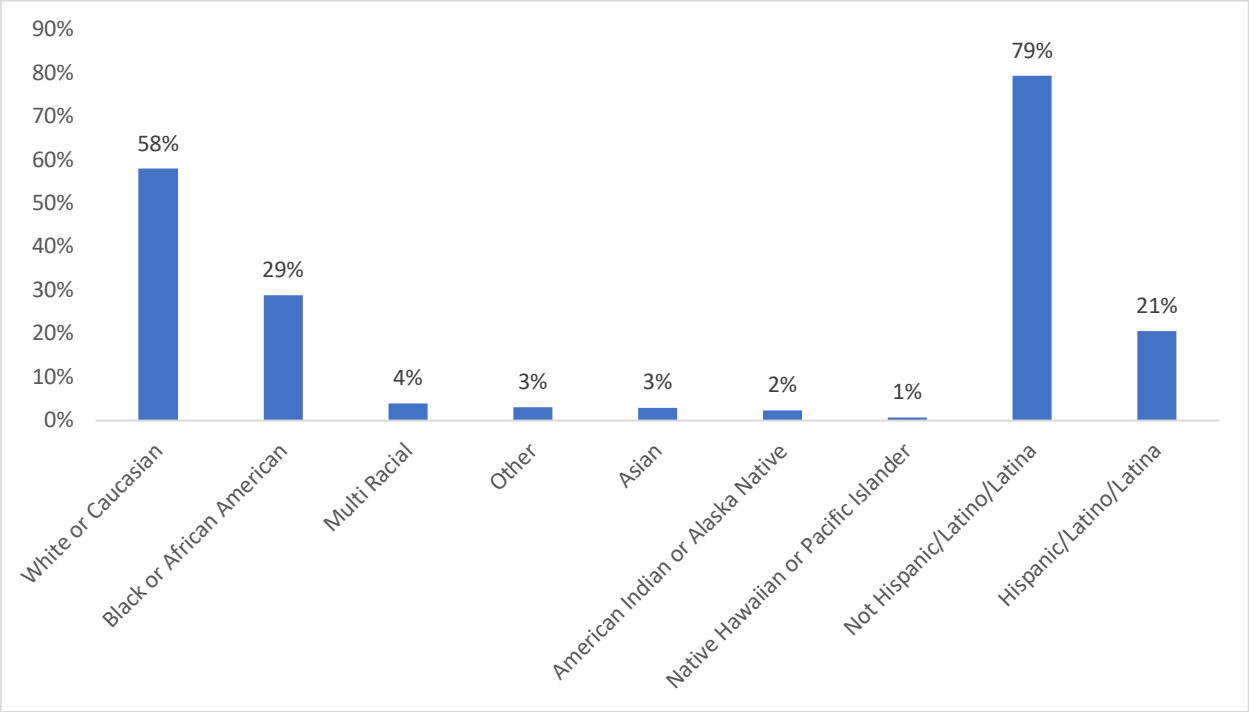
In the final dataset, there were 8,833 youth across 65 organizations. The most complete variable was gender, available for 8,825 youth, or 99% of youth in the final dataset. Race was available for 58% of youth in the final dataset, while ethnicity (46%) and free/reduced lunch eligibility (45%) were available for less than half of the youth in the final dataset.

Variable-level analysis

Underrepresented races and ethnicities. One goal of the *FIRST* LEGO League Jr. Season Pass Model was to reach youth from races and ethnicities traditionally underrepresented in STEM, including African American/Black youth, Hispanic/Latino(a) youth, American Indian or Alaska native youth, Hawaiian or Pacific Islander youth, or youth from other racial/ethnic backgrounds not including White/Caucasian or Asian. In the registration data, information on student's race was available for 5,110 youth across 44 organizations, and information on ethnicity for 4,065 youth across 35 organizations.

Broken down by racial/ethnic category, more than half of youth for whom we have race information were White/Caucasian (58%), 29% were Black/African American, 4% were multiracial, 3% were Asian, 2% were American Indian or Alaska native, 1% were Native Hawaiian or Pacific Islander, and 3% were Other. Of the 4,065 youth for whom we have ethnicity information, 21% were Hispanic/Latino/a and 79% were Not Hispanic/Latino/a. Figure 2 below shows the breakdown for all race/ethnicity categories. Note that most youth with "Other" for race were Hispanic/Latino/Latina (87%).

Figure 2. Race and Ethnicity Breakdown of Season Pass Youth in the 2017-2018 Season



We also looked at the aggregated race and ethnicity information to find the percentage of youth from racial/ethnic groups traditionally underrepresented in STEM. For this analysis, we considered White/Caucasian and Asian youth to be well-represented in STEM, and all other racial and ethnic groups to be underrepresented in STEM. Cases were omitted if they were missing *both* race and ethnicity information. Regardless of race, if ethnicity was reported as Hispanic/Latino/Latina, the student was considered underrepresented. Likewise, regardless of ethnicity, if the student was a race other than White/Caucasian or Asian, they were considered underrepresented. Of the 5,057 youth across 43 organizations for whom we have race and/or ethnicity information, 51% were from racial or ethnic groups underrepresented in STEM.

Underserved

Another goal of the *FIRST* LEGO League Jr. Season Pass Model was to reach youth typically underserved in STEM, including females and youth from low-income families. For these analyses, eligibility for free/reduced lunch is used as a proxy for low-income status. The

registration dataset included gender data for 8,825 youth from 65 organizations, 51% of whom were female. Looking at the site level, 83% of sites (54/65) met the criterion that 40% of youth be female. Of the 3,997 youth across 33 organizations for whom have free/reduced lunch eligibility data, 74% were Eligible for Free/Reduced Lunch.

Aggregate Underrepresented/Underserved Analysis

Next, we aggregated youth's registration information to create a single Underrepresented/Underserved (UR/US) variable, incorporating gender, free/reduced lunch eligibility, and race/ethnicity. If a student met the above-listed criteria for "underrepresented" or "underserved" for any single variable, they were considered to be UR/US.

Analytical Approach. Because there was so much missing data, it was difficult to definitively determine for all youth whether or not they were UR/US. Anyone with complete data was easy to classify as UR/US or not based on whether they met the established criteria. Youth with partial data, however, were more challenging to address. A student could be *confirmed* as UR/US with just one variable if they met the criterion for that variable: for example, a student with only gender data could easily be identified as UR/US if she was female. On the other hand, it is impossible to *confirm* a student is *not* UR/US unless we have data for all variables. For example, if we only have gender data and know the student is male, the male student could be a White male from a middle or high income family, or he could be a Black/African American male, a low-income male, and so on. In other words, without complete data, it was impossible to definitively determine a student was *not* underserved or underrepresented. This created a challenge for how to handle the cases with missing data.

Because of these challenges in classifying youth using incomplete data, we conducted overall UR/US analyses in three ways: complete data analysis, partial data analysis - upper estimate, and partial data analysis - lower estimate. Each of these approaches is described below.

Complete data analysis. For this analysis, only youth with complete data across gender, free/reduced lunch eligibility, and race/ethnicity are included. The benefit of this analysis is that we are not making any assumptions about the missing data. However, the drawback of this analysis is that it greatly reduced the sample size. Taking this approach, there was complete data for 3,242 youth across 27 organizations. Of these youth, 91% were Underrepresented and/or Underserved overall. At the site level, 18 of 27 sites (67%) met the criterion that 90% of youth be UR/US.

Partial data analysis - Upper Estimate. Because complete data were available for relatively few youth, the next step was to include youth with partial data, including an upper estimate of youth UR/US, and a lower estimate. For the upper estimate, youth with confirmation of UR/US features were included in calculations in addition to those with complete data. For example, a female would be included even if we did not know her race/ethnicity or free/reduced lunch eligibility. In contrast, only *confirmed* cases of *not* underrepresented (i.e., those with complete data) were included, and *unconfirmed* cases were excluded because not all information was available. For example, if we knew a student was Male and White/Caucasian, but we did not know if he was eligible for free/reduced lunch, he would have been excluded from this analysis. A White/Caucasian male who is not eligible for free/reduced lunch would be included, because he had complete data. In sum, this analysis is an upper estimate of UR/US percentages because <u>all</u> cases of UR/US are included, but *not* underrepresented cases are only included if they have

complete data. This means we likely underestimate the percentage of *not* UR/US, consequently overestimating the percent UR/US.

Including US/UR cases with partial data increased the sample size to 6027 youth across 54 organizations. Of these 6,027 youth, 95% were underrepresented or underserved. At the site level, 81% of sites (44 of 54) met the criterion that 90% of youth be UR/US.

Partial Data Analysis - Lower Estimate. The lower estimate classified all youth according to the information available, including classifying youth as *not* UR/US if the data we had failed to meet the criteria for UR/US. Here, if we knew a student was Male but didn't have information on his race/ethnicity or free/reduced lunch eligibility, he would still be included in the analysis and classified as *not* UR/US. Thus, this approach is a lower estimate of UR/US because it includes *unconfirmed* cases of *not* UR/US. This means we likely overestimate the percent *not* UR/US, and consequently underestimate the percent UR/US. However, it includes the most youth of the three approaches.

With the conditions for a lower estimate, we had information on 8,831 youth across 65 organizations. Of these youth, 72% were UR/US. This means that, at a minimum, 72% of youth across sites were UR or US. Using this approach, just 23 of 65 sites (35%) met the criterion that 90% of youth be underrepresented.

Table 1. Overall Percentage of Underrepresented/Underserved, Calculated in Three Ways

Sample	Description	n	%UR/US Overall	Sites meeting 90% UR/US
Complete Data	Only youth with data for all variables included	3,242	91%	18 of 27 sites (67%)
Partial Data - Upper Estimate[a]	Youth with confirmed US/UR (partial or complete data) PLUS confirmed not US/UR (complete data)	6,027	95%	44 of 54 sites (81%)
Partial Data - Lower Estimate[b]	All youth included. Unconfirmed UR/US considered to not be US/UR	8,831	72%	23 of 65 sites (35%)

a. This approach likely overestimates the % UR/US
b. This approach likely underestimates the % UR/US

Site-level Analysis

Next, we analyzed the percentage of reporting sites that met both criteria in the grant: 90% UR/US, and 40% female. Depending on how we estimate the percentage UR/US, the percentage of sites meeting both criteria ranged from 35% to 81%.

Table 2. Summarizing the Percentage of Sites Meeting Grant Criteria for UR/US and Gender, Calculated in Three Ways

Sample	Number of Sites	% Meeting 40% Female	% Meeting 90% UR/US	% Meeting Both Criteria
Complete Data	27	96%	67%	63%
Partial Data - Upper Estimate[a]	54	98%	81%	80%
Partial Data - Lower Estimate[b]	65	83%	35%	34%

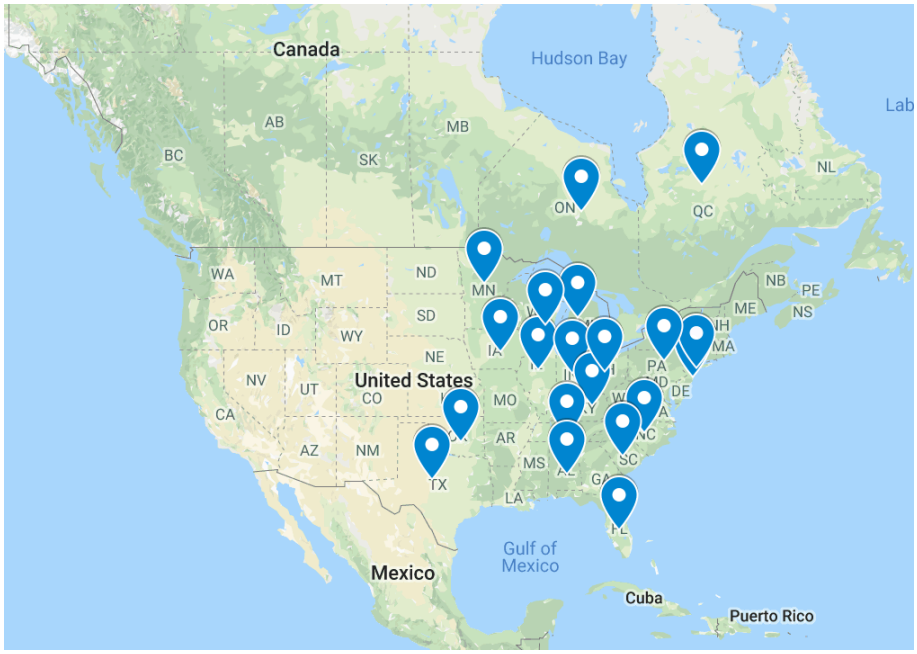
a. This approach likely overestimates the % UR/US
b. This approach likely underestimates the % UR/US

Survey Analyses

Site Administrator Participants

Out of 111 total sites, a total of 49 individuals completed the site administrator survey (44% response rate). Site administrators represented sites from 18 states and 2 Canadian provinces, but they were not uniformly distributed geographically. The response patterns do seem to reflect the location of sites, however, as only 13 of the 111 sites were in the western half of the United States or Canada.

Figure 3. Map of Site Administrator Survey Responses



Most site administrators had either a master's degree (45%) or bachelor's degree (36%), with 10% having a doctorate and 2% having a high school diploma or GED. Their fields of study were usually related to education (73%), though 10% had a STEM degree. Most were female (83%) and White/Caucasian (88%), with 12% Black/African American, and 2% Middle Eastern/North African. (5% declined to state.) While some had experience working in a STEM field either currently (7%) or in the past (17%), the majority (76%) had never worked in a STEM occupation. When asked to report their current occupation, 87% were currently working in education, whether as a teacher, program director/coordinator, or some other educational position (e.g., STEM Lab Instructor and Technology Integration Specialist, Curriculum Specialist).

When asked how they had become involved with *FIRST* LEGO League Jr. this season, most site administrators fell into one of three categories: either they first heard of the program through the grant opportunity (29%), through word of mouth (29%), or through previous experience with *FIRST* (29%). Those who first heard about the program through word of mouth,

found out from colleagues, supervisors, or were contacted by *FIRST*. In five cases (12%), site administrators were told specifically about the Season Pass model. Those with previous experience with other *FIRST* programs had experience leading either other *FIRST* LEGO League Jr. teams using the standard, non-Season Pass model, or they had experience with *FIRST* LEGO League or *FIRST* Robotics Competition. A smaller group (12%) had a pre-existing interest in STEM or STEM education and learned about *FIRST* LEGO League Jr. through their own research.

Site administrators were qualified for their roles in a variety of ways. Nearly all site administrators had previous experience working in education (79%), with youth (64%) and/or with diverse youth (64%). Others had previous experience with a *FIRST* program (62%), a background in STEM (48%), and/or experience in business, administration, or program management (24%). When asked which background experience was most useful, over half (57%) reported that their education background was most useful, while 20% thought their prior experience with *FIRST* was most useful, and 17% thought their prior experience with program administration/management was most useful. A few respondents thought their STEM background or experience working with youth was most useful.

Coach Participants

According to site administrator responses, a total of 339 adults coached at the 42 sites. The average number of coaches per site was 8.07 (*SD* = 9.48; min = 1, max = 50). Extrapolating from these 42 sites, we can estimate that 895.77 coaches participated across the 111 sites. Many sites had just one (19%) or two (12%) coaches, while 21% of sites had 3-5 coaches, 24% had 6-10 coaches, 12% had 11-15 coaches, and 12% had 16 or more coaches. Most coaches were teachers at the host site (83%), but others were administrators at the host site (21%), parents

(21%), high school students (12%), teachers from another nearby school (10%), other community members or volunteers (10%), or college students (7%). Site administrators estimated that 93% of coaches had experience working with youth, 78% had a background in education, and 59% had a background in STEM.

*Coach Survey Respondents***.** As for the Coach Survey, a total of 23 individuals from 9 of the 12 case sites completed the survey. Based on estimates extrapolating from the site administrator survey responses, these 23 coaches represented 2% of all coaches. The majority of respondents (73%; n=16) were teachers or student teachers at the site, while 2 were parents, 2 were volunteers, 1 was a student, and 1 was an "affiliated partner."

There was a wide range of educational backgrounds, with 50% having a high school diploma or less (many of whom were current college students), 25% having an associate's or bachelor's degree, and 25% having a Master's degree. Those with an associate's degree or higher typically held their degree in Education (85%), with the remaining 15% holding a STEM degree (i.e., computer science, technology). Most respondents (86%) were female, and the majority were White/Caucasian (87%), with 9% identifying as Black/African American, and 4% as Hispanic/Latinx. Most (78%) had a background in education, 57% had experience working with diverse student populations, 48% had a background in STEM, and 44% had a background working with youth. A few (13%) were *FIRST* alumni, having previously participated as a student. Just 20% had ever been employed in a STEM-related career; 80% had never held a position in the field.

Most coaches became involved with *FIRST* LEGO League Jr. because it was part of their school or youth organization (50%) and because it was a way to get youth interested in STEM (46%). Others cited a personal interest in *FIRST* LEGO League Jr. as a major motivation in their

decision to coach (38%). Just a quarter (25%) reported that they became involved in *FIRST* LEGO League Jr. because it was part of their job. In terms of their familiarity with *FIRST* in general, coaches were highly familiar with *FIRST* LEGO League, with 82% having heard of the program. Coaches were more mixed in whether they had heard of *FIRST* Tech Challenge (50%) or *FIRST* Robotics Competition (50%).

Coaches reporting overseeing an average of 27 youth (*SD* = 22.34), ranging from 4 to 98. More than half of coaches (61%) reported having another coach or other helper assist them in coaching their youth, while 39% of coaches worked alone.

Youth Participants

According to site administrators, a reported total of 6,230 youth enrolled in *FIRST* LEGO League Jr. across the 42 sites, and a reported 5,768 attended regularly. The average number of youth per site was 148.33 (*SD* = 140.64). Extrapolating from these 42 sites, we can estimate that 16,465 youth participated across the 111 sites (or 14,563 total based on the registration data). Site administrators reported equal representation of boys (*mean* = 76.18) and girls (*mean* = 73.61). Overall, youth were most commonly in 2nd or 3rd grade, and least commonly in kindergarten or 5th grade.

Table 3. Site Administrator Reports of Participating Youth by Grade

	Total	Average	Range
Kindergarten	204	12.75	0 - 146
1st grade	211	11.11	0 - 47
2nd grade	1,626	54.20	0 - 453
3rd grade	1,524	49.16	4 - 175
4th grade	910	30.33	0 - 149
5th grade	207	11.50	0 - 41

Youth were most frequently enrolled as part of an ongoing class or program (49%). Otherwise, recruitment strategies included an announcement (47%), personal invitations (31%),

information sessions (18%), or parent communications (14%). When asked how successful they were at recruiting youth traditionally UR/US in STEM, 79% of site administrators reported they were very successful, while 5% reported they were somewhat successful, and 7% unsuccessful. In describing their experiences recruiting youth traditionally UR/US in STEM, most site administrators responded that it was either easy or difficult depending on which dimension was being considered. Many sites reported that the demographic makeup of their youth mirrored that of the school they were working with. In some cases, site administrators reported that the school they partnered with was primarily Black/African American or Hispanic/Latino/Latina, so there were no issues recruiting youth who were traditionally UR/US. On the other hand, some sites reported that its schools focused on recruiting girls and economically disadvantaged youth. A few specific challenges mentioned in reaching UR/US populations included transportation issues and parent buy-in. A few strategies that worked well were partnering with organizations that had established relationships with youth traditionally UR/US in STEM, and providing additional supported when needed (i.e., ELL resources, additional sessions to reach a wider audience).

To gain a sense of the participation of youth from backgrounds historically underrepresented or underserved in STEM, the survey also asked site administrators to estimate the approximate percentage of youth representing an array of demographic characteristics. Table 4 below summarizes their responses. Overall, the program served high numbers of youth who qualified for free/reduced lunch (*mean* = 66%), and just over half (*mean* = 52%) of youth were races/ethnicities other than White/Caucasian or Asian.

Table 4. Youth Demographics as Estimated by Site Administrators

	0%	1-24%	25-49%	50-74%	75-100%	Mean
Qualify for Free/Reduced Lunch	5%	8%	16%	24%	**47%**	66%
Youth with Disabilities	10%	**63%**	23%	0%	3%	17%
English Language Learners	21%	**59%**	14%	3%	3%	16%
White/Caucasian	11%	19%	19%	22%	**28%**	43%
Black/African American	8%	**42%**	33%	3%	14%	28%
Hispanic/Latinx	6%	**69%**	22%	0%	3%	16%
Other Races/Ethnicities	17%	**67%**	17%	0%	0%	11%

*Youth Survey Respondents***.** A total of 426 youth from 10 of the 12 case sites completed the youth survey. Based on estimates extrapolating from the site administrator survey, these youth represented 3% of participating youth across the 111 sites. The number of surveys received from sites ranged from 9 to 109 (*mean* = 42.6, *median* = 38). Across the 10 sites, youth were, on average, 7.97 years old (*SD* = 1.44; *range* 5-12; *mode* = 9 [27%]). Most respondents were in the 3rd grade (31%) or 1st grade (22%). In terms of gender, 51% of participants identified as girls, 42% identified as boys, and 7% responded that they didn't want to report their gender.

Table 5. Youth Survey Responses by Gender and Grade

	# Boy	# Girl	# Don't want to say	Total #	Total%
Kindergarten	19	19	0	38	9%
1st grade	38	44	9	91	22%
2nd grade	31	38	10	79	19%
3rd grade	46	77	5	128	31%
4th grade	37	30	5	72	17%
5th grade	3	6	0	9	2%
Total #	174	214	29	417	
Total %	42%	51%	7%		100%

Of the 419 respondents to the question "What languages do you speak at home?," most spoke English (81%) or Spanish (24%) at home. Other languages respondents reported speaking at home were often spoken by only 1 or 2 youth. One exception is Creole, which 14 youth (3%) reported speaking at home. Most youth reported only speaking English at home (68%).

Parent Survey Respondents

A total of 88 individuals from 9 of the 12 case sites completed the parent survey. The majority of respondents (70%) were the mother or female guardian, while 28% were fathers or male guardians, and 2% identified as "other." (Henceforth, respondents will be referred to as "parents.") The majority of parents (82%) reported that they had one child participating in *FIRST* LEGO League Jr., but 16% had two children, and 2% had three or four children participating. Parents with multiple children participating were asked to focus on a single child for the survey.

Parents reported that their children were, on average, 7.83 (*SD* = 1.24) years old. Most youth were in first grade (39%) or third grade (30%). Just over half of youth (54%) were female, and two fifths (40%) of youth were White, followed by 26% who were Hispanic/Latinx. About 11% of youth were being educated under an individualized education plan (IEP) or behavioral intervention plan (504 plan). Table 6 below summarizes the characteristics of children whose parents completed the parent survey.

Table 6. Child Characteristics of Parent Survey Respondents

Grade		**Race/Ethnicity**		**Gender**		**IEP/504 Plan**	
K	2%	White/Caucasian	40%	Female	54%	Yes	11%
1st	39%	Hispanic/Latinx	26%	Male	46%	No	80%
2nd	16%	Black/African American	12%			Not Sure	7%
3rd	30%	Asian	11%			Prefer not to answer	2%
4th	10%	Other races/ethnicities	9%				
5th	2%						

Parents were asked about their family demographics and background. Overall, parent respondents were a diverse group in terms of education, income, and race/ethnicity. The highest degree in the household (whether the responding parent's or the child's other guardian's, if applicable) was less than a Bachelor's degree in 37% of responses, a Bachelor's for 30% of responses, or more than a Bachelor's for 33% of responses. Respondents were asked to report their annual household incomes in $10,000 increments. A third of parents skipped this question or selected "Decline to answer." Of the available responses, 32% of families had incomes of less than $50,000 per year, 19% had incomes from $50,000-99,999, 27% had incomes from $100,000-149,000, and 22% had incomes greater than $150,000. Parents most often identified as White (40%) or Hispanic/Latinx (30%), while 12% identified as Black/African American, 12% as Asian, and 4% as another race or ethnicity. Over half (62%) of families speak exclusively English at home, while 24% are bilingual households with English plus another language (primarily Spanish), and 14% were non-English household (again, primarily Spanish).

STEM Environments. About a third of parents (32%) reported being employed in STEM either currently (20%) or formerly (12%). Beyond employment, parents were also asked to report their own comfort and experience with STEM subjects. Parents responded to five items indicating their attitudes across a five-point Likert scale, from 1=Very Untrue to 5=Very True. Parents responded most positively to "I encourage my child to do well in science and math" (86% "Very true") and less positively to "People in our family are interested in science and math" (61% "Very True"), but overall responses indicated highly positive attitudes and high comfort towards STEM.

Table 7. Parental Comfort with Science and Math

Item	1 Very Untrue	2 Slightly Untrue	3 Neither	4 Slightly True	5 Very True	Mean (SD)
I encourage my child to do well in science and math	1%	6%	0%	7%	86%	4.71 (0.83)
I would like my child to learn about science and math jobs and careers.	2%	1%	5%	15%	77%	4.63 (0.82)
I am comfortable helping my child with homework in science and math	2%	7%	3%	12%	76%	4.52 (1.01)
I talk to my child about what he/she learns in science and math classes.	3%	2%	9%	21%	64%	4.40 (0.99)
People in our family are interested in science and math.	6%	1%	11%	22%	61%	4.30(1.10)

Parents were asked to report on other STEM activities their child has participated in other than *FIRST* LEGO League Jr. Responses indicate that there may be few formal opportunities for youth to explore STEM, with just 17% of parents reporting that their child was in another STEM club or program. Most often, youth engage in informal STEM activities, such as playing computer games, watching television shows, or reading books. A little less than half of parents (42%) reported that their child had engaged in 1-3 of these activities, 39% reported they had engaged in 4 or more activities. One fifth (20%) reported their child had never engaged in any of the following activities.

Table 8. Other STEM Activities

Other than *FIRST* LEGO League Jr., has your child participated in any of the following?	% yes
Playing computer games with a STEM focus *(e.g., games about building things, math, puzzles)*	49%
Watching STEM television shows *(e.g., Magic School Bus, SciGirls, Sid the Science Kid)*	42%
Reading STEM books for fun *(e.g., books about space, dinosaurs, bugs)*	42%
Visiting STEM museums *(e.g., science museum, museum of natural history)*	40%
Experimenting with STEM kits at home *(e.g., chemistry kits, Rock candy kits)*	32%
STEM clubs/afterschool programs *(e.g., robotics club, science fairs, math team)*	17%
Other (please specify)	3%

Summary of Evaluation Question 1 Results

Approximately 15,500 youth participated in *FIRST* LEGO League Jr. during 2017-2018, and approximately 22,000 youth participated across two years (from 2016-2018). This is roughly 50% of the stated goal to reach 45,000 youth across three years, not including participants from international sites. Most youth enrolled in *FIRST* LEGO League Jr. were White/Caucasian, about 20% were Hispanic/Latinx and about 25% were Black/African American. About half of youth were female, and nearly three quarters were eligible for free/reduced lunch. Sites were generally very successful in meeting the grant criteria for at least 40% female youth, but were less successful in terms of meeting the grant criteria for at least 90% UR/US youth. Depending on how estimates were calculated, between 30-80% of sites met grant criteria for both females and UR/US youth. Geographically, sites were not evenly distributed across the United States, with a much higher concentration in the eastern half of the country, a smaller representation of sites from the West/Pacific Northwest, and no sites in the Mountain States/Mountain West.

Site administrators were well prepared for their role. Most had a degree and/or previous experience in education, experience working with diverse youth, and a background in STEM. Approximately 900 coaches participated across the 111 sites in 2017-2018, with a wide range of the number of coaches per site (1-16). Most coaches were teachers at the host sites and had experience working with diverse student populations and in STEM. Coaches who were not teachers were often high school or college students. Neither site administrators nor coaches were very diverse, with most being female and White/Caucasian. A total of 88 individuals completed the parent survey. Households were very diverse in terms of education, income, race/ethnicity, language, and professional experience. Parents had highly positive attitudes towards STEM, but reported few opportunities for their child to explore STEM outside of class.

Results – Evaluation Question 2: Implementation

This next section focuses on Evaluation Question 2: *How did sites implement the Season Pass Model? What challenges did coaches and site administrators face? Is the Season Pass Model implemented with fidelity across sites?* To explore this question, we drew data from site administrator surveys, coach surveys, and observations.

Site Organization

Most sites (85%) reported hosting the small season pass, while 7% of sites reported hosting the large season pass, and 9% were not sure. Overall, 64% of sites reported holding *FIRST* LEGO League Jr. activities in a single location or building, while 36% reported using multiple locations or buildings. Of those using multiple buildings/locations, the number of locations ranged from 2 - 15, with an average of 5.57 (*SD* = 3.18). Nearly all sites hosted

activities in one or more schools (98%). Teams typically met during the school day (70%) and/or after school (43%). Some met on weekends (23%), and a few met before school (4%).

Though some sites had individual teams as small as 2 and as large as 9 youth, the average team size per site ranged from 4 to 7.5 youth, with a mean of 5.46 (*SD* = 0.92). Interestingly, when asked for the ideal team size, many site administrators reported preferences for smaller teams, saying "six was too large" and/or "four/three would be better." At the majority (91%) of sites, youth worked with the same team over the course of the season, but 9% of sites reported that youth worked with different teams each session.

Sites varied widely in their schedules for implementing *FIRST* LEGO League Jr. activities. On average, they hosted 1.57 meetings per week (*SD* = 1.04), for 70.21 minutes (*SD* = 29.92), over 11.91 weeks (*SD* = 5.11). The majority of programs (61%) met once a week, with meeting frequency ranging from every other week (2%) to every day (4%). The most common meeting length was 1 hour (38%), with 30% of teams meeting less than an hour, and 32% of teams met for more than an hour. When asked for the ideal schedule, responses were quite similar to the actual schedules reported. On average, their ideal schedule would be about 1.63 meetings per week (*SD* = 0.52), for 68.4 minutes (*SD* = 23.4 minutes).

Coach responses were very consistent with site administrator responses. According to coaches, the most common team meeting schedule was to meet once a week (74%), with 17% of teams meeting about twice a week, and 8% of teams meeting 3-4 times per week. Most commonly, teams met for 1 hour (61%), though 27% of teams met for 1.5-2 hours, and 9% of teams met for just 45 minutes.

Compared to last year, when only 42% of site administrators reported running more than one implementation cycle, more sites (72%) were able to complete multiple implementation

cycles this year, with an additional 4% still planning to implement a second one at the time of the survey. The majority of sites implementing more than one cycle ran two cycles (92%), though one site ran as many as 7 cycles. Those who did not implement more than one implementation cycle neglected to respond to a follow-up question to explain, but responses from the 2016-2017 season typically related to running out of time.

Sites also varied in whether they believed they could fund a *FIRST* LEGO League Jr. Season Pass in a typical year. Less than a fifth (17%) of site administrators reported they should be able to fund the program without support from *FIRST* LEGO League Jr., and 27% thought it was possible, but were not sure. There were significantly more sites who said they probably could *not* afford to fund *FIRST* LEGO League Jr. this year (51%) compared to last year (36%), perhaps reflecting the increase in price between seasons. Comments on this question frequently indicated that sites would need to secure funding through alternative grant opportunities, as there was no room in their standard budget for this expense. A few sites mentioned that without support from *FIRST* LEGO League Jr., some schools within their site would be able to afford the program, but not all. Nevertheless, when asked whether their sites would continue to host *FIRST* LEGO League Jr., 93% of site administrators answered "Yes," with just three site administrators answering "No" or "Not sure." Of the site administrators who were not certain they would host the program again, one explained that they would if they had funding.

Team Organization

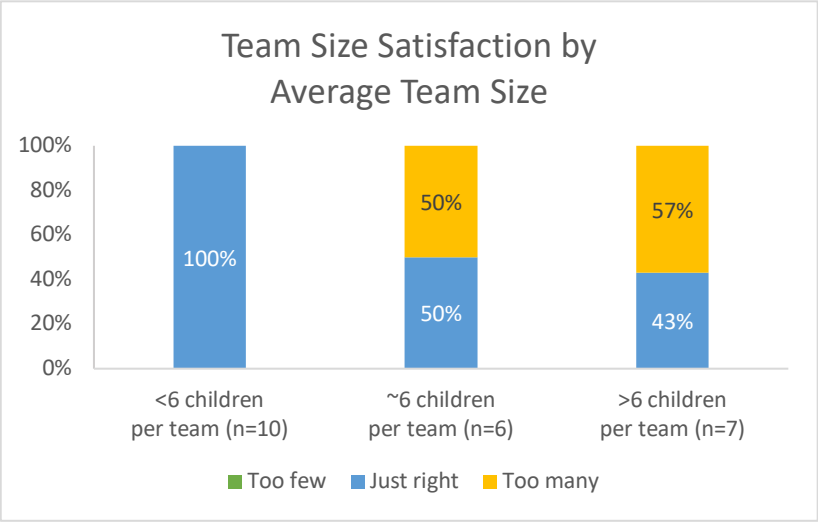
Almost two-thirds of site administrators (64%) reported that coaches oversaw multiple teams at once, while 31% of sites had just one team per coach, and 5% had some other arrangement, typically that the number of teams per coach varied by location within the site or varied between implementation cycles. For those coaches overseeing multiple teams, the average

was 3.75 teams (*SD* = 1.27), ranging from 1.5 to 6 teams per coach, according to site administrators.

As reported in the Coach Survey, there was a wider range and higher average in the number of teams each coach was overseeing. While 13% of coaches worked with only one team, other coaches worked with anywhere from 2 teams (9%) to 19 teams (4%). Most often, coaches worked with 4 teams (i.e., the mode; 39%), or an average of 6.25 teams (i.e., the mean; *SD* = 6.25). Teams also ranged in size. Coaches were asked to report the smallest team at their site (*mean* = 5.20, *SD* = 1.36; range 3-8), the largest team at their site (*mean* = 6.71, *SD* = 1.71; range 4-10), and the average team size at their site (*mean* = 5.86*, SD* = 1.90; range 3-8). Overall, 27% of coaches reported that the average team at their site was 6 youth, the *FIRST* LEGO League Jr. recommended size; 41% of coaches reported that the average team was smaller than 6, and 32% reported that the average team was larger than 6.

Coaches were also asked to reflect on whether the number of youth on each team was too few, too many, or just right. As shown in Figure 4, responses varied depending on the size of the average team at a site. Specifically, coaches following the recommended team size (6 youth per team) were equally split between thinking the teams were too big and thinking they were just right. On the other hand, coaches with smaller teams were in agreement that team sizes were just right.

Figure 4. Coach Satisfaction with Team Size



Coaches were asked to reflect on whether team size had a positive or negative impact on youth's experiences in the program and why. 100% of coaches with average teams smaller than 6 said that the small teams had a positive impact on youth's experiences, usually commenting on ease of keeping youth on task or of dividing labor without leaving anyone out. For example, one coach shared "If I had more than 5, it was tough to keep all students on task." Another shared "4 students was the perfect size. Everyone had a job."

Coaches with about 6 youth per team were mixed, with 50% thinking team size had a positive impact and 50% thinking it had a negative impact. Positive responses included "good student to teacher ratio" and "They needed to work together, but there weren't too many that it was hard for everyone to participate." Negative responses included "With too many team members it was difficult for them to agree and for all to contribute" and "I think teams should be smaller to allow for more hands-on time with the LEGO [bricks]."

Only 14% of coaches with teams larger than 6 thought team size had a positive impact. Coaches reported difficulty getting youth engaged ("some students were not interested in helping

because there were too many students to share with") or to stay on task ("larger groups were a bit more difficult to maintain").

FIRST LEGO League Jr. Program Delivery

Implementation of the 12 Sessions

Nearly all sites reported making connections with a broader curriculum. Most commonly, 84% of sites reported making connections with a science curriculum. Other common connections included language arts/reading (57%), social studies (37%), and mathematics (35%). Interestingly, sites were much more likely to make connections with social studies this year (37%) than last year (14%), likely reflecting the social implications of clean water availability.

Coaches reported spending, on average, less than 30 minutes per week (44%) or 30-60 minutes per week (44%) preparing for meetings, though 13% spent as much as 1-2 hours preparing. Only 61% of coaches reported having completed all 12 sessions laid out in the Team Meeting Guide, and many of these coaches reported having "combined" or "squeezed together" sessions to be able to complete them within their schedule constraints. Similarly, coaches who did not complete all 12 sessions reported inadequate time (26%), or that some sessions were developmentally or logistically inappropriate for their youth or setting (9%).

Expos

Despite these constraints, the majority of coaches reported that youth at their site attended an Expo on-site (74%) or at a different location (17%). Just 9% reported that youth had not participated in an Expo or culminating event. Parents were usually invited to these events (86%), and, on average, about half (54%) of parents attended, ranging from 20% (n=1) to 100% (n=1). According to the Site Administrator Survey, the majority of sites (77%) reported hosting an Expo(s) at their site, and an additional 2% were still planning to host one after the time of survey

completion. For those sites that had hosted an Expo, the number of Expos ranged from 1 to 7 (*mean* = 2.27, *SD* = 1.42), and 86% of these sites reported that "most youth" attended the Expo, while 14% reported "Some youth attended or will attend an Expo, but not a majority." An additional 31% of site administrators reported that some (11%) or most (20%) youth at their site attended an Expo elsewhere, off-site. These included citywide (3 sites), regional or county-wide (4 sites), and statewide Expos (1 site), plus the World Championships (1 site).

All coaches (100%) thought the Expo was a positive experience for youth. They reported that youth enjoyed presenting and sharing their ideas and models (62%), as well as seeing other teams' projects (24%). Coaches said youth seemed proud of themselves (29%), though one coach thought the youth were nervous. When asked what they would change about the Expos to make it a more positive experience for youth, the most common response was nothing (44%). Suggested recommendations included higher turnout of parents and the community (17%), and adjusting the length of the event, with some wanting to make it shorter (17%), and some make it longer (11%).

WeDo 2.0

A requirement of the *FIRST* LEGO League Jr. program this season was that all teams use the WeDo 2.0 kits. Compared to last year, when 85% of site administrators reported that all teams used the kits, there were even higher rates of uptake this year. Specifically, 95% of site administrators reported that 100% of teams used the kids, while the remaining 5% of sites reported that either 80% (2.5%) or 90% (2.5%) of teams used the kits. Unfortunately, similar to last year, a significant portion of sites reported issues using the WeDo 2.0 kits (41% this year, 47% last year). Site administrators described technological issues, such as difficulties with batteries, Bluetooth, or power supply (16%), as well as issues with insufficient materials, lost

pieces, or missing parts (10% of sites). However, it is worth reiterating that over half (62%) of sites were "very satisfied" with WeDo 2.0, so these challenges were not universal.

Parent Awareness and Participation

When asked if they had been aware that their child participated in *FIRST* LEGO League Jr., 61% of parents had been aware, but 39% of parents only became aware when asked to take the survey. Of those parents who were previously aware (n=54), 42% were not involved in *FIRST* LEGO League Jr. activities, 25% were slightly involved (i.e., attended one or two meetings, or attended Expo only), 8% were moderately involved (i.e., attended several meetings), and 26% were very involved (i.e., attended almost all meetings). About half (52%) of parents reported receiving some sort of materials about the program, such as emails, flyers, and information packets. Over half (60%) of parents reported attending the Expo. Most parents (85%) reported that their child attended most or all meetings, though 13% were not sure.

Observation Ratings

At each of the 12 site visits, the evaluation team rated the quality of the program across five dimensions. Overall, there was "Good" to "Excellent" quality of interactions among teammates and between coaches and youth, with less consistent quality of connections to STEM.

Table 9. Observer Ratings of Program Quality

	1 Very Poor	2 Poor	3 Mixed or Neutral	4 Good	5 Excellent	Mean (SD)
Quality of interactions among teammates	0%	0%	18%	45%	36%	4.18 (0.72)
Quality of interactions between coach and youth	0%	0%	9%	36%	55%	4.32 (0.72)
Quality of instruction/guidance by the coach	0%	0%	18%	64%	18%	3.91 (0.56)
Quality of connections to STEM	0%	27%	55%	0%	18%	3.05 (1.01)
Fidelity of implementation to the FLL Jr. program	0%	0%	18%	64%	18%	3.91 (0.56)

The variation in STEM connections was captured in the evaluators' notes. For example, below is a description of a site with rich STEM connections:

> *The opening full group lesson* [on simple machines and the engineering design process] *is fantastic. The teacher is an engineer and is excellent with kids. Even though his audience is K-5, he presents the material in a way that is engaging for the 5th graders without overwhelming the kindergartners. Once in small groups, the coaches do a good job asking questions to guide student thinking.*

On the other hand, notes from other sites indicated "*the coach did not make any explicit connections to STEM*" or "*little mention of STEM concepts. Mentioned germs/bacteria in quick comment after PlayPump video.*"

Materials

Coach Ratings

Coaches were asked to share how frequently they used various materials to guide or support their team meetings. There was a wide range in use of materials, even those required by the *FIRST* LEGO League Jr. program. For instance, 19% of coaches reported that they "never" or "rarely" used the Team Meeting Guide, and 14% "never" or "rarely" used the Engineering Notebooks. On the other hand, 68% of coaches "often" or "always" used these resources.

Table 10. Frequency of Use of Materials

	1 Never	2 Rarely	3 Sometimes	4 Often	5 Always	Mean (SD)
Engineering Notebook	5%	9%	18%	36%	32%	3.82 (1.14)
Team Meeting Guide	5%	14%	14%	46%	23%	3.68 (1.13)
Other instructional materials related to the challenge topic, AQUA ADVENTURE	10%	14%	29%	24%	24%	3.38 (1.28)
Other instructional materials related to STEM	9%	14%	24%	38%	14%	3.33 (1.20)
Materials from the Training*	17%	11%	17%	33%	22%	3.33 (1.41)

*17% of coaches reported that they did not know about any coach training or materials. The above percentages reflect the responses of the 83% of coaches who were aware.

Notwithstanding the range of frequency of use, most coaches believed the Team Meeting Guide was helpful in describing the sequence of activities, the steps within each activity, and the purpose and goals of the activity. There were mixed responses, however, on whether the Team Meeting Guide was helpful in pointing coaches toward additional resources that may be helpful, with 24% of coaches finding the Guide unhelpful in this area. In addition, there was variability in responses on the age-appropriateness of activities, with 19% of coaches saying the Team Meeting Guide was unhelpful in this regard. Follow-up analyses revealed that three of the four coaches who felt the activities may be age-inappropriate had some proportion of youth in kindergarten. However, the other eight coaches working with kindergartners did not have these concerns, indicating that it may depend on the particular group of youth and/or the coach's facility with scaffolding for this age group.

Table 11. Coach Ratings of Team Meeting Guide's Helpfulness

	1 Very Unhelpful	2 Unhelpful	3 Slightly Helpful	4 Very Helpful	Mean (SD)
Describing the sequence of activities over the entire season	0%	10%	24%	67%	3.57 (0.68)
Describing the steps within each activity	0%	10%	24%	67%	3.57 (0.68)
Describing the purpose and goals of each activity	0%	5%	38%	57%	3.52 (0.60)
Providing necessary background information	5%	5%	24%	67%	3.52 (0.81)
Providing additional resources (for the research or programming)	5%	19%	24%	52%	3.24 (0.94)
Providing age-appropriate activities	14%	5%	29%	52%	3.19 (1.08)

Relatedly, in debrief conversations with coaches after the evaluation team finished observations, it frequently came up that the Team Meeting Guide was insufficient preparation for some of the activities. For instance, several coaches chose to search online for videos of PlayPumps so that youth could see them in action. Other sites expressed a need for more support finding age-appropriate resources for the research component of the design process.

Observation Ratings of Materials

At each of the 12 site visits, the evaluation team rated the materials (e.g., LEGO kits, tablets, notebooks) across four dimensions: quality, availability, sufficiency, and range. Overall, materials were highly rated, particularly on quality, but received a lower rating on sufficiency.

Table 12. Observer Ratings of Materials

	1 Very Poor	2 Poor	3 Mixed or Neutral	4 Good	5 Excellent	Mean (SD)
Quality of materials (i.e., are they organized, labeled, well-maintained)	0%	0%	9%	0%	91%	4.88 (0.48)
Availability of materials (i.e., can kids easily access what they want/need?)	0%	9%	0%	9%	82%	4.63 (0.78)
Sufficiency of materials (i.e., are there enough materials to go around?)	0%	9%	9%	45%	36%	4.06 (0.83)
Range of materials (i.e., is there a variety of materials?)	0%	11%	0%	44%	44%	4.23 (0.89)

At one site, the insufficiency of materials was particularly salient. As documented in the notes:

> *This school was one sub-site of multiple within the [site's] Season Pass. Because of this, the site admin had to split up six kits across sub-sites. This school only received two kits originally, for three classes of five groups each. The teacher was able to convince her district to purchase an additional three kits, so each group within each class can have its own kit. However, at the end of each session, the groups must disassemble their designs so that the next class can build. The teacher takes photos to remind students of what they built last time.*

In the case described above, insufficient materials actually influenced fidelity, as students were not able to add on to their designs from the previous session.

Program Comparisons with 2016-2017

Two-thirds (65%) of site administrators reported that they had also participated in the 2016-2017 CREATURE CRAZE^{SM} season of *FIRST* LEGO League Jr. Site administrators were mixed in their feedback comparing the two seasons, as shown below in Table 13. Overall, they preferred the CREATURE CRAZE^{SM} challenge theme over AQUA ADVENTURE^{SM} challenge theme, but thought logistics and coordination went more smoothly the second year.

Table 13. Site Administrators' Year by Year Comparison of Program

	1 Much better last year	2 Slightly better last year	3 About the same	4 Slightly better this year	5 Much better this year	Mean (*SD*)
Challenge Theme	30%	20%	27%	10%	13%	2.57 (1.38)
Materials Provided (e.g., Coach's Guide, Site Administrator Guide)	13%	13%	55%	10%	10%	2.90 (1.08)
Overall organization and logistics (i.e., coordinating the program)	13%	3%	42%	29%	13%	3.26 (1.15)
Overall program/content delivery	26%	3%	36%	23%	13%	2.94 (1.37)

Comments suggested that many site administrators thought the AQUA ADVENTURE theme and materials were too advanced, particularly for their youngest youth.

> *I work in a site that only hosts kindergarten teams. Last year's materials were much better suited to that age group. This year, I had to make a lot of adjustments to the curriculum.*

> *The concepts were slightly more accessible last year for the students.*

> *The 2nd grade teachers thought this theme was harder for the students to understand and connect with, but students still enjoyed the program and learned.*

One site administrator pointed out the difficulty of the research component, specifically.

> *This year's theme was challenging to implement as there were few research resources that were available on a 2nd or 3rd grade reading level. This limited the student's ability to conduct their own research and gain their own understanding of the challenge.*

In contrast, other site administrators reported that the AQUA ADVENTURE theme was more relevant given recent local and national news events.

> *The content was very relevant to the student's lives. There have been many water incidents in the United States in which the students have heard.*

> *The [AQUA ADVENTURE] theme was much more relevant to the students this year. I was able to relate problems with the journey of water to local issues that have made news.*

In terms of implementation, organization, and logistics, 61% of site administrators reported that they had made minor changes, 26% had made major changes, and 13% had not made any notable changes this year compared to last year. Several respondents made changes to the program implementation in order to have more time for implementation, address the difficulty of the AQUA ADVENTURE theme for younger students, or create a richer experience for youth.

> *This year our program was started earlier in the school year. Last school year I felt as though we were rushing to complete the program. Therefore, I knew that I needed to start the [program] earlier in the school year.*
>
> *We added more supplemental learning activities such as water testing, aquifer demonstrations, a groundwater patch program, etc. to get the [students] more interested and invested in their research (water is less exciting than animals).*
>
> *The curriculum wasn't appropriate for kindergarteners this year both for the interest level and ability levels. I had to recreate the curriculum to meet our needs. So every week I sent out a coaches email outlining what to follow and what to ignore in the Engineering Notebook.*
>
> *Some of the changes I made were to take the time to incorporate the 'spontaneous' challenge builds that are in each session…This year, I did most of them and then allowed students to share about what they built/designed. It worked out better in creating a 'team-centered' environment.*

One to two respondents reported making improvements to their coach trainings, changing the location of their Expo, expanding the program, giving coaches more autonomy, partnering with different schools or organizations, or restructuring their internal team.

Participant Experiences

Site Administrator Experience

Site administrators were also asked to rate the level of challenge experienced in different aspects of being a site administrator, from 1 ("Not at all challenging") to 5 ("Extremely

Challenging"). Just as in the 2016-2017 season, the most challenging aspects were scheduling (*mean* = 2.83, *SD* = 1.17) and having sufficient resources (e.g., funding, materials, space; *mean* = 2.57, *SD* = 1.23). The least challenging aspects were recruiting youth (*mean* = 1.49, *SD* = 1.14) and handling coach attendance problems (*mean* = 1.49, *SD* = 0.90).

Table 14. Site Administrator Ratings of Challenges, From Most Challenging to Least Challenging

	1 Not at all challenging	2 Slightly challenging	3 Moderately challenging	4 Very challenging	5 Extremely challenging	Mean (SD)
Scheduling	18%	15%	45%	13%	10%	2.83 (1.17)
Having sufficient resources	19%	38%	19%	14%	10%	2.57 (1.23)
Coordinating multiple teams	26%	31%	26%	7%	10%	2.43 (1.23)
Managing the number of children on teams	22%	34%	29%	12%	2%	2.39 (1.05)
Recruiting coaches	37%	24%	18%	11%	11%	2.34 (1.36)
Following the program's guidelines or expectations	36%	31%	17%	14%	2%	2.17 (1.15)
Managing building or technical aspects of the activities	54%	15%	15%	10%	5%	1.97 (1.27)
Training coaches	48%	25%	18%	8%	3%	1.93 (1.10)
Managing youth behavior problems	49%	26%	21%	3%	3%	1.85 (1.01)
Recruiting traditionally underrepresented youth	56%	21%	10%	8%	5%	1.85 (1.20)
Handling youth attendance problems	62%	16%	16%	3%	3%	1.68 (1.03)

	1 Not at all challenging	2 Slightly challenging	3 Moderately challenging	4 Very challenging	5 Extremely challenging	Mean (SD)
Handling youth attrition/drop outs	68%	16%	11%	3%	3%	1.55 (0.98)
Handling coach attendance problems	70%	16%	11%	0%	3%	1.49 (0.90)
Recruiting youth	80%	8%	5%	0%	8%	1.49 (1.14)

N = 39

When asked how *FIRST* LEGO League Jr. could improve their experience as a site administrator, suggestions varied. Of the 29 respondents to this question, 6 did not have any recommendations. Of those who did provide suggestions, most commented on materials (mentioned by 8 site administrators) or training (mentioned by 6 site administrators). More specifically, site administrators requested additional Inspire Sets, materials in more languages, and more visual materials such as posters. One site administrator requested receiving materials earlier. Again, two site administrators brought up the issue of materials that didn't work well with younger audiences. In regards to training, site administrators requested additional training on working with the WeDo 2.0, preparing for Expos, and preparing for the program. They also requested more clarity around the training available to them and to their coaches,

> *This was year 2 as program administrator and I still had no idea about [online training] or where it is located. I am also unaware of any training for me. Where is this located because I would find it extremely helpful.*

These challenges notwithstanding, most (93%) site administrators planned to return as the site administrator next year, while just 2% did not plan to return and 5% were unsure. When those expecting to return were asked what they will do differently, they reported they will provide more supports to coaches (47%), adjust their approach to scheduling (31%), recruit more support (13%), and create smaller groups (13%). When those not returning or not sure were

asked why, one indicated they were retiring, one was unsure about funding, and one did not know their plans yet.

Coach Experience

Coaches were recruited through a variety of means. Over half (53%) of site administrators described processes for engaging internal staff as coaches. In some cases this meant recruiting teachers, while in other cases it meant recruiting staff from non-profit and other professional organizations. In cases where teachers were recruited, they were asked to volunteer, not assigned to participate. Other times, site administrators described embedding the program in regularly scheduled classes (29%), and therefore not needing to recruit coaches, since they were automatically involved. Site administrators also reached out to parents (16%), or held information events (11%). While many site administrators reported no issues with recruitment (often because they were working with teachers for a school-based program), several reported that they found some potential coaches to be hesitant to take on more work or to have conflicting obligations. A couple of site administrators also noted that it was hard to reach parents in order to ask them to be coaches, and some potential coaches were hesitant to participate because they lacked confidence in the coding element or with the breadth of the program in general. Nearly a quarter of site administrators (21%) reported issues with unexpected coach turnover. Explanations for these turnover problems included personal issues and changes of employment.

We asked coaches how prepared they felt to engage in various activities with their team members. They felt most prepared to connect activities with STEM content, which is interesting in comparison with observation data (see Table 9 on page 33), which found inconsistent quality of connections to STEM. On the other hand, coaches felt least prepared to work with youth with a range of prior knowledge. This is consistent with themes from other data sources, expressing

challenges with using the program materials with youth of different ages, and coaches' desire for additional resources for research or programming in the Team Meeting Guide .

Table 15. Coach Reports of Preparation, from Most Prepared to Least Prepared

	1 Very Unprepared	2 Unprepared	3 Neither	4 Prepared	5 Very Prepared	Mean (SD)
Connecting activities with STEM content	5%	0%	5%	62%	29%	4.10 (0.89)
Engaging students in the activities and content	10%	0%	10%	48%	33%	3.95 (1.16)
Leading youth through the different challenge components	5%	5%	10%	57%	24%	3.90 (1.00)
Supporting students with a range of prior knowledge	10%	10%	14%	38%	29%	3.67 (1.28)

Similarly, coaches were asked to report how challenging certain aspects of leading *FIRST* LEGO League Jr. activities were on a 1-3 scale (1=Not at all challenging, 3 = Very Challenging), as well as how successful they were at these aspects (1 = Not at all successful, 3 = Very successful). Coaches reported very limited challenges and high success in developing relationships with kids and supporting learning through play. On the other hand, they reported high challenges and low success with time management.

Table 16. Coach Perceptions of Challenges and Level of Success

	Challenging Mean (SD) (1=Not at all Challenging, 3 = Very Challenging)	**Successful Mean (SD)** (1=Not at all Successful, 3 = Very Successful)
Developing relationships with kids	1.42 (0.69)	2.68 (0.48)
Incorporating STEM content into instruction	1.79 (0.63)	2.53 (0.61)
Managing behavior	2.00 (0.49)	2.50 (0.51)
Getting kids excited about STEM	1.59 (0.71)	2.68 (0.48)
Supporting kids in learning through play	1.39 (0.61)	2.70 (0.47)
Time Management	2.50 (0.62)	2.06 (0.54)
Following the program's guidelines or expectations	2.16 (0.50)	2.21 (0.42)
Scaffolding instruction for children at different ages/stages of development	2.11 (0.66)	2.21 (0.54)
Maintaining interest or engagement of youth	1.94 (0.54)	2.55 (0.51)

Note: Green cells indicate the least challenging and most successful items. Red cells indicate the most challenging and least successful items.

Next we asked about coaches' priorities for the program and participating youth. Specifically, we asked how important 16 goals were to the coaches, on a scale of 1-5 (1= Not at all important, 5 = Extremely important). Their responses are presented in Table 16. Overall, coaches thought it was most important that youth learn to work together as a team (*mean* = 4.91, *SD* = 0.29) and have fun (*mean* = 4.73, *SD* = 0.46). They thought it was least important that youth graduate to *FIRST* LEGO League (*mean* = 3.77, *SD* = 1.45) and learn about robotics (*mean* = 4.00, *SD* = 1.16).

Table 17. Coach Ratings of Importance of Goals for Team, from Most to Least Important

I want my team to…	1 Not at All Important	2 Slightly Important	3 Moderately Important	4 Very Important	5 Extremely Important	Mean (SD)
Learn how to work as a team	0%	0%	0%	9%	91%	4.91 (0.29)
Have fun	0%	0%	0%	27%	73%	4.73 (0.46)
Build something	0%	0%	5%	24%	71%	4.67 (0.58)
Be actively engaged during activities	0%	0%	5%	32%	64%	4.59 (0.59)
Be accessible to all students	5%	0%	5%	14%	77%	4.59 (0.96)
Communicate their thoughts and share ideas with their peers	0%	0%	9%	23%	68%	4.59 (0.67)
Learn about a real-world problem	0%	0%	0%	41%	59%	4.59 (0.50)
Learn about STEM	0%	0%	9%	32%	59%	4.50 (0.67)
Share or present what they've learned	0%	5%	9%	23%	64%	4.45 (0.86)
Be diverse and inclusive	5%	0%	9%	23%	64%	4.41 (1.01)
Relate new experiences to something familiar	0%	0%	18%	23%	59%	4.41 (0.80)
Enjoy tasks for their own sake	5%	0%	18%	9%	68%	4.36 (1.09)
Learn programming/coding	00%	5%	23%	18%	55%	4.23 (0.97)
Try out possibilities, revise hypotheses, and discover new questions	0%	9%	14%	23%	55%	4.23 (1.02)
Learn about robotics	5%	5%	23%	23%	46%	4.00 (1.16)
Graduate to *FIRST* LEGO League	9%	14%	18%	9%	50%	3.77 (1.45)

Per site administrator reports, coaches were successful in their roles, with slightly higher levels of success in establishing relationships with kids (*mean* = 4.40 out of 5) than in incorporating STEM content (*mean* = 4.00) or managing behavior (*mean* = 3.98 out of 5).

Table 18. Site Administrator Ratings of Coach Success, from Most to Least Successful

	1 Very Unsuccessful	2 Unsuccessful	3 Mix of successful and unsuccessful	4 Successful	5 Very Successful	Mean (*SD*)
Developing relationships with kids	7%	0%	2%	26%	64%	4.40 (1.08)
Getting kids excited about STEM	7%	0%	2%	38%	52%	4.29 (1.07)
Incorporating STEM content into instruction	7%	2%	10%	44%	37%	4.00 (1.12)
Managing behavior	8%	0%	20%	33%	40%	3.98 (1.14)

Site administrators selecting "N/A" were excluded itemwise from percentage calculations. n=42.

Summary of Evaluation Question 2 Results

Most sites hosted *FIRST* LEGO League Jr. with the small season pass, in a school, and during the school day. On average, sites hosted 1.5 meetings per week, for slightly over an hour, over approximately 12 weeks. The average team size was approximately 5 youth, ranging from 4 to 7.5. There was a consensus that having smaller teams had a positive impact on youth's experiences. Over half of site administrators described processes for engaging internal staff as coaches, either by asking for volunteers from the host site, or implementing the program in an established classroom, and therefore automatically engaging teachers in that classroom. The majority of coaches reported that youth at their site attended an Expo, either on-site or at a different location. Parents were usually invited to these events and about half attended. All coaches thought the Expo was a positive experience for youth.

In terms of implementation, organization, and logistics, over three quarters of site administrators reported that they had made changes this year compared to last year, often in

order to have more time for implementation or address the difficulty of the AQUA ADVENTURE theme for younger students. Despite efforts to make more time for implementation, only three-fifths of coaches reported having completed all 12 sessions in the Team Meeting Guide, and many of these coaches reported having "combined" sessions to be able to complete them within their schedule constraints. This is consistent with challenges reported by site administrators and coaches, who reported scheduling and time management as some of the most challenging aspects of running the program.

Site administrators and coaches reported primarily positive feelings towards *FIRST* LEGO League Jr. materials, with some reservations. Though there were occasional technical issues with the WeDo 2.0, most sites reported being satisfied with the WeDo 2.0 overall. In terms of other materials, most coaches believed the Team Meeting Guide was a helpful starting point for program implementation, but some thought the activities were too advanced for younger youth, and several searched for supplementary materials. Site administrators voiced similar feedback. Additionally, many site administrators reported challenges with having sufficient resources to serve all the youth at their site. These challenges notwithstanding, most site administrators planned to return as the site administrator next year.

Results – Evaluation Question 3: Outcomes

We now shift to address Evaluation Question 3: *Is there evidence of promise in increasing youth's positive attitudes toward STEM, Emerging Activation, learning through play, and 21st century skills? Is there variation in outcomes by subgroup? That is, do participants show equal levels of learning regardless of gender, age, or other variables?* Analyses for these questions drew from Youth Surveys, Parent Surveys, Coach Surveys, and Observations.

Activation/Emerging Activation

On the Youth Surveys, youth showed moderate increases in both Emerging Activation and Emerging 21st century skills. For Emerging Activation, youth showed statistically significant increase ($t(417) = 8.55$, $p < 0.01$) with a small to medium effect size (Cohen's $d = 0.39$). For Emerging 21st Century Skills, youth showed a statistically significant increase ($t(381) = 3.17$, $p < 0.01$) with a small effect size (Cohen's $d = 0.15$).

Table 19. Retrospective-pre and Post-test Scale Scores by Program

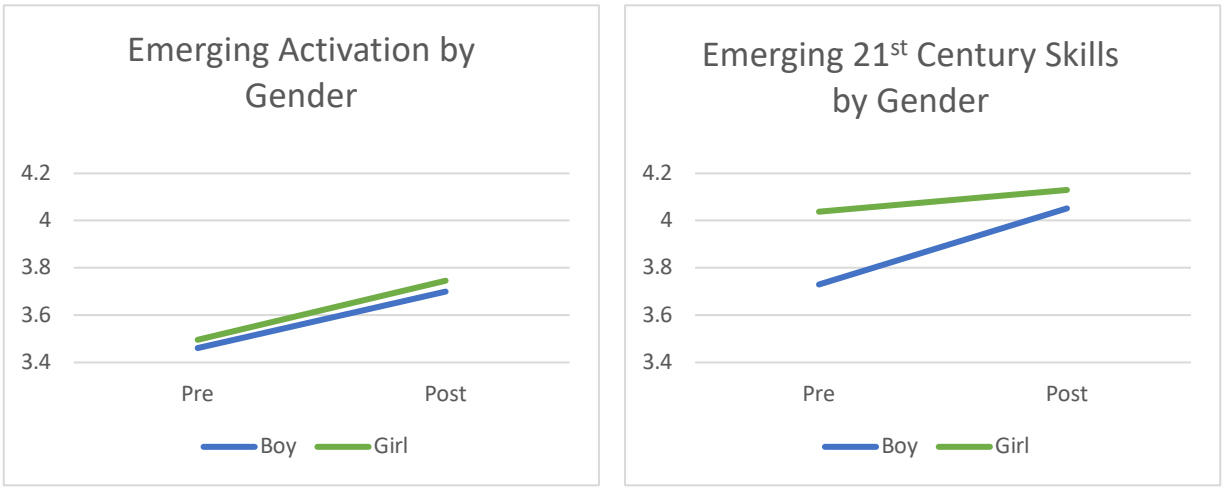
Construct	Average Scores (min=1, max=5)		Effect Size
	Pre Mean (SD)[b]	Post Mean (SD)[c]	Cohen's *d*[d]
Emerging Activation[a]	3.47 (0.65)	3.73 (0.67)	0.39
Emerging 21st Century Skills[b]	3.90 (1.30)	4.09 (1.17)	0.15

a. Emerging Activation ($n = 420$)
b. 21st Century Skills ($n = 382$)
c. SD = Standard deviation, a measure of the spread of responses. Higher standard deviations indicate more variability.
d. Statistician Jacob Cohen provided the following guidelines for interpreting Cohen's *d*: 0.2 = small effect size, 0.5 = medium effect size, and 0.8 = large effect size (Cohen, 1977).

Analyses by Gender

Girls consistently scored slightly higher than boys on Emerging Activation and Emerging 21st Century Skills, but the two groups both showed comparable gains over time. A series of repeated measures ANOVAs, comparing changes from retrospective pre-scores to post-scores for girls and boys (i.e., with time as the within-subject factor, and gender as the between subject factor) found no significant time by gender interaction for Emerging Activation (time x gender $F(1,385 = .04$, $p > 0.05$) or Emerging 21st Century Skills (time x gender $F(1,380 = 3.46$, $p > .05$). In other words, both boys and girls experienced comparable gains in Emerging Activation and Emerging 21st Century Skills.

Figure 5. Gender Changes in Emerging Activation and Emerging 21st Century Skills

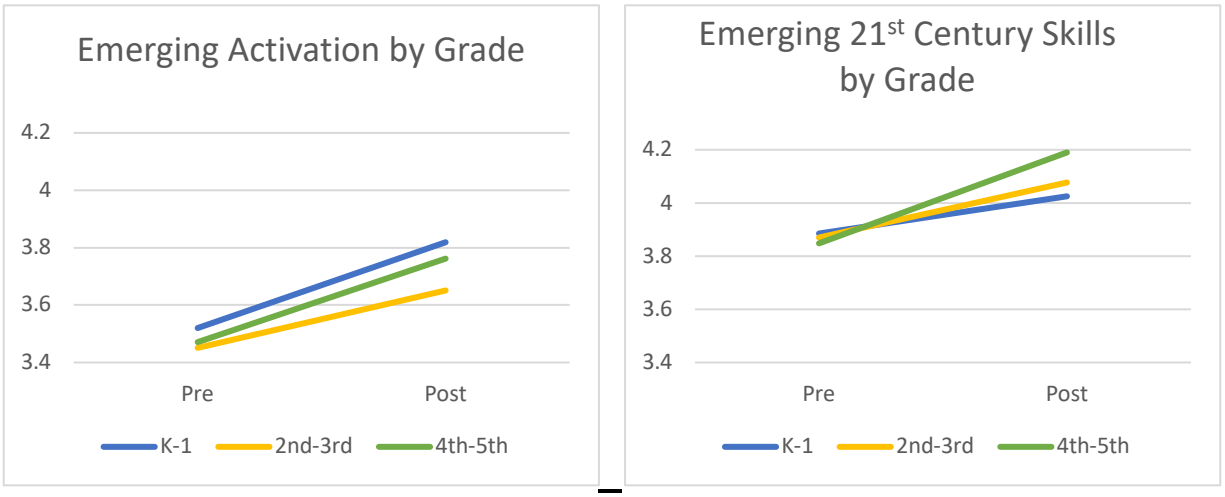


Analyses by Grade

For analyses by grade, grades were grouped together as K-1, 2-3, and 4-5. All three grade groupings showed comparable growth in Emerging Activation after participation. Though Kindergarteners and 1st graders consistently scored highest, and 2nd-3rd graders scored lowest on the Emerging Activation, the three groups both showed comparable gains over time. A series of repeated measures ANOVAs, comparing changes from retrospective pre-scores to post-scores for the grade groupings (i.e., with time as the within-subject factor, and grade as the between subject factor) found no significant time by grade interaction (time x grade $F(2,410 = 1.38, p > 0.05)$. In other words, all three grade groupings (K-1, 2-3, and 4-5) experienced comparable gains in Emerging Activation.

All three grade groupings also showed comparable growth in Emerging 21st Century Skills. A series of repeated measure ANOVAs, comparing changes from retrospective pre-scores to post-scores for the three grade groupings, found no significant time by grade interaction for *FIRST* LEGO League Jr. participants (time x grade $F(2,405 = .66, p > .05)$.

Figure 6. Grade Changes in Emerging Activation and Emerging 21st Century Skills



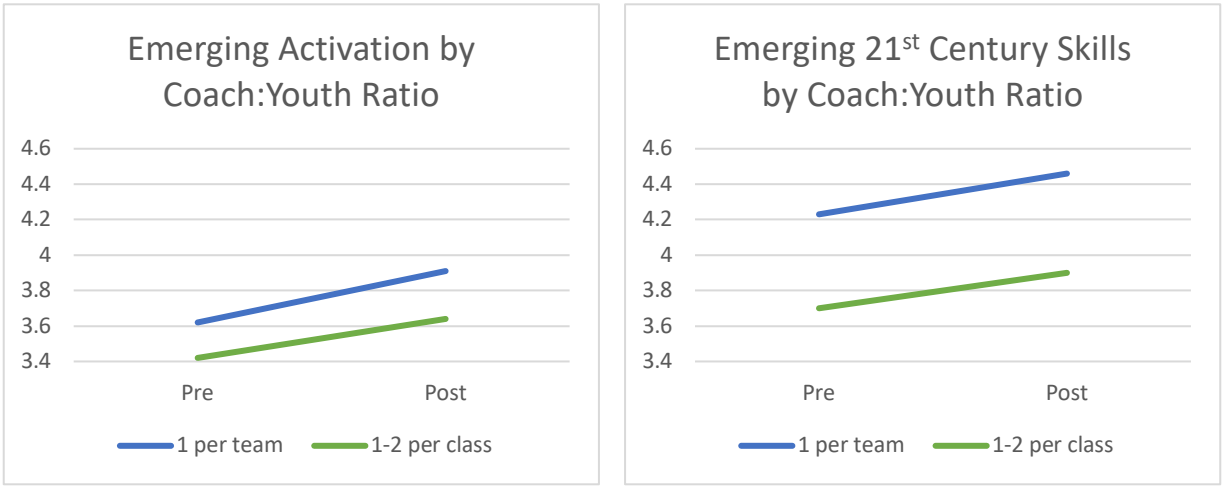
Analyses by Site Implementation Differences

In an attempt to understand the range of changes in Emerging Activation and Emerging 21st Century Skills, we considered a number of factors, such as whether activities were in a school (11 of 12 sites) or another location type (1 of 12 sites); whether they were during the school day (8 sites) or after-school/on weekends/mixed (4 sites); and whether the site was in a large urban area (5 sites), small urban or suburban area (6 sites), or rural area (1 site). None of these variables was systematically related to scores or growth observed.

One factor that was systematically related to scores, though not growth, was coach to youth ratio. Five out of the 10 responding sites had one coach per team, while the other five had one or two coaches overseeing an entire class or site (i.e., a single coach would oversee multiple teams at the same time). When comparing these groups, there were comparable gains over time. However, programs with one coach per team consistently scored higher than programs with 1-2 coaches per class on both Emerging Activation and Emerging 21st Century Skills. (Note that we did not test for statistical significance because there were only 5 cases per group, and thus statistical power was extremely low.) This consistent different in scores may suggest there were

bigger differences between these sites beyond just the coaches. Perhaps the sites that had the possibility of having one coach per team had more established STEM education programs, more established or accessible networks of educators and volunteers, or simply more resources available.

Figure 7. Changes in Emerging Activation and Emerging 21st Century Skills by Coach to Youth Ratio



More discussion of site-level differences can be found in the case site section of this report (Appendix C).

Youth Self-Reported Learning

In the Youth Survey, youth were asked to reflect, "What did you learn in *FIRST* LEGO League Jr. this year?" In their open-ended responses, participants frequently reported learning STEM content and 21st century skills. For instance, 42% of participants reported learning more about working with LEGO bricks, including how to build different structures with LEGO bricks and how to program LEGO bricks to move. Many participants (30%) reported learning more about water, including how people get water, how PlayPumps can help communities get clean water, the different methods by which people clean their water, and the importance of having

access to water. One participant responded, "I learned that some people don't have water and since everyone needs water they should have it." Youth also reported learning about various 21st century skills (28%), particularly teamwork (21%), but also perseverance, creativity, problem-solving, and learning through play. One respondent said, "It isn't just work, you get to have fun!" Another commented, "In this year what I learn is to be a team and never give up."

Coach Survey – What Youth Learned

Coaches were asked to reflect on student learning across three STEM dimensions: awareness, attitudes, and skills; as well as three 21st century skills: teamwork, problem solving, and creativity.

STEM Learning

Coaches first were asked the extent to which *FIRST* LEGO League Jr. helped increase youth's awareness of STEM subjects along a 5-point scale (1 = Not at All, 5 = A lot). Table 20 below summarizes their responses. Coaches reported increases across all subject areas, with the greatest increase in awareness of computer programming (*mean* =4.52, *SD* = 0.73) and the least increase in awareness of math (*mean* = 3.67, *SD* = 1.24).

Table 20. Coach Perceptions of Increased STEM Awareness

	1 Not at all	2	3 Somewhat	4	5 A lot	Mean (SD)
Science	4%	0%	9%	39%	48%	4.26 (0.96)
Computer Programming	0%	0%	13%	22%	65%	4.52 (0.73)
Engineering	4%	0%	8%	25%	63%	4.42 (0.97)
Math	4%	17%	21%	25%	33%	3.67 (1.24)
Local Water Issues	0%	8%	17%	25%	50%	4.17 (1.01)

Coaches also reflected on the extent to which *FIRST* LEGO League Jr. helped improve youth's attitudes toward STEM, in terms of their interest, confidence, and persistence. Table 21 below summarizes their responses. Coaches perceived improvements on all three dimensions, with

slightly higher increases in interest (*mean* = 4.43, *SD* = 0.66) than either confidence or persistence (which had equivalent ratings; *means* = 4.35, *SD*s = 0.78).

Table 21. Coach Perceptions of Improved STEM Dispositions

	1 Not at all	2	3 Somewhat	4	5 A lot	Mean (SD)
STEM Interest	0%	0%	9%	39%	52%	4.43 (0.66)
STEM Confidence	0%	0%	17%	30%	52%	4.35 (0.78)
STEM Persistence	0%	0%	17%	30%	52%	4.35 (0.78)

In reflecting on these changes in interest, confidence, and/or persistence, one coach reflected:

> *The students hear all about STEM and learn the basics to those subjects, but rarely have time to actually complete projects that have real-life application of those skills. This brought their dreams and goals closer to their reach because they felt like real scientists and engineers.*

Lastly, coaches reflected on the extent to which *FIRST* LEGO League Jr. helped youth develop five STEM skills. Table 22 below summarizes their responses. Overall, coaches reported development in all five skills, with the most development occurring in spatial construction and reasoning (*mean* = 4.30, *SD* = 0.82) and the least development in research skills (*mean* = 3.63, *SD* = 1.21).

Table 22. Coach Perceptions of STEM Skill Development

	1 Not at All	2	3 Somewhat	4	5 A lot	Mean (SD)
Spatial construction and reasoning	0%	4%	9%	39%	48%	4.30 (0.82)
Programming and coding through LEGO software	0%	0%	17%	44%	39%	4.22 (0.74)
Understanding science content as it relates to the challenge topic	8%	4%	13%	33%	42%	3.96 (1.23)
Knowing how to ask questions that will advance their understanding and knowledge	4%	17%	8%	33%	38%	3.83 (1.24)
Research Skills	8%	8%	21%	38%	25%	3.63 (1.21)

21st Century Skill Learning

Next, coaches were asked to rate their perceptions of student learning across three 21st century skills: teamwork, problem solving, and creativity. For teamwork, coaches rated the extent to which *FIRST* LEGO League Jr. helped youth develop across seven dimensions on the same 5-point scale as above (1=Not at All, 5 = A lot). Table 23 presents their responses. Reported growth in teamwork was high across dimensions. Coaches rated the most improvement in fairly dividing up work (*mean* = 4.63, *SD* = 0.58) and making decisions as a team (*mean* = 4.64, *SD* = 0.65). They perceived the least improvement in accepting feedback or criticism (*mean* = 4.26, *SD* = 0.92).

Table 23. Coach Perceptions of Growth in Teamwork Abilities

How to…	1 Not at All	2	3 Somewhat	4	5 A lot	Mean (SD)
Make decisions as a team	0%	0%	8%	21%	71%	4.63 (0.65)
Fairly divide up work	0%	0%	4%	29%	67%	4.63 (0.58)
Work with others	0%	0%	13%	17%	71%	4.58 (0.72)
Contribute to the team and help with the project	0%	0%	13%	17%	71%	4.58 (0.72)
Listen to and understand others	0%	0%	8%	29%	63%	4.54 (0.66)
Explain one's own ideas to others	0%	4%	8%	33%	54%	4.38 (0.82)
Accept feedback or criticism	0%	0%	30%	13%	57%	4.26 (0.92)

Coaches provided a number of positive examples of teamwork in their open-ended follow-up responses.

> *At the end right before the Expo one group had their pump come apart. A few students from each group collaborated to help that team rebuild the pump and put their model back together.* [Coach from 1st grade classroom]

> *Students who were not always included in teamwork [in other classes], such as those with an IEP, were excelling at working with others to complete a task, especially the building.* [Coach from kindergarten and 2nd grade classrooms]

There were also some examples of teamwork challenges, which typically related to youth becoming disengaged if their idea or work was not the center of the activity.

> *I think that the roles of the group need to be thoroughly explained in the beginning. I noticed that some of the students became disengaged if they did not think they had an important part.* [Coach from 2nd grade classroom]

> *I had students that were very eager to share their own ideas and bring their own visions to fruition, but they struggled to listen to others. They would share their idea and then forge ahead rather than waiting to hear feedback or further suggestions from teammates.* [Coach from 2nd grade classroom]

Next, coaches rated the extent to which participating in *FIRST* LEGO League Jr. helped youth develop along seven dimensions of problem solving. Responses are presented below in Table 24. Again, coaches reported high ratings for development across all dimensions. The most

improvement was reported for exploring and trying out ideas (*mean* = 4.61, *SD* = 0.50), and the least improvement was reported in understanding a real-world problem (*mean* = 4.26, *SD* = 1.01), problem solving with technology (*mean* = 4.26, *SD* = 1.01), or thinking through the steps of a problem (*mean* = 4.26, *SD* = 0.81).

Table 24. Coach Perceptions of Growth in Problem Solving

How to…	1 Not at All	2	3 Somewhat	4	5 A lot	Mean (SD)
Explore and try out ideas	0%	0%	0%	39%	61%	4.61 (0.50)
Keep trying, even when things don't work out	0%	0%	4%	35%	61%	4.57 (0.59)
Figure out a solution to a problem	5%	0%	5%	32%	59%	4.41 (0.96)
Resolve conflict and negotiate with others	0%	4%	13%	30%	52%	4.30 (0.88)
Think through the steps of a problem	0%	4%	9%	44%	44%	4.26 (0.81)
Understand a real-world problem	4%	4%	0%	44%	48%	4.26 (1.01)
Problem-solve in regards to technology	4%	0%	13%	30%	52%	4.26 (1.01)

Coaches provided a number of positive problem solving examples.

> *I had a group that couldn't understand why their pump was popping apart upon turning and completion. They were able to make the decision to back up in the steps until they found out why their problem was occurring.* [Coach from 2nd grade classroom]

> *…it was tough hearing 5 different ideas at once. So students would come up with their own solutions and write all the ideas down, and each student would then present their idea in front of the team. The students would then vote on what they wanted to do for this Challenge.* [Coach from 4th & 5th grade classroom]

Examples of challenges in problem solving included the following:

> *Sometimes I would have to show them how to fix it because they gave up.* [Coach from kindergarten classroom]

> *The largest team had the hardest time because they could not decide. They did not have a leader or a clear vision and sat helpless many times. Many of them wanted to play with the LEGO bricks or build their own thing, but had a hard time when it came to solving problems with their model.* [Coach from 1st grade classroom]

Finally, coaches rated the extent to which participating in *FIRST* LEGO League Jr. helped youth develop along six dimensions of creativity. As shown below in Table 25, coaches again reported high ratings for development across all dimensions. The most improvement was reported for coming up with creative and original ideas (*mean* = 4.45, *SD* = 0.60), and the least improvement was reported in viewing problems from multiple perspectives (*mean* = 4.09, *SD* = 1.11).

Table 25. Coach Perceptions of Growth in Creativity

How to…	1 Not at all	2	3 Somewhat	4	5 A lot	Mean (SD)
Come up with creative and original ideas	0%	0%	5%	46%	50%	4.45 (0.60)
Think about multiple different solutions to problems	5%	0%	5%	41%	50%	4.32 (0.95)
Adapt ideas to solve new problems	0%	5%	5%	46%	46%	4.32 (0.78)
Learn from failure	0%	10%	10%	20%	60%	4.30 (1.03)
Build from what they learn to come up with new ideas	0%	5%	10%	43%	43%	4.24 (0.83)
View problems from multiple perspectives	5%	5%	14%	32%	46%	4.09 (1.11)

Parent Survey - What Youth Learned

Teamwork

Parents were also asked to talk with their children about what they had learned about teamwork through *FIRST* LEGO League Jr. Youth most commonly reported learning about the value of sharing ideas (39% of responses).

> *Being on a team is fun because hearing other people's ideas helps me learn.* [2nd grade female]
>
> *Sharing our ideas is important so you can figure out new ways.* [2nd grade male]

Youth also reported learning that collaborating/working together (36%) could help the team work more quickly and build better models.

Working together you can get the work done and combine ideas to make something better and faster. [3rd grade female]

It's not always about your way but the team has to work together to accomplish something great. [4th grade male]

Problem Solving

Parents also asked their children what they learned about problem solving. Most commonly, youth replied that they had learned that it is an iterative, trial-end-error process (16%), but it can be made easier by working as a team (16%).

Try, test, fix, repeat test, fix ... not always right the first time. [4th grade male]

She used her observations to retrace her steps and find a solution to the problems. [1st grade female]

We can solve problems by talking through them. [1st grade male]

It's easier with someone else. [3rd grade female]

Though the question was intended to learn about problem solving in general, youth often talked about what they had learned about real world problems (22%).

Learned how simple machines help with problems we face day to day. [2nd grade female]

We learned how to save water and why it was important. [2nd grade male]

Solving problems helps your community. [3rd grade female]

Creativity

Lastly, parents asked their children what they had learned about creativity. Youth were most likely to comment that it was fun to be creative and build the models with their teams (30%).

[I] liked being creative and [I] can make really cool things. [3rd grade female]

It was fun! [kindergarten female]

They also commented on the importance of trying new things (21%), even if it can be difficult or scary sometimes.

When you try new things you just might love it. [1st grade male]

It's fun to try new things and see what happens. [3rd grade male]

Similarly, youth reported that they learned it is good to try to think of new ideas (11%), and that it is okay if the idea does not work out or if you make a mistake (10%).

I learned how to think of new ways to make things better. [1st grade female]

There is no bad idea- you can try as many new ideas as you want. [1st grade male]

It is ok for it not to work, but you cannot give up. [2nd grade male]

Observations– What Youth Learned

Learning through Play

At each of the 12 site visits, the evaluation team rated the extent of evidence that youth were learning through play along the six dimensions specified by the LEGO Foundation: joyfulness, meaningful play/participation, iterative play/participation, socially interactive play/participation, teamwork and communication, and active engagement. Overall, there was "Good" to "Excellent" quality of evidence that youth were experiencing joyfulness (from observation notes: *"When kids enter the room, one says "Oh yay! We get to build!"*), and active engagement. There was less consistent evidence of iterative play/participation and meaningful play/participation.

Table 26. Observer Ratings of Learning Through Play

	1 Very Poor	2 Poor	3 Mixed or Neutral	4 Good	5 Excellent	Mean (SD)
Joyfulness	0%	0%	9%	9%	82%	4.55 (0.58)
Meaningful Play/Participation	0%	0%	18%	55%	27%	3.95 (0.58)
Iterative Play/Participation	0%	9%	27%	27%	36%	3.86 (1.00)
Socially Interactive Play/Participation	0%	0%	0%	73%	27%	4.18 (0.44)
Teamwork and Communication	0%	0%	9%	64%	27%	4.09 (0.56)
Active Engagement	0%	0%	0%	27%	73%	4.59 (0.42)

One site visit captured a fantastic example of iterative play.

> *Students were trying different speeds and other settings to try to get their street sweeper to stop breaking. They were trying to make it slower so it wouldn't break. They would adjust something, test it, and then readjust.*

Likewise, the following situation describes an iterative problem solving process by one student.

> *When water droplets got jammed in the water pump, one student took apart the top level, looked to see what the problem was, and then tried using it again once she had made some changes to the design.*

Other observations, however, found less compelling evidence.

> *This definitely felt more like "free-play" than iterative play. Students weren't trying out possibilities, revising hypotheses, or discovering questions.*

Similarly, for meaningful play, there was high variability across sites. For instance, one site visit included youth engaged in the research phase without much direction (observation notes: "*sometimes, it seemed like students were just aimlessly clicking around on their laptops")*, while a different site visit found youth making several connections between the program content and their lives:

> *Students made several connections from the PlayPumps story to their own lives, with the support of their coaches. They talked about the types of things the kids from the video could do with the time they used to spend collecting water.*

21st Century Skills

Throughout their observations, observers took note of any salient examples of 21st century skills displayed by youth.

Teamwork. There were several instances of notable teamwork during the observations. Sometimes, teamwork involved physically helping each other connect pieces: "*At one point, a girl had trouble pushing two pieces together. Her teammate helped by holding one part while she pushed in the other."* Other times, teamwork included turn-taking, passing pieces to each other, and dividing up the work.

Sometimes, the coaches were extremely deliberate in structuring the activities to promote teamwork. At one site, the coach structured each step of the build so that it was divided into three parts: "Build, Fetch [the next piece], Check." Each part was completed by a different person, and then the model was passed one person to the left so that everyone got an equal turn. At another site, the coaches scaffolded teamwork by encouraging teams to vote on decisions.

> *The coach has the group take a vote on decisions, so everyone gets a voice and knows why decisions were made. Co: okay, hands up. Where do you think our PlayPump should go? {kids point} right there? Does everyone think it should go right there? {kids raise hands} So, as a team, we've decided it should go right there. Okay, should we try it? [YEAH!]*

Communication. Like teamwork, communication was also evident across observations. At one site, a student gently suggested the team move on to a different task after getting too focused on one detail: "*Can we please stop building the walls and build the ramp?"* An observation at another site found a compelling example of teamwork and communication:

> *There were two students working with the WeDo 2.0 (they would press the start and stop buttons) and two students troubleshooting directly with the water pump. The students working on the water pump would have to communicate their needs to the students using the tablet. "Can you stop the pump?" "Okay you can start it" "Can you change the direction of the pump?"*

Again, some coaches were much more deliberate than others about scaffolding opportunities for their youth to develop communication skills. At one site, the coach ended every session by asking youth to compliment other teams' models.

> *T: Okay, what do you notice that you'd like to compliment?*
> *C1: I really liked that table's pump because they made a ramp and pushed the water down to the door.*
> *C2: I liked this table's model because it was kinda like ours.*
> *C3: I liked this table's because it has lots of details.*

<u>*Creativity/Innovation*</u>. Creativity and innovation were frequently apparent during brainstorming phases, such as drawing in the Engineering Notebook. For instance, when learning about the PlayPump, one student drew a ferris wheel and roller coaster to power the pumps. At one site, the teacher not only provided structured time for youth to discuss ideas with their friends, but she even gave them a new vocabulary word: brainstorming.

> *T: How will you collect [the water] without your hands? Turn to your friends and brainstorm.*[after a few minutes] *K, what ideas came up?*
> *C1: You could make a bucket*
> *C2: We could make a bowl or box*
> *C3: We could make a hand.*
> *C4:You could make tubes that go into the ground and go to houses.*
> *T: oh, yes. Tubes. What else do we call those tubes?*
> *Class: pipes!*

<u>*Problem-Solving*</u>. Instances of problem-solving often coincided with instances of teamwork or iterative play. For example, the following describes a situation where a team worked together to solve problems with their water pump.

> *The water pump was getting stuck with water droplets for most of the time that the students were working on it: the students were putting in water droplets too quickly, the water droplets were getting stuck on part of the inside mechanism, the band was coming loose from the motor, or the water pump was shaking around. Each time, the students found the problem and fixed it, together. Eventually, when one student had the idea to change the direction of the spin, the water pump worked smoothly!*

Another instance of problem solving involved backtracking to find where they went wrong.

At one point, the girls noticed that their water pump was looking different from the instructions. They back tracked to the most recent point where their water pump looked correct, and went from there.

Summary of Evaluation Question 3 Results

Overall, data from multiple sources indicate highly positive outcomes for youth who participated in *FIRST* LEGO League Jr. under the Season Pass Model. On survey responses across 10 sites, youth showed small to medium increases in both their Emerging STEM Activation and Emerging 21st Century Skills, with no differences in growth by gender or grade level. There was some variability in growth by site, but with multiple differences across sites in terms of implementation characteristics, student populations, and context, it is challenging to isolate specific variables without a more controlled study. Coaches were highly positive in their ratings of youth's STEM learning, and coaches, parents, and evaluation team observers were highly positive in their assessments of youth's growth in 21st century skills, including teamwork, communication, creativity, and problem solving. The evaluation observers also found the program to be also highly successful in supporting learning through play, with rich opportunities to experience joyfulness, meaningful play, iterative play, socially interactive play, teamwork/communication, and active engagement.

Results - Evaluation Question 4: Strengths, Weakness, and Areas for Improvement

Lastly, we now will focus on addressing Evaluation Question 4: *What are the FIRST LEGO League Jr. Season Pass Model's strengths and weaknesses? What are the areas for program improvement?* Data for this question were drawn from Site Administrator Surveys, Coach Surveys, Parent Surveys, Youth Surveys, and evaluation team reflections.

Site Administrator Satisfaction

Overall, site administrators reported high satisfaction with the *FIRST* LEGO League Jr. program. When asked to rate their overall satisfaction on a 10-point scale, their average response was 8.86 (*SD* = 1.12), and 33% of site administrators selected 10, or "Very Satisfied." Similarly, when asked to rate the quality of the *FIRST* LEGO League Jr. program on a 10-point scale, the average response was 9.05 (*SD* = 1.21), and 57% of respondents selected 10, or "Very high quality." When asked to rate satisfaction with specific components of the *FIRST* LEGO League Jr. program, they were most satisfied with the WeDo 2.0 Robotics kits (*mean* = 4.62, *SD* = 0.39), and least satisfied with the training provided by *FIRST* LEGO League Jr. for both coaches (*mean* = 3.68, *SD* = 0.95) and site administrators (*mean* = 3.78, *SD* = 1.07). This parallels results from the 2016-2017 season, when site administrators requested more training. Table 27 below shows site administrator ratings of satisfaction across 8 elements of *FIRST* LEGO League Jr.

Table 27. Site Administrator Satisfaction with *FIRST* LEGO League Jr., from Most Satisfaction to Least Satisfaction

	1 Very Dissatisfied	2 Dissatisfied	3 Neutral	4 Satisfied	5 Very satisfied	Mean (*SD*)
WeDo 2.0 Robotics Kits	0%	0%	0%	38%	62%	4.62 (0.39)
Other physical materials provided by FLL Jr. (e.g., LEGO bricks)	0%	2%	7%	29%	62%	4.50 (0.74)
Program Administrator Guide (for your use)	0%	0%	17%	34%	49%	4.32 (0.76)
Team Meeting Guide (for coaches)	0%	2%	12%	38%	48%	4.31 (0.78)
Engineering Notebook (for youth)	0%	7%	12%	33%	48%	4.21 (0.93)
The year's challenge theme, "AQUA ADVENTURE"	5%	7%	10%	38%	41%	4.02 (1.12)
Training provided by FLL Jr. (for site administrators)	3%	6%	31%	28%	31%	3.78 (1.07)
Webinar Training for coaches	0%	5%	50%	18%	27%	3.68 (0.95)

Site administrators selecting "N/A" were excluded itemwise from percentage calculations. N's ranged from 22-42.

Open-ended responses throughout the survey suggested that more communication may be needed around training. For example, one site administrator shared, "I was not aware that there was training available," and another reflected "I was not aware of webinars or any other training available to me." Another theme in site administrators' dissatisfaction with the program related to the AQUA ADVENTURE theme, with some saying it was "too abstract," "not interesting," or that it "did not grab the students' attention." Nevertheless, these comments came from just a handful of site administrators, and overall ratings were quite high; across the 8 elements, site administrators averaged 4.23 (*SD* = 0.60) out of 5.

Coach Satisfaction

Overall, coaches had very positive experiences in *FIRST* LEGO League Jr. When asked to rate their satisfaction on a scale of 0-10, (0 = Very Dissatisfied, 10 = Very Satisfied), coaches averaged 8.32 (*SD* = 1.96). They rated the quality of the program as 8.59 out of 10 (0 = Very low quality, 10 = Very high quality; *SD* = 1.53). Given this high satisfaction and the high quality ratings, it is unsurprising that when asked whether they would recommend the program to a colleague (0 = Not at all likely, 10 = Extremely likely), they averaged 8.73 (*SD* = 1.75). Considering those who selected "9" or "10" as promoters and those who selected "6" or less as detractors, the resulting net promoter score (percentage promoters minus percentage detractors) was 59.2, which is considered excellent. The majority (77%) of coaches planned to coach *FIRST* LEGO League Jr. again next year, with the remaining 23% unsure about their plans.

When asked what they liked most, coaches most frequently mentioned seeing youth develop 21st century skills (35% of responses), including problem solving (18%), teamwork (12%), and creativity (6%). They also commonly liked seeing the youth excited about and proud of what they were doing (24%), as well as seeing them learn about STEM in a hands-on way (24%). Other responses included the student-centered nature of the activities (18%), and the opportunity for youth to learn through play with LEGO bricks (12%). One coach summarized,

> *I loved that the kids could lead themselves through the process and that they SMILED THE WHOLE TIME! They LOVED this and really were able to learn from it.* [coach from 2nd grade classroom]

When asked what they liked least, the most common response theme was related to materials (53% of responses). Specifically, 29% of responses reported negative feedback on the Engineering Notebook, with one coach explaining, "I think that the workbooks are much too long" (other responses were too vague to interpret). Other common response themes were time

management (24% of responses) and classroom management (12%), with additional responses related to "training" (6%) and "cost" (6%).

Parent Satisfaction

Overall, parents were highly satisfied with their child's experience in *FIRST* LEGO League Jr. On a scale of 0-10, where 0 = Very dissatisfied and 10 = Very satisfied, 67% of parents selected 10, and the remaining 33% ranged from 7-9 (overall *mean* = 9.41, *SD* = 0.94). Parents were asked what they thought was the greatest benefit their children received from participating in *FIRST* LEGO League Jr. Open-ended responses most frequently cited "teamwork" (33%), STEM content and skills (31%), and having fun (20%). Within STEM content and skills, parents typically described engineering concepts and skills (e.g., steps of the engineering design process; 20%), but also concepts and skills related to technology (9%), coding (7%), and water (7%).

Youth Satisfaction

In general, youth across programs had fun in *FIRST* LEGO League Jr., and nearly all would participate again. Across the 10 sites, when asked if they had fun in *FIRST* LEGO League Jr., 78% said yes, 18% had mixed feelings, and only 4% said no. When asked if they would like to do *FIRST* LEGO League Jr. again with a different challenge, 89% of respondents said yes, 7% were unsure, and 5% said no.

Table 28. Youth Satisfaction

Item	1 NO!	2 no	3 Mixed	4 yes	5 YES!	Mean (SD)
Did you have fun in *FIRST* LEGO League Jr.?	3%	1%	18%	0%	78%	4.48 (1.03)
Would you like to do *FIRST* LEGO League Jr. again next year?	4%	1%	7%	11%	77%	4.57 (0.95)

Overall, youth appreciated multiple aspects of their experiences. Participants were asked, "What was your favorite part of *FIRST* LEGO League Jr.?" Most reported enjoying building with LEGO bricks (46%) and programming with the WeDo 2.0 (11%). Several (14%) said their favorite part was working with others on their team.

Parent Survey Reflections. In addition to providing their own reflections, parents were asked to talk with their child about their experiences in the program, including how much they liked different aspects of the program. As shown in Table 29, youth had highly positive reactions to all aspects of their *FIRST* LEGO League Jr. experience. Their most positive ratings were about building, including "building with LEGO bricks" (*mean* = 4.86 out of 5, *SD* = 0.45), followed by "building with friends" (*mean* = 4.63, *SD* = 0.69), and "programming a small robot" (*mean* = 4.57, *SD* = 0.97).

Table 29. How much did your child like the following activities?

Item	1 Not at all	2	3 Somewhat	4	5 Very much	Mean (SD)
Building with LEGO bricks	0%	0%	4%	6%	90%	4.86 (0.45)
Building with friends	0%	1%	8%	18%	73%	4.63 (0.69)
Programming a small robot	3%	2%	12%	3%	81%	4.57 (0.97)
Learning about ideas related to science, technology, engineering, or math	1%	0%	8%	27%	64%	4.53 (0.75)
Working as a team	0%	4%	9%	19%	68%	4.52 (0.81)
Learning about water	0%	3%	13%	19%	65%	4.47 (0.82)
Sharing ideas with others	1%	4%	6%	23%	65%	4.47 (0.88)

Parents were also asked to have their child reflect on the Expo. At the time of the survey, 54% of youth had attended an Expo and 5% were planning to in the near future. A third (31%) of parents and youth reported that an Expo was not offered, and 10% reported that an Expo was offered but the child did not attend. Of those youth who attended an Expo, 80% "liked it a lot," 18% "liked it at little," and 2% "neither disliked nor liked it."

Summary of Evaluation Question 4 Results

Data indicate that site administrators, coaches, parents, and youth were all highly satisfied with their experiences in *FIRST* LEGO League Jr. under the Season Pass Model and found the program to be very high quality. Coaches particularly enjoyed watching youth learn 21st century skills through play. Youth enjoyed building with friends and were enthusiastic about participating again in the future. Suggestions for improvement from participants included providing more flexibility or tips for adapting materials and activities and providing more training to support coaches and site administrators in managing the program's complex implementation. In the section below, we expand upon these suggestions, and add additional recommendations from the Evaluation Team, particularly related to expanding data collection methods to gain a clearer understanding of implementation variability and the program's impact.

Recommendations

Based on surveys, observations, and reflections, the evaluation team has identified a number of recommendations to improve the experiences of site administrators, coaches, and youth; support even more positive outcomes for youth; increase the reach of the Season Pass Model; and improve data collection procedures.

Because the Season Pass Model is frequently implemented in schools during the school day, consider providing suggestions for dividing sessions into shorter segments. The majority of Season Pass holders were either schools or school districts, and nearly all sites included activities hosted at a school, often during the school day. Often, school day schedules are not structured to allow for a full hour to be dedicated to *FIRST* LEGO League Jr. activities. In fact, when asked how *FIRST* LEGO League Jr. could improve their experience, coaches frequently suggested shorter activities or suggestions for "chunking" activities to fit within time constraints (17%), and more guidance for how to manage time within activities (17%). Many sites reported making their own decisions about which activities to keep as written, shorten for time, or cut out entirely. To guide sites on making these decisions, and to improve overall fidelity, it may be worthwhile to either break activities into shorter sessions (e.g., 15 sessions that are 45-minutes each), or provide suggestions for optimal ways to fit sessions into compressed timeframes.

Provide suggestions for modifications to reach youth at different developmental levels or with different background knowledge. *FIRST* LEGO League Jr. is designed for youth from kindergarten through fifth grade, a pivotal period for developing early STEM skills, knowledge, and dispositions. However, it is also a very wide developmental range, with vast differences in experience, knowledge, and capabilities between the youngest and oldest participants. Therefore, though the themes and general structure of the activities are entirely appropriate for the age

range, there may be a need for increased supports to help coaches make slightly modifications for different age groups or developmental levels. Indeed, when asked how *FIRST* LEGO League Jr. could improve their experience, coaches frequently suggested more guidance on how to differentiate the activities for younger learners (17%). This request often came up around activities in the Engineering Notebook (e.g., too much writing) or in the research stage (e.g., not enough developmentally appropriate research materials available).

Consider providing tips on how to structure the room or activities to allow for more student-driven explorations. Under the Season Pass Model, it is common for there to be a single coach leading multiple teams at once. As seen in the observations (see Appendix B and C), some coaches in these types of settings approached *FIRST* LEGO League Jr. activities as if they were leading traditional classroom instruction. In these classes, activities felt extremely teacher-driven. Whether teachers settled into this routine because it was what they were used to, because they thought it was the best way to manage the classroom, or for some other reason, is unknown. However, it may be helpful for *FIRST* to provide specific structural suggestions for ways to support more student-driven activities. For instance, perhaps it may be worthwhile to consider decreasing the number of youth per team to 4 or 5. Many coaches suggested that having smaller teams would decrease behavior problems and make it easier for youth to stay on task. Perhaps by decreasing the likelihood of behavior problems, smaller teams would allow coaches to feel more comfortable allowing youth more autonomy in activities.

Provide more training for coaches and site administrators, and/or provide clearer communication on the training that is available. In both the 2016-2017 site administrator survey and the 2017-2018 surveys, site administrators and coaches requested more training from *FIRST* on how to implement the program. This year, when asked for recommendations for

improvement, as many as half of coaches suggested more training (50%). In fact, some coaches appear to have received no training whatsoever. One coach reflected "The lack of any training set us up for a rough start and first session." Though the materials provided by *FIRST* LEGO League Jr. are thoughtfully designed and provide high quality guidance on coaching activities, it still seems like more support or training is needed. In addition, increased communication about the training that is available would also be appreciated. A number of coaches or stakeholders were completely unaware that a webinar was offered, for example.

Increase recruitment in the Mountain States and Southwest. These states (MT, ID, WY, NV, UT, CO), as well as ND and SD, had no participating sites during the 2016-2018 seasons. Thus, they may be important to target if geographic diversity is a goal of *FIRST*. Moreover, many of these states have high populations of Native American families, who were largely unrepresented in the participating sample. Working with these communities would serve to benefit *FIRST* in making progress towards its goal of working with underrepresented and/or underserved populations, as Native American communities in the United States face a number of educational and social barriers, making them among the least represented in STEM fields.

Prepare coaches for youth survey administration. Without direct access to youth at Deep Dive sites at the end of the season, the present study relied on coaches to assist with youth survey administration. Unlike traditional *FIRST* LEGO League Jr. coaches, who worked with just 6 youth, Season Pass Model coaches oversaw as many as 30 youth at a time. This increased the challenges with administering student surveys. To gauge the difficulty of the endeavor, the evaluation team surveyed coaches for their feedback on the process of administering the survey. Less than half of the responding coaches (7 of 17) indicated challenges administering the survey. Those who reported challenges either thought it took longer than expected (3 responses) or that it

was hard to get youth to think back to *before FIRST* LEGO League Jr. (3 responses). Nearly all (16 of 17) reported reading the survey aloud to the group (whether a whole class or a whole team was not clear), with just one person completing the surveys one-on-one. Responses came from coaches working with all grade levels, from kindergarten up. Therefore, it seems feasible to have Season Pass Model coaches lead survey administration with this age group; however, they should be provided with clear guidance on the time involved and be given clear scripts on how to get youth to think retrospectively.

Collect both student-level and site-level demographic data. The registration dataset contained extensive missing data, with only about a third of participants having complete data on all variables. It is highly probable that many site administrators simply lacked access to data at this level of specificity. Moreover, even site administrators who had such access may have been reluctant to enter the data due to the demands required enter it for each participating child, when, under the Season Pass Model, many sites had of upwards of a hundred youth participating. It may be worthwhile to *also* ask for site-level estimates from site administrators (e.g., estimated percentage of youth at a site within each racial/ethnic category). Though decidedly less precise, these site-level estimates are less demanding on site administrators and may increase the number of sites responding. For example, 78% of site administrators provided estimates for free/reduced lunch eligibility in the site administrator survey. In the registration data, however, student-level data on free/reduced lunch eligibility was available for just 45% of participants. To be sure, child-level data is more informative and necessary for calculating percentages underrepresented/underserved because it allows for calculations across variables. Nevertheless, collecting data through multiple approaches may provide a clearer picture of the youth being served under the Season Pass Model.

Oversample for parent and coach surveys. Low response rates are a recurring challenge in evaluation, and the current project's low response rates were challenging but not surprising. In some cases, however, response rates in the present project were lower than those seen in previous collaborations between The Lawrence Hall of Science (evaluators) and *FIRST* LEGO League Jr. For instance, in the 2014 evaluation (Lee, Dorph, Newton, Chung, & Cannady, 2014), there was about 53% response rate for parents who were invited to take surveys. The context for that evaluation was different in a number of ways, but one key distinction may be in the different dynamics between the Season Pass Model and the more traditional *FIRST* LEGO League Jr. Model. Traditionally, once the evaluation team received the consent of a coach to participate in the evaluation, the coach could then assist in making contact between the evaluator and the six parents on that team. Under the Season Pass Model, however, there may be multiple degrees of separation between evaluator and parent (e.g., evaluator to site administrator to coach to parent), parents may not even be aware of the program, or parents may face unique challenges completing surveys due to their personal situations (e.g., work schedules, living arrangements, access to Internet) as individuals from communities that are traditionally underrepresented and/or underserved. Additionally, instead of having just six parental contacts, coaches and/or site administrators may have as many as 100+, placing a higher burden on individuals in coordinating any sort of data collection or outreach, particularly at sites where the site administrator is also the sole coach. Therefore, the low response rates in the present work likely reflect these many challenges in collecting data under the Season Pass Model.

One potential approach to increase the number of responses (though not response *rates*), would be to oversample. For instance, inviting coaches and/or parents from *all* sites, and not just Deep Dive sites, to complete surveys would have increased the number of responses and data

available for analyses. Though there would still be issues with representativeness (e.g., do parents who respond differ from parents who do not?), larger samples would allow for more robust analyses of the variability of experiences and outcomes under the Season Pass Model.

Appendices

Appendix A: International Case Site Report

Appendix B: Domestic Case Site Details – Site by Site

Appendix C: Case Site Comparison: Two Domestic Sites

Appendix D: *FIRST* LEGO League Jr. Season Pass Logic Model

Appendix A – International Case Site Report

Understanding the International Context – *FIRST* LEGO League Jr. in Mexico as a Case Study in International Expansion of the Season Pass Model

Historically, the *FIRST* programs, including *FIRST* LEGO League Jr., have demonstrated great success in reaching international audiences, but disseminating the *FIRST* LEGO League Jr. Season Pass Model presented unique challenges for working with international partners. Traditionally, other *FIRST* programs have operated with individual teams working independently, coordinated by a partner organization. Because the Season Pass Model requires multiple teams participating within sites, an additional layer of coordination is needed to coordinate both multiple sites and multiple teams within sites. Thus, though the Season Pass Model would enable the participation of more youth around the world, it may also place an increased demand on partnering organizations in terms of managing implementation. Moreover, if the goal of the Season Pass Model is to enable a more diverse audience to participate in the program, the question of what counts as "diverse" – of what populations are traditionally underrepresented in STEM—will vary in different countries and contexts, and the international organization will be an essential partner in defining and reaching populations of interest. Therefore, under the Season Pass Model, the role of the partner organization may take on a different form.

To gain insight into factors influencing international dissemination of the *FIRST* LEGO League Jr. Season Pass Model, the evaluation team worked together with *FIRST*'s partner for program delivery in Mexico, RobotiX®. RobotiX is an educational organization, based in Mexico City, Mexico, that offers classroom-based and extracurricular STEM educational programming through games, robotics, and technology-infused activities and challenges for youth from early

childhood through high school. In operation for over 10 years, the organization has sites across the country and has established a wide network of partnering schools and organizations.

The present case study sought to understand the following evaluation question: In what ways can a case study of implementation in Mexico inform dissemination of *FIRST* LEGO League Jr. to other international sites? Specifically, the key ideas we wanted to explore were:

- *How FIRST LEGO League Jr. fits into educational goals of the country or region*
- *How FIRST LEGO League Jr. Season Pass Model extends the country or region's ability to achieve those goals*
- *How the FIRST LEGO League Jr. Season Pass Model compares to other activities, programs, and initiatives available in the country or region*
- *How implementation is influenced by context: how local and regional elements, such as resources, cultural norms, or other factors, influence implementation*
- *Which populations are traditionally underrepresented in STEM in the country or region; challenges of reaching those populations (including culture, geography, resources, language, political factors, infrastructure)*
- *Strengths of the FIRST LEGO League Jr. Season Pass Model and challenges faced in implementing the FIRST LEGO League Jr. Season Pass Model; how can the FIRST LEGO League Jr. Season Pass Model be improved*
- *Special concerns to consider when partnering with international agencies to implement the FIRST LEGO League Jr. Season Pass Model*

Design

To investigate the above questions, we used a mixed methods approach. First, we conducted an interview with RobotiX staff in the fall of 2017. Through this conversation, we co-developed a plan for data collection in the Spring of 2018. Specifically, we decided that the evaluation team would (1) conduct an interview with RobotiX staff to learn more about program operations and the broader context, and (2) survey Mexican site administrators to learn more about site-level implementation. RobotiX agreed to translate the survey to Spanish and distribute the survey to its sites.

The Spring 2018 interview with RobotiX staff covered topics related to *FIRST* LEGO League Jr. coordination and implementation, and the context in which *FIRST* LEGO League Jr.

was implemented. The interview, which lasted about an hour, was conducted over Zoom, an online video conferencing platform, with the Chief Operations Officer and an Academic Advisor. Questions for the semi-structured interview protocol were developed in consultation with the *FIRST* Research and Evaluation team to ensure alignment with their goals for international data collection.

The site administrator survey was administered online via *Qualtrics*. In general, survey items were similar to those of the Site Administrator Survey for sites in the United States and Canada, with a few modifications. Questions covered topics related to general site information, site administrators, coaches, participating youth, and the program in general. There were 11 responding sites, though the number of respondents per question varied from 7-11. As with the interview protocol, the Mexican version of the Site Administrator Survey was developed in consultation with the *FIRST* Research and Evaluation team before being translated into Spanish by RobotiX. Following the initial translation, a fluent Spanish-speaking member of the evaluation team back-translated the survey to confirm that it captured the original intent of the questions.

Context

The interview with RobotiX staff yielded interesting insights into the context of *FIRST* LEGO League Jr. in Mexico. RobotiX staff reported that there had been a recent shift in focus to STEM education by Mexico's Ministry of Education. However, they noted that the effects had been seen more in private schooling than public schooling. The RobotiX Chief Operations Officer reflected on the role of RobotiX and *FIRST* LEGO League Jr. in supporting change on a broader spectrum (i.e., beyond private schools). He noted that the sites RobotiX approached for participation in *FIRST* LEGO League Jr. might not have previously prioritized STEM education,

but after participating in the program, they reported or demonstrated shifts in their goals or priorities. The Academic Advisor, too, noted that after introducing sites to the Inspire Sets and explaining the *FIRST* LEGO League Jr. program goals, sites were interested in continuing to offer STEM opportunities to youth, and helping youth grow in their STEM skills.

Additionally, RobotiX staff commented that while the Ministry of Education's recent emphasis on STEM caused an increase of STEM program offerings, other STEM programs did not focus on younger students in the same way that *FIRST* LEGO League Jr. did. They noted that other STEM programs used materials such as Arduinos, which were not practical for use with younger students. They felt that the use of LEGO bricks and the WeDo 2.0 were better suited for younger audiences. Even within RobotiX itself, staff felt that *FIRST* LEGO League Jr. was complementary to the other programs they offered, which were aimed at an older age group.

RobotiX Implementation Experience

During the interview, the Chief Operations Officer and Academic Advisor reflected on their experience implementing *FIRST* LEGO League Jr. Season Pass in Mexico. The Chief Operations Officer's role was to focus on communication with site administrators and community partners. The Academic Advisor focused on communicating with coaches. Both agreed that implementation of the program was smoother during the second year, due to the experience they had gained in the first year. The Academic Advisor noted that during the first year, RobotiX did not have a team in place that was ready to implement the program, and they were unsure how to structure a team and delegate roles to implement the program efficiently and effectively. It was challenging to learn how to work as a team to balance the new roles and responsibilities required of *FIRST* LEGO League Jr. implementation. In this way, the first year

was a learning process for the whole team, but after going through the process, they were better prepared for the second year.

Additionally, during the interview, the Chief Operations Officer shared that there was uncertainty around how to handle site administrator and coach training during the first year. After providing sites with the program materials, they were unsure about what follow-up would be helpful. Again, however, the first year served as a learning process. During the second year, RobotiX was prepared to provide a training and site visit prior to program implementation, and weekly webinars and newsletters during the program. The Academic Advisor was responsible for these supports.

After learning about the broader context for *FIRST* LEGO League Jr. activities, we next turned to Site Administrator Surveys to learn more about site-level implementation. The following section provides a summary of implementation findings.

Program Implementation in Mexico

Overall, implementation in Mexico was extremely similar to implementation in the United States and Canada, in terms of site, team, and program organization; site administrator, coach, and youth characteristics; challenges experienced; and recommendations for improvement.

General Site Information

A total of 11 site administrators, located in 5 states, completed the Site Administrator Survey. Overall, 8 out of 9 sites reported holding *FIRST* LEGO League Jr. activities in a single location or building, while only one site reported using multiple locations or buildings. The site that reported using multiple locations specified that *FIRST* LEGO League Jr. activities took place at 3 locations. Overall, most respondents reported meeting at a youth or community organization

(6 out of 10) or in a school (5 out of 10). Most respondents reported hosting *FIRST* LEGO League Jr. after school (7 out of 10), while only a few provided activities during the school day (2 out of 10) or on weekends (2 out of 10). Though most sites did not meet during the school day, 8 out of 10 still made connections to a broader curriculum. All eight made connections to science, three made connections to social studies, and two made connections to language arts/reading.

Sites varied slightly in their schedules for implementing *FIRST* LEGO League Jr. activities. On average, the 9 responding sites hosted an average of 1.33 meetings per week (*SD* = .50), for 2.75 hours (*SD* = .92), over 11.5 weeks (*SD* = 1.65). The majority of programs (6 out of 9) met once a week, while the remaining 3 programs met twice a week. Common meeting lengths were 3 hours (6 sites), and 2 hours (2 sites).

RobotiX reflections on site variation. When asked to reflect on variations in program implementation, RobotiX staff reported differences in implementation as being a product of the number of youth, community, or events outside the control of RobotiX staff or site administrators. For example, RobotiX staff mentioned that some sites had as few as eight youth, while others had up to 30 youth. They also mentioned that depending on the site, the youth had varying levels of comfort with LEGO bricks and building. At one site, many of the youth had parents who worked at the LEGO factory, increasing their exposure to and comfort with LEGO bricks and building. Meanwhile, there were sites with limited access to computers and materials such as LEGO bricks prior to the implementation of *FIRST* LEGO League Jr., and these youth came into the experience with limited familiarity or comfort.

Another factor that influenced variation between sites was the disastrous earthquake in central Mexico in September 2017. After the earthquake, some sites had to make significant

adjustments to implementation, including schedule changes and site relocation. At one site, teams could not meet in their building because there were structural and electrical issues. Youth at this site had to work outside for some of the program, until they could get back into the classroom. Interestingly, RobotiX mentioned that these youth were among the first to be allowed entry back into their school because *FIRST* LEGO League Jr. was highly valued by the school. RobotiX staff also mentioned that, as sites saw interruptions to their regularly scheduled programming, they adjusted their meeting schedule to make up for the time lost. For example, if they could only meet once during a particularly difficult week, they would meet multiple times the next week. RobotiX staff mentioned supporting all sites by sending out weekly newsletters with the week's focus and having weekly webinars, so sites were more or less on the same page despite individual disruptions that came up.

Participant Information

Site administrators were asked to provide a count of the youth who attended at least 75% of *FIRST* LEGO League Jr. sessions. Across the 9 responding sites, approximately 248 youth regularly attended *FIRST* LEGO League Jr. activities in Mexico. The average number of youth attending per site was 28 (*SD* = 3.24). There were significantly more boys (62%) than girls (38%) who regularly attended the program. In terms of grade, most regularly attending youth were in the 4th (28%), 3rd (27%), or 2nd (21%) grade.

Table 1. Site Administrator Counts of Regularly Attending Youth by Gender and Grade in Mexico

Gender	Count (%)	Total
Boys	153 (62%)	248
Girls	95 (38%)	
Grade	**Count (%)**	**Total**
Kindergarten	3 (1%)	215*
1st grade	22 (10%)	
2nd grade	46 (21%)	
3rd grade	57 (27%)	
4th grade	60 (28%)	
5th grade	18 (8%)	
6th grade	9 (4%)	
N=9		

* Note: we calculated overall participant rates using the Gender totals, as there appeared to be less missing data.

Site administrators were also asked to report what percentage of youth were from backgrounds typically underrepresented and/or underserved in STEM. Across sites, approximately half of participants (*mean* = 52%, SD = 32%) were from low socioeconomic backgrounds, and over a third (*mean* = 39%, SD = 25%) were girls, but other underrepresented/underserved groups were less involved.

Table 2. Approximate Percentage of Youth who were Underrepresented or Underserved in Mexico

	Min	Max	Mean (*SD*)
Low socioeconomic background	20%	100%	52% (32%)
Girls	10%	70%	39% (25%)
Students with disabilities	0%	10%	4% (5%)
Indigenous students	0%	5%	1% (2%)
Other	0%	30%	15% (21%)

RobotiX reflections on Participants. When asked if there was a focus on a specific population typically underrepresented or underserved in STEM, RobotiX staff reported focusing on reaching girls. They noted that about 65% of participants were boys, and the remaining 35%

were girls, which is consistent with the averages reported by site administrators. As previously mentioned, RobotiX staff reported that there were more STEM education opportunities for youth in private schools in Mexico than for those in public schools. Therefore, public school youth could be considered an unanticipated audience underserved in STEM within the Mexican context. All sites participating in the program included student populations from public schools and community centers. Thus, if we include this variable as an indicator of being underserved, there would be a high degree of success in reaching underserved populations than the above table suggests.

Expo

The majority of sites (10 out of 11) reported hosting an Expo(s) at their site. For those hosting an Expo, the number of Expos ranged from 1 to 2 (*mean* = 1.22, *SD* = .44), and 9 of these sites reported that "most youth" attended or will attend an Expo, while 1 site reported "Some youth attended or will attend an Expo, but not a majority." For the 1 site *not* hosting their own Expo, they reported that "most" attended an Expo elsewhere. All of the sites, including the one that did not host its own Expo, reported that parents were invited to attend the Expo. The percentage of parents that actually attended an Expo varied from 60% to 100% (*mean* = 87%, *SD* = 12%), with 5 out of 10 sites responding that 90% of invited parents attended the Expo.

When asked if and in what ways the Expo was a positive or negative experience for youth, all 10 respondents reported that it was a positive experience. Site administrators appreciated the opportunity for youth to present their work (6 out of 10 respondents) and interact with their families at the Expo (4 respondents). A couple of site administrators also thought it was a good opportunity for youth to overcome a challenge and gain confidence. As one site administrator put it,

> *It was challenging because they have to overcome their nerves of presenting in front of their families and the LEGO judge. Most of the students accomplished it with success and it's one of the experiences that pushes them to their limits and at the end of the Expo they're proud to have overcome a challenge.*

When asked for recommended changes to the Expo to maximize positive experiences for youth, most site administrators (5 out of 9) did not have any suggestions. The remaining four site administrators recommended inviting more guests, finding a larger venue, adding in an exhibition area for other activities, and providing hands-on materials for guests to interact with.

RobotiX reflections on the Expos. RobotiX staff complimented the celebration of all teams during the Expo, especially in contrast with the more competitive nature of other programs' culminating festivals. RobotiX staff appreciated that the Expo was not focused on prizes like "Best Robot," but instead was a place where all teams could celebrate together and share something about the experience.

Site Administrators

Most site administrators had either a bachelor's degree (4 out of 7) or master's degree (2 out of 7). Typically, their fields of study were either education (3 out of 7) or psychology/social work (2 out of 7). About half of the respondents were female (4 out of 7). The majority (6 out of 7) had never worked in a STEM occupation. When asked to report their current occupation, six site administrators provided responses, three of whom worked in the field of education and three of whom worked in community centers as either an administrator or manager.

Site administrators were also asked to rate the level of challenge experienced in different aspects of being a site administrator, from 1 ("Not at all challenging") to 5 ("Extremely Challenging"). The most challenging aspects were having sufficient resources (e.g., funding, materials, space; *mean* = 3.14, *SD* = 1.21), managing the number of youth on teams (*mean* = 2.57, *SD* = 1.51), and coordinating multiple teams (*mean* = 2.57, *SD* = 1.40). The least

challenging aspects were recruiting youth (*mean* = 1.00, *SD* = 0.00) and handling youth attrition/dropouts (*mean* = 1.57, *SD* = 1.13).

Table 3. Mexican Site Administrator Ratings of Challenges, From Most Challenging to Least Challenging

	1 Not at all challenging	2 Slightly challenging	3 Moderately challenging	4 Very challenging	5 Extremely challenging	Mean (SD)
Having sufficient resources	0	3	1	2	1	3.14 (1.21)
Coordinating multiple teams	1	4	0	1	1	2.57 (1.40)
Managing the number of children on teams	2	2	1	1	1	2.57 (1.51)
Scheduling	2	1	3	1	0	2.43 (1.13)
Following the program's guidelines or expectations	1	4	0	2	0	2.43 (1.13)
Handling youth attendance problems	1	4	2	0	0	2.14 (.69)
Handling coach attendance problems	2	3	2	0	0	2.00 (.82)
Managing building or technical aspects of the activities	3	2	1	1	0	2.00 (1.16)
Recruiting coaches	5	0	0	2	0	1.86 (1.46)
Training coaches	4	1	1	1	0	1.86 (1.22)
Recruiting traditionally underrepresented youth	4	1	1	1	0	1.86 (1.21)
Managing youth behavior problems	3	2	2	0	0	1.86 (.90)
Handling youth attrition/drop outs	5	1	0	1	0	1.57 (1.13)
Recruiting youth	7	0	0	0	0	1.00 (.00)

Site administrators selecting "N/A" were excluded itemwise. N = 8

When asked how *FIRST* LEGO League Jr. could improve site administrators' experiences, respondents suggested modifications related to the WeDo 2.0 programming, Expo logistics, and materials. Specifically, respondents wanted more training on the WeDo 2.0 programming, clarity about what is involved in taking youth to the Expo (including expenses and transportation), and more LEGO kits.

These challenges notwithstanding, 7 out of 8 site administrators planned to return as the site administrator next year. When those expecting to return were asked how prepared they felt to host the program next year, on a scale from 1 ("Very unprepared") to 5 ("Very prepared"), 4 out of 6 responding site administrators reported feeling "Prepared." When asked what they will do differently, they reported they will distribute resources better, work on relationships with parents, and consider the distribution of their time and budget. When the one site administrator not returning was asked why, they reported that another employee would be given the opportunity.

Coaches

A total of 26 adults coached in Mexico through the *FIRST* LEGO League Jr. Season Pass 2017-2018 season. The number of coaches per site ranged from 1-6, with an average of 3.5 (*SD* = 1.92) coaches per site. Coaches came from a range of sources, including teachers from the hosting site (at 5 out of 10 sites), non-educator staff from the hosting site (3 out of 10 sites), and community volunteers (1 out of 10 sites). Across all sites, site administrators estimated that approximately 78% of coaches had a background in education, 65% had experience working with children of this age, and 49% had a background in STEM. The average number of coaches per site was 3.50 (*SD* = 1.93), meeting the RobotiX requirement to have a minimum of 2 coaches for 30 children. Most sites (9 out of 11) reported that coaches oversaw multiple teams at once, while 2 out of 11 sites had just one team per coach. For those coaches overseeing multiple teams, the average was 4.80 teams (*SD* = 1.10) per coach, ranging from 3 to 6 teams.

Though site administrators rated them as "successful" across a range of responsibilities, coaches had especially high levels of success in developing relationships with kids and getting youth excited about STEM (*mean* = 4.50 out of 5 for both), and slightly lower levels of success

in managing behavior (*mean* = 4.25) and incorporating STEM content into instruction (*mean* = 4.13).

Table 4. Site Administrator Ratings of Coach Success

	1 Very Unsuccessful	2 Unsuccessful	3 Mix of successful and unsuccessful	4 Successful	5 Very Successful	Mean (*SD*)
Developing relationships with kids	0	0	1	2	5	4.50 (.76)
Getting kids excited about STEM	0	0	1	2	5	4.50 (0.76)
Managing behavior	0	0	2	2	4	4.25 (.89)
Incorporating STEM content into instruction	0	0	2	3	3	4.13 (.84)

Youth Outcomes

Site administrators were asked to report how youth were recruited. Youth were most frequently recruited through an announcement (reported by 5 out of 9 sites). One site each reported using the following techniques: personal invitations, an information setting, recruitment through other adults, or "Other." The "Other" strategy was not having to do much recruitment since the program was part of the school day.

The average team size per site ranged from 4 to 6 youth, with a mean of 5.50 (*SD* = .76). When asked whether they thought their team sizes were too big, too small, or just right, some site administrators (3 out of 8) thought 5-6 youth was just right, while others (5 out of 8) thought it would be ideal to have a maximum of 4 youth on each team. One site administrator commented that they thought 5-6 youth per team worked well because "they had the opportunity to disagree and reach a solution together. They complemented each other's abilities." Another site

administrator thought 5-6 "was too many students. The teams should have 4 members since not every student could have a role in the activities. What ends up happening are distractions to the team and the group in general." At all of the sites, youth worked with the same team over the course of the season.

Respondents were asked to report the extent to which participating in *FIRST* LEGO League Jr. helped youth develop in STEM attitudes, including confidence, persistence, and interest. On a scale from 1-5, site administrators reported the biggest development in STEM confidence (*mean* = 4.89, *SD* = .33). When asked to elaborate on their responses, one site administrator commented:

> *[FIRST LEGO League Jr.] opened a new world for them, where they were capable of understanding science topics, helping their self-esteem. And the confidence that they developed in being able to explain things that at first could be complex made them mature a lot. It was not easy work, questions were constantly used for students to investigate the answers. There was frustration at times, but by overcoming [that frustration] the students were empowered.*

Table 5. Mexican Site Administrator Ratings of Student Growth in STEM

	1 Not at all	2	3 Somewhat	4	5 A lot	Mean (*SD*)
More confident in STEM	0	0	0	1	8	4.89 (.33)
More persistent in STEM	0	0	0	3	6	4.67 (.50)
More interested in STEM	0	0	1	2	6	4.56 (.73)

Respondents were also asked to report the extent to which youth developed in STEM skills and 21st century skills. In regards to STEM skills, the 9 responding site administrators reported the greatest development in spatial construction and reasoning (*mean* = 4.78, *SD* = 0.67), and the smallest amount of development in programming and coding through LEGO software (*mean* = 4.11, *SD* = 1.36).

Table 6. Mexican Site Administrators' Ratings of STEM Skill Growth in Youth

	1 Not at all	2	3 Somewhat	4	5 A lot	Mean (*SD*)
Spatial construction and reasoning	0	0	1	0	8	4.78 (0.67)
Research skills	0	1	0	1	7	4.56 (1.01)
Understanding science content as it relates to the challenge topic	0	0	2	2	5	4.33 (0.87)
Knowing how to ask questions that will advance their understanding and knowledge	0	0	3	0	6	4.33 (1.00)
Programming and coding through LEGO software	1	0	1	2	5	4.11 (1.36)

In regards to 21st century skills, site administrators reported the largest change in youth's ability to explore and try out ideas (*mean* = 4.89, *SD* = 0.33) and fairly divide up work (*mean* = 4.89, *SD* = 0.33). Site administrators reported the smallest development in youth's ability to adapt ideas to solve new problems (*mean* = 4.22, *SD* = 0.83) and think through the steps of a problem (*mean* = 4.22, *SD* = 1.09).

Table 7. Mexican Site Administrators' Ratings of 21st Century Skill Growth in Youth

	1 Not at all	2	3 Somewhat	4	5 A lot	Mean (*SD*)
How to explore and try out ideas	0	0	0	1	8	4.89 (0.33)
How to fairly divide up work	0	0	0	1	8	4.89 (0.33)
How to explain one's own ideas to others	0	0	0	2	7	4.78 (0.44)
How to make decisions as a team	0	0	0	2	7	4.78 (0.44)
How to keep trying, even when things don't work out	0	0	1	0	8	4.78 (0.67)
How to listen to and understand others	0	0	1	1	7	4.67 (0.71)
How to come up with creative and original ideas	0	0	0	4	5	4.56 (0.53)
How to accept feedback or criticism	0	0	1	2	6	4.56 (0.73)
How to view problems from multiple perspectives	0	0	2	2	5	4.33 (0.87)
How to problem-solve in regards to technology	0	0	2	2	5	4.33 (0.87)
How to learn from failure	0	0	2	2	5	4.33 (0.87)
How to adapt ideas to solve new problems	0	0	2	3	4	4.22 (0.83)
How to think through the steps of a problem	0	1	1	2	5	4.22 (1.09)

When asked what they felt was the greatest benefit youth received from participating in *FIRST* LEGO League Jr., 4 out of 7 site administrators reported teamwork. Other responses included the balance between play and developing concrete skills and social skills, developing imagination and creativity, and the opportunity to learn new things. RobotiX staff echoed these thoughts, reporting feeling confident that if a student participates in *FIRST* LEGO League Jr. they will learn to work in a team and problem solve. Site administrators were also asked what *FIRST* LEGO League Jr. could do to improve youths' experiences. Of the 6 responding site administrators, 2 thought youth would have benefitted from more materials, another 2 thought

more time with programming and technology would have benefitted youth. Other site administrators thought coaches could have used more training and there could have been more categories for recognitions at the expo.

Program

Overall, site administrators reported high satisfaction with the *FIRST* LEGO League Jr. program. When asked to rate their overall satisfaction on a 10-point scale, their average response was 9.00 (*SD* = 0.93), and 3 of the 8 responding site administrators selected 10, or "Very Satisfied." Similarly, when asked to rate the quality of the *FIRST* LEGO League Jr. program on a 10-point scale, the average response was 9.13 (*SD* = 0.64), and 2 of the 8 responding site administrators selected 10, or "Very high quality." When asked to rate satisfaction with specific components of the *FIRST* LEGO League Jr. program, they were most satisfied with the year's challenge theme, "AQUA ADVENTURE" (*mean* = 4.67, SD = 0.71) and the training provided for site administrators (*mean* = 4.56, SD = 0.53). Site administrators reported being least satisfied with the webinar training for coaches (*mean* = 3.89, *SD* = 0.93). However, none of the respondents selected "Very dissatisfied" or "Dissatisfied" for any of the categories. During the interview with RobotiX staff, the Academic Advisor reported having extensive communication with coaches, including an initial training, weekly webinars, and weekly newsletters. Table 8 below shows site administrator ratings of satisfaction.

Table 8. Mexican Site Administrator Satisfaction with *FIRST* LEGO League Jr., from Most Satisfaction to Least Satisfaction

	1 Very Dissatisfied	2 Dissatisfied	3 Neutral	4 Satisfied	5 Very satisfied	Mean (*SD*)
The year's challenge theme, "AQUA ADVENTURE"	0	0	1	1	7	4.67 (.71)
Training provided by FLL Jr. (for site administrators)	0	0	0	4	5	4.56 (.53)
Program Administrator Guide (for your use)	0	0	1	4	4	4.33 (.71)
Engineering Notebook (for youth)	0	0	2	2	5	4.33 (.87)
Team Meeting Guide (for coaches)	0	0	2	2	4	4.25 (.89)
Other physical materials provided by FLL Jr. (e.g., LEGO Bricks)	0	0	1	5	3	4.22 (.67)
Availability of materials and resources in Spanish	0	0	3	1	5	4.22 (.97)
WeDo 2.0 Robotics Kits	0	0	2	4	3	4.11 (.78)
Webinar Training for coaches	0	0	4	2	3	3.89 (.93)

Site administrators selecting "N/A" were excluded itemwise. N's ranged from 8-9.

In addition to the elements described above, open-ended responses throughout the survey demonstrated a desire for more support in transporting youth and materials to the Expo. For example, one site administrator reported that it was a challenge to arrange the "transportation of models" and "get the budget for transportation to another state for the [Expo]." RobotiX staff also reflected on the challenge of getting team models to the Expo site. During the first year, they actively supported teams in transportation to Expo sites, and they witnessed a few models breaking during transit. With the program's expansion in the second year, RobotiX was no longer able to actively support team's transportation, and thus sites were left on their own to not

only manage the physical challenge of transporting their models, but the financial challenge of paying for this transportation.

Another common challenge reported by site administrators was related to WeDo 2.0. When asked what percentage of teams used the kits, all 9 of the responding site administrators reported that 100% of teams used the kits. Unfortunately, most sites (6 out of 9) reported challenges using the LEGO WeDo 2.0 robotics kits. In their open-ended responses, site administrators described issues with insufficient materials (2 out of 6), connection issues between the WeDo 2.0 and computer (1 out of 6), teaching youth to share the WeDo 2.0 (1 out of 6), and not being prepared to work with the WeDo 2.0 (2 out of 6) However, 2 respondents did not experience any challenges using the WeDo 2.0, and 7 were either "very satisfied" or "satisfied" with WeDo 2.0, so these challenges were not universal and for most site administrators, did not prevent the WeDo 2.0 from positively contributing to the program.

RobotiX staff were aware of the issues with Bluetooth and WiFi connectivity. They reported visiting sites prior to the start of the program in order to get an idea of the unique materials and capabilities at each site. However, it was still difficult to find solutions that would work for everyone because the sites had different resources, tablets, and computers. While RobotiX provided the support they could by sending along different components and occasionally visiting sites in person to support with troubleshooting throughout the program, they noted that 3-4 sites still had significant issues.

Besides concerns with transportation to the Expo and occasional issues with WeDo 2.0, open-ended feedback was overall quite positive. For example, when asked if they would like to share anything else about their experience, one site administrator spoke of the program's benefits for youth and their families.

> *These experiences are the ones that positively impact students and give them the tools to better themselves and create a new future for them and their families.*

Another site administrator reflected on the program's benefits for themselves.

> *...working with children is to my liking, they generate energy and boost my mood, and I learn a lot from them every day.*

Additionally, when asked if their site would host *FIRST* LEGO League Jr. again next year, all 8 responding site administrators reported they would.

RobotiX staff, too, had very positive things to say about the program. As previously mentioned, they commented that the materials used by *FIRST* LEGO League Jr. are particularly well suited for use with younger students. RobotiX staff also appreciated the emphasis on Core Values, noting "the program has the Core Values in the center and [everything else, like building and coding, goes along] with the Core Values." Finally, RobotiX staff reflected on how much the enjoyed being part of the program in Mexico, commenting "I love the program and we are excited for the next season!"

Recommendations

Support international sites in preparing for program implementation, particularly with regards to staffing needs. RobotiX staff mentioned they did not have a team in place that was prepared for *FIRST* LEGO League Jr. prior to program implementation. Additionally, because it was a new experience for them, it took RobotiX some time to delegate responsibilities among its staff and find a way to work together efficiently. *FIRST* can be extremely helpful in this regard, due to its extensive experience implementing programs with various sites. An overview of the types of tasks international partners need to be prepared for and a suggestion of staff roles could be helpful in informing program implementation prep for international partners.

Provide examples of trainings to site administrators and coaches. RobotiX staff also mentioned that it took them some trial and error to figure out how to provide training for their coaches and supports for their site administrators. Again, *FIRST* can be a resource for international partners as they work to prepare their site administrators and coaches for program implementation. To start, it would be helpful to provide a suggested schedule and/or list of topics to be covered during trainings. Even better, providing actual training materials, such as PowerPoint slides, handouts, videos, or activities, that could be translated into each country's local language, would be very helpful in supporting international partners in preparing their site administrators and coaches.

Provide tips or tools for moving completed models to Expo site. One seemingly minor task that brought unexpected challenges for both domestic and international sites was the transportation of models from local sites to Expos. In Mexico, site administrators and RobotiX staff both reported issues with moving models to the Expo site. Previously, RobotiX staff assisted sites in transportation to Expos, and even with their support, there were multiple issues with models breaking. As *FIRST* LEGO League Jr. expanded in Mexico, RobotiX had to take a step back from transportation, as coordinating these logistics for all sites would not be sustainable. However, if sites are expected to take care of the transportation themselves, there should be processes or resources in place to support them. It may be helpful for *FIRST* to provide potential strategies for transporting models smoothly, perhaps providing a list of tips in the *Team Meeting Guide*. Such a resource would also be helpful for domestic sites.

Consider technological limitations, and provide strategies for troubleshooting and/or adapting the program to fit within technological capacities. RobotiX staff and site administrators both reported difficulties with WiFi, Bluetooth connections, and other tech issues.

RobotiX staff put in a lot of work to minimize the disruptions caused by these issues, including providing sites with materials that might help, and visiting sites that needed further help with troubleshooting. To decrease the burden on international partners, it might be helpful to standardize the technology at different sites, and/or provide alternative activities for sites with limited connectivity. It might also be helpful to ask international partners to dedicate one day of training to troubleshooting their sites' tech ahead of time, rather than waiting for challenges to arise.

Work with organizations to develop a shared understanding of what UR/US means for each country. Because one of the goals of the Season Pass Model of *FIRST* LEGO League Jr. is to reach traditionally underserved and underrepresented populations, *FIRST* should consider working with international partners, before program implementation, to talk through what each party means by "underserved and underrepresented." These conversations will ensure that international partners are clear on how to focus their recruitment time and resources, and that *FIRST* is satisfied with the final results. In Mexico, even though RobotiX focused primarily on recruiting girls, only about 40% of the participating youth were girls. Perhaps through conversations with *FIRST*, additional target populations may have been identified.

Appendix B – Domestic Case Site Details – Site by Site

Over the course of the 2017-2018 school year, the Research & Impact Group visited the 12 Deep Dive sites to observe programming. The table below summarizes the 12 site visits.

Table 1. Summary of Twelve Case Study Site Visits

Site	Observer	Region	Type of Site	Observation Date	Number of Teams Observed	FLL Jr. Session Number	Timing
1	MC	South	School	12/13/17	5	2	During School
2	RN	West	School District	12/14/17	8	9	During School
3	RN	Midwest	Nonprofit	1/10/17	4	Varied	During & After
4	MC	Canada	Nonprofit	1/28/18	5	Varied	Weekend
5	MC	Midwest	School	2/9/18	6	8	During School
6	RN	Southeast	Nonprofit	2/16/18	5	2	During School
7	MC	Northeast	School	3/8/18	8	2	During School
8	MC	Northeast	Nonprofit /School	3/22/18	2	6	After school
9	RN	Midwest	School	4/16/18	11	2	During School
10	RN	Midwest	Nonprofit	4/21/18	3	Local expo	During & After
11	RN	Southeast	For-profit	5/8/18	13	Varied	During & After
12	MC	West	School District	5/29/18	10	5	During School

Analyses by Site

Across the 10 sites submitting surveys, average growth in Emerging Activation varied from 0.04 to 0.60 (Table 2). A repeated measures ANOVA, comparing changes from retrospective pre-scores to post-scores for all sites, found a significant time by site interaction for *FIRST* LEGO League Jr. participants (time x site $F(9,407 = 4.20, p = .00)$, with a moderate effect size ($\eta^2 = 0.09$). Follow-up analyses revealed that participants at Site 1 showed significantly

different patterns over time than those at 8 of the 9 other sites, reflecting both unusually low scores and unusually low growth.

Table 2. Emerging Activation Dimension Changes by Site

Site	Average Scores (min=1, max=5)		Change
	Pre Mean (SD)	Post Mean (SD)	
1	2.80 (0.56)	2.84 (0.82)	0.04
2	3.15 (0.51)	3.75 (0.50)	0.6
3	3.80 (0.49)	3.97 (0.55)	0.17
4	3.70 (0.48)	3.90 (0.48)	0.21
5	3.45 (0.58)	3.87 (0.66)	0.42
6	3.65 (0.59)	3.76 (0.69)	0.11
7	3.65 (0.59)	3.97 (0.51)	0.32
8	N/A	N/A	N/A
9	N/A	N/A	N/A
10	3.50 (0.81)	3.68 (0.86)	0.18
11	3.51 (0.61)	3.76 (0.60)	0.25
12	3.73 (0.64)	3.78 (0.57)	0.05

Across the 10 sites, average change in Emerging 21st Century Skills varied from -0.36 to 0.52 (Table 3). A repeated measures ANOVA, comparing changes from retrospective pre-scores to post-scores for all sites, found a significant time by site interaction (time x site $F(9,402 = 1.92$, $p = .048$) with a small effect size ($\eta^2 = 0.04$). Follow-up analyses revealed that participants at Site 1, again, showed different patterns over time that participants at 4 of the 9 other sites, reflecting their unusually low responses.

Table 3. 21st Century Dimension Changes by Site

Site	Average Scores (min=1, max=5)		Change
	Pre Mean (SE)	Post Mean (SE)	
1	2.95 (1.56)	3.41 (1.48)	0.46
2	4.42 (1.01)	4.06 (1.08)	-0.36
3	4.56 (0.85)	4.44 (0.85)	-0.12
4	4.24 (0.94)	4.48 (0.75)	0.24
5	3.89 (1.39)	4.19 (1.22)	0.3
6	3.52 (1.19)	4.04 (1.13)	0.52
7	4.25 (1.13)	4.67 (0.70)	0.42
8	N/A	N/A	N/A
9	N/A	N/A	N/A
10	4.44 (0.73)	4.67 (0.50)	0.23
11	3.67 (1.34)	3.92 (1.26)	0.25
12	3.71 (1.44)	3.98 (1.25)	0.27

In the following pages, we summarize each of the site visits, focusing on general site and team information, *FIRST* LEGO League Jr. program details, and overall reflections.

Site Visit #1

Site Visit #1 was conducted by Melissa Collins on December 13, 2017. The observation was conducted over two hours in a classroom during the school day. The coach, a computer science specialty teacher, was both the site administrator and only coach for this site. Teams at this site met once a week for 45 minutes and were single gender (because sessions occurred as part of Technology Classes, which were single gender). The teams observed included second and fourth grade females. The youth were working on building their Inspire Set PlayPumps with very heavy control by the teacher, who projected step-by-step instructions on the SMART Board and had girls assemble one piece at a time before passing the model to the next team member.

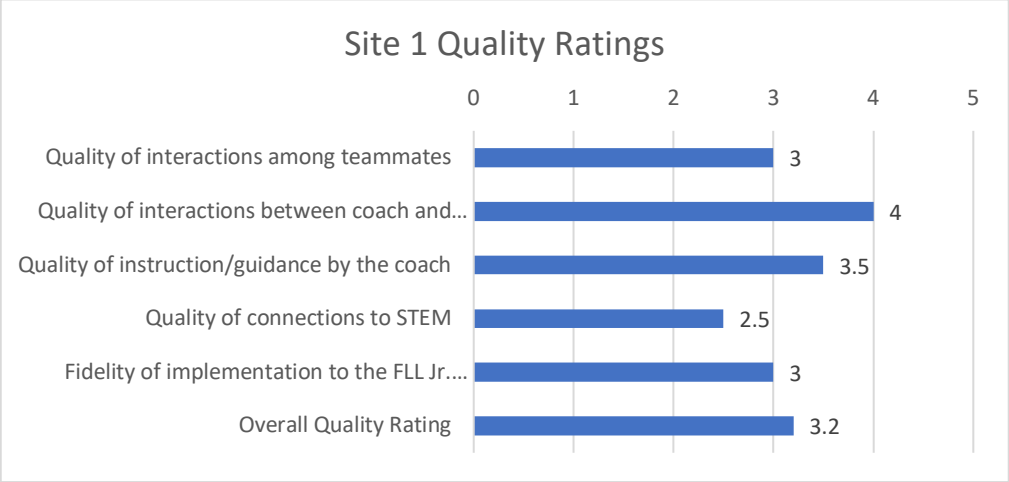
Overall, this site started and ended with the lowest Emerging Activation scores (10th out of 10 sites submitting student surveys), and showed the least change.

Table 4. Site 1 Emerging Activation Growth

Average Scores (min=1, max=5)		**Change**	**Comparison with other sites**		
Pre Mean (SD)[b]	**Post Mean (SD)[c]**		**Starting Rank**	**Ending Rank**	**Change Rank**
2.80 (0.56)	2.84 (0.82)	0.04	10th	10th	10th

Quality ratings were also quite low, the lowest of all 12 sites.

Figure 1. Site 1 Quality Ratings



Site Visit #2

Site Visit #2 was conducted by Rosalinda Nava on December 14, 2017. The observation was conducted over two hours during the school day. Teams at this site met 1-3 times a week for 30-60 minutes. The season pass at this site was held at the district level and, due to limited resources, there were 2 rounds of programming. Half of the first and second grade teachers had the LEGO materials until February, at which point they gave the materials to the second half of first and second grade teachers. Each classroom teacher was the only coach for the teams in their class. The first round of implementation consisted of 4 classes: 2 first grade classes, 1 class with a mix of first and second graders, and 1 second grade class. The teams were working on finishing

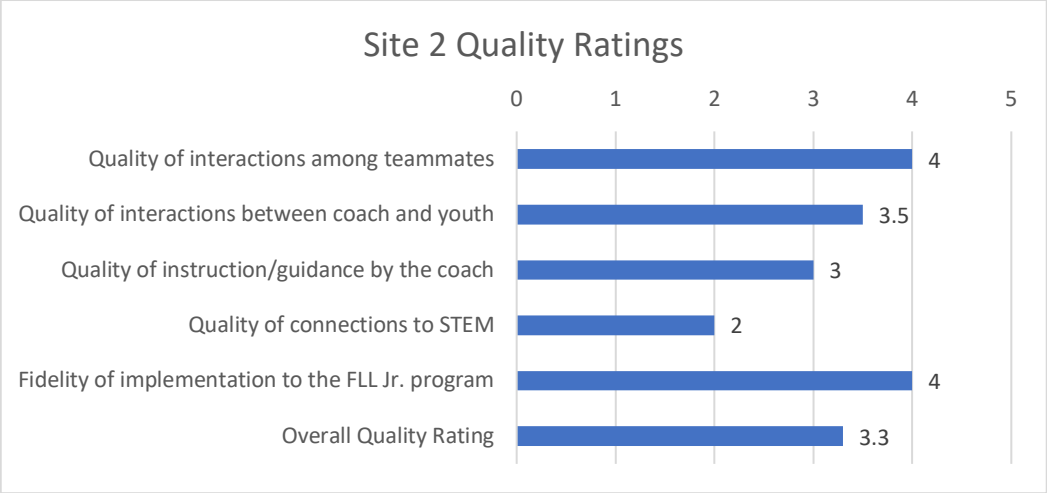
up their Team Models, with almost completely autonomy. The classroom teacher walked around to check on each group, but allowed them to troubleshoot problems and make decisions on their own. This site showed the largest growth from pre to post in Emerging Activation.

Table 5. Site 2 Emerging Activation Growth

Average Scores (min=1, max=5)		**Change**	**Comparison with other sites**		
Pre Mean (SD)[b]	**Post Mean (SD)**[c]		**Starting Rank**	**Ending Rank**	**Change Rank**
3.15 (0.51)	3.75 (0.50)	0.60	9th	8th	1st

Interestingly, though, quality ratings were quite low, scoring second lowest on average quality, perhaps indicating that the session observed was unusual, or that youth at this site benefited from other aspects of the experience not captured in the quality ratings.

Figure 2. Site 2 Quality Ratings



Site Visit #3

Site Visit #3 was conducted by Rosalinda Nava on January 10, 2018. Two sites were observed over the course of two hours, one during the school day, and a second during an after-school program. A nonprofit served as the site administrator for the sites in this region. The nonprofit had their educators go to different schools in the region to facilitate the program. At the

two observed sites, teams met 2 times a week for 1 hour. In each case, there were two educators serving as coaches, and two participating teams, so there was 1 coach per team. At the first site, the program took place during the school day, with participants selected based on interest and attendance. There were 4 second grade girls, and 5 third grade girls. The teams were split by grade level. At the second site, the program took place after-school, with participants self-selecting. There were a range of girls from first grade to third grade. The teams were randomly assigned. Most youth at the second site had previously been involved with the nonprofit organization and were very familiar with their coaches. During both observed sessions, the teams talked about the importance of having access to clean water, and watched a video about PlayPumps. Afterwards, the first site continued building their LEGO water pump, which they had started during a previous session. Teams at the second site decided on a water use, and started building their team model. In each case, the coaches were fairly hands-off once the youth started building with LEGO bricks.

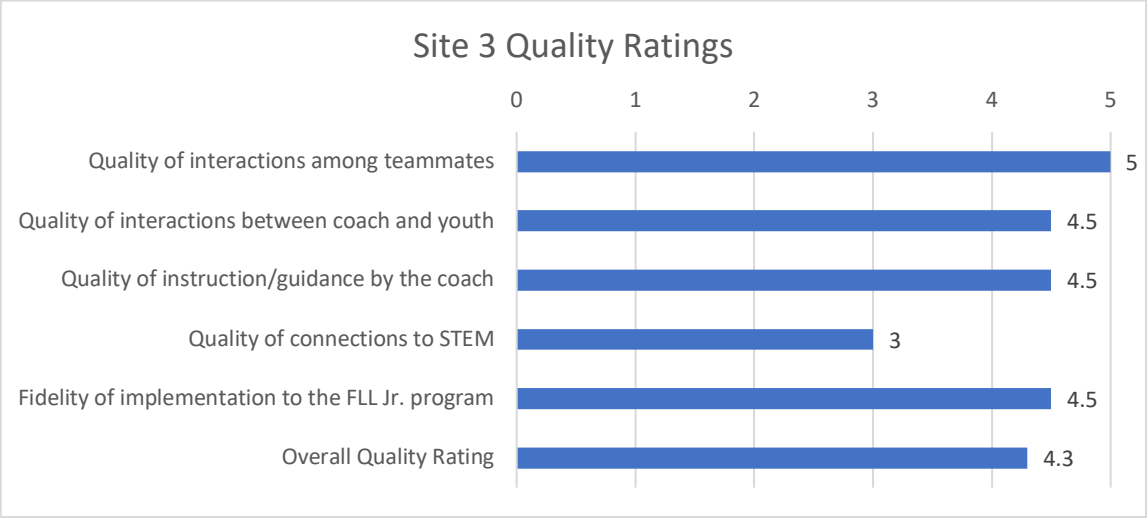
Overall, youth showed below average growth from pre to post (7th of 10), but this is likely because they started so high (1st).

Table 6. Site 3 Emerging Activation Growth

Average Scores (min=1, max=5)		**Change**	**Comparison with other sites**		
Pre Mean (SD)[b]	**Post Mean (SD)**[c]		**Starting Rank**	**Ending Rank**	**Change Rank**
3.80 (0.49)	3.97 (0.55)	0.17	1st	1st	7th

Quality ratings were quite high as well, scoring 2nd out of 11 for overall quality.

Figure 3. Site 3 Quality Ratings



Site Visit #4

Site Visit #4 was conducted by Melissa Collins on January 28, 2018. The observation was conducted over two hours in a community center on a Sunday afternoon. Parents served as coaches at this site, with one head coach who appeared to have a rich background in STEM. Each team had at least two coaches. Teams at this site meet once a week for two hours and ranged from kindergarten through fifth grade. Teams were split into classrooms by grade level: one classroom for kindergarteners, one for second-third graders, and one for fourth-fifth graders. The three groups were at different stages of the process. In the two older rooms, the youth were around Session 4 or 5. They were working building or adding to the models in various ways. The youngest group, in Room 2, was still naming their team and building their Inspire Set PlayPump. In general, activities were not rigidly following the Team Meeting Guide, as the site administrator had put together a lesson plan for coaches. Though the lesson plans referenced the Team Meeting Guide and Engineering Notebook, they also included their own content.

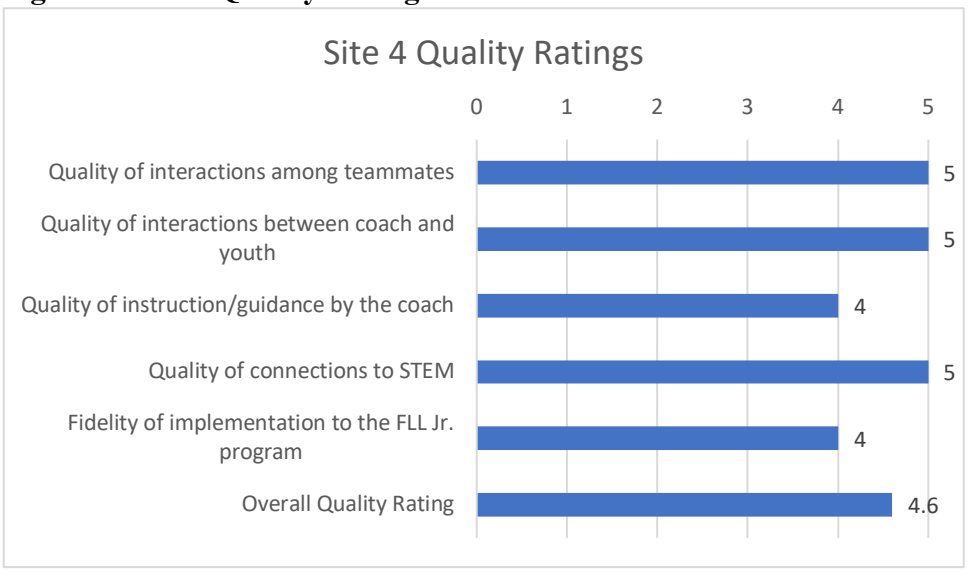
Overall, youth started and ended above average on Emerging Activation (3rd), and demonstrated average growth (5th)

Table 7. Site 4 Emerging Activation Growth

Average Scores (min=1, max=5)		Change	Comparison with other sites		
Pre Mean (SD)[b]	**Post Mean (SD)[c]**		**Starting Rank**	**Ending Rank**	**Change Rank**
3.70 (0.48)	3.90 (0.48)	0.21	3rd	3rd	5th

Quality ratings were high, with the second highest average rating across all sites.

Figure 4. Site 4 Quality Ratings



Site Visit #5

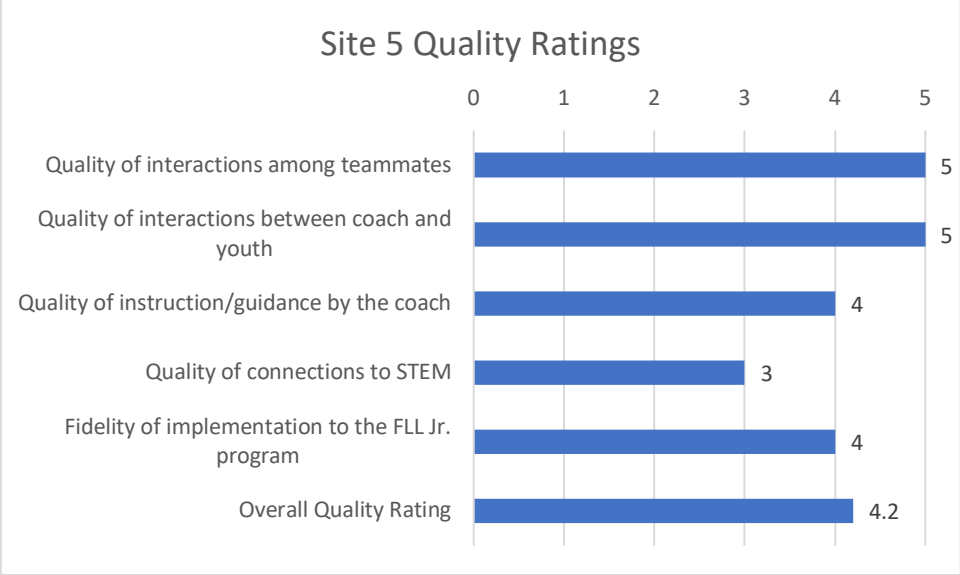
Site Visit #5 was conducted by Melissa Collins on February 9, 2018. The observation was conducted over two hours in two classrooms during the school day. Teams meet once a week for one hour. Local college students, studying education, served as coaches at this site. All kindergarteners at this school had the opportunity to participate as part of their science class. Each team had its own coach. Teams were around session 7 or 8, all adding to their team model in various ways. All groups seemed to be working at their own place, to an extent, and not closely following the sessions as laid out in the Team Meeting Guide or Engineering Notebooks. Overall, youth showed large gains in Emerging Activation (2nd highest growth).

Table 8. Site 5 Emerging Activation Growth

Average Scores (min=1, max=5)		Change	Comparison with other sites		
Pre Mean (SD)[b]	**Post Mean (SD)[c]**		**Starting Rank**	**Ending Rank**	**Change Rank**
3.45 (0.58)	3.87 (0.66)	0.42	8th	4th	2nd

Correspondingly, quality ratings were also quite high (4th out of 11).

Figure 5. Site 5 Quality Ratings



Site Visit #6

Site Visit #6 was conducted by Rosalinda Nava on February 16, 2018. The observation was conducted over one hour in a classroom during the school day. At this site, there was a STEM specialist who was the site administrator and lead coach at two different schools. In each case, she had 1-4 assistant coaches. At the first site, she had a classroom dedicated to *FIRST* LEGO League Jr. At the second site, she conducted the program from the library, though she had access to a separate room for storage. At each site, classroom teachers would bring their class to her once a week for an hour. The observed session consisted of 5 teams. The session started with a group discussion about the different uses of water. Afterwards, youth got into their teams and

read the PlayPump story. Next, the coaches asked the youth to build their own version of a PlayPump. Finally, youth started presenting their PlayPump to the class, though not all youth got a chance to do so.

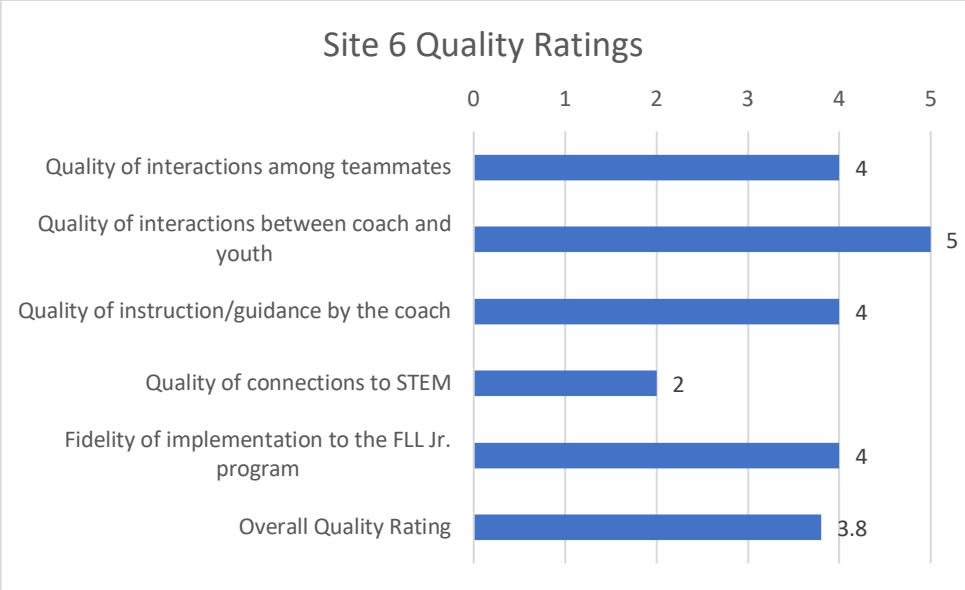
Overall, youth at this site were average in their Emerging Activation scores and showed below average growth (8th out of 10).

Table 9. Site 6 Emerging Activation Growth

Average Scores (min=1, max=5)		**Change**	**Comparison with other sites**		
Pre Mean (SD)[b]	**Post Mean (SD)[c]**		**Starting Rank**	**Ending Rank**	**Change Rank**
3.65 (0.60)	3.76 (0.69)	0.11	4th	6th	8th

The site also was rated as average in quality (6th out of 11).

Figure 6. Site 6 Quality Ratings



Site Visit #7

Site Visit #7 was conducted by Melissa Collins on March 8, 2018. The observation was conducted over two hours in one classroom during the school day. The STEM specialty teacher

served as both the site administrator and head coach, with classroom paraprofessionals and YMCA volunteers also serving as coaches. Teams met once a week for 45 minutes. Each team had its own coach or coaches, and they were on session 2 of the Engineering Notebook. Groups started as a whole group to recite the Core Values. Next, they watched a PlayPump video from YouTube, followed by drawing in their Engineering Notebooks. The session ended with 5-10 minutes of completely free building with LEGO bricks.

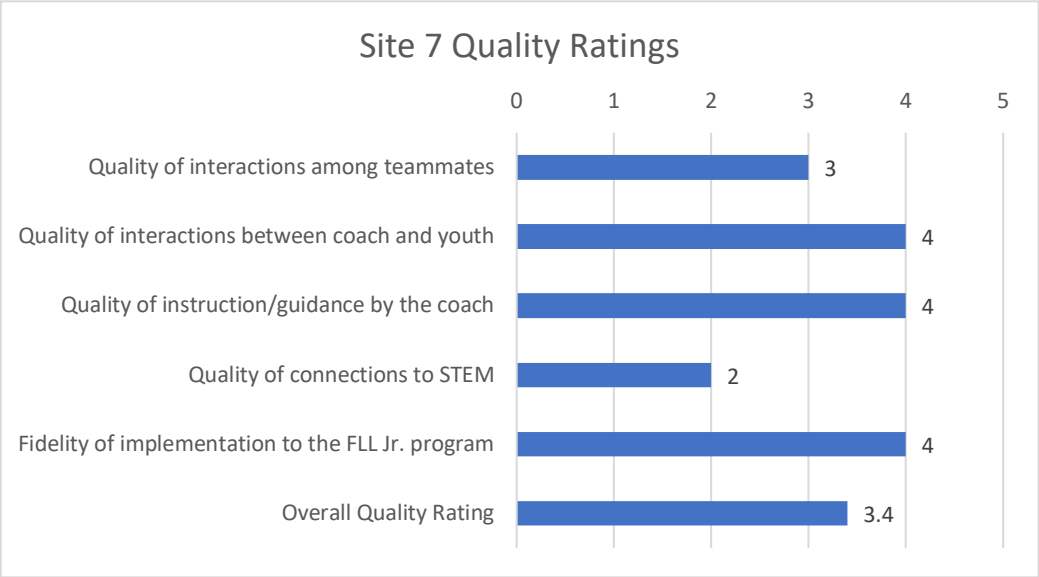
Overall, youth at this site started fairly high, showed above average growth, and finished with the highest scores on Emerging Activation.

Table 10. Site 7 Emerging Activation Growth

Average Scores (min=1, max=5)		**Change**	**Comparison with other sites**		
Pre Mean (SD)[b]	**Post Mean (SD)**[c]		**Starting Rank**	**Ending Rank**	**Change Rank**
3.65 (0.59)	3.97 (0.51)	0.32	4th	1st	3rd

Interestingly, however, similar to Site 2, quality ratings were quite low. In this case, the observation occurred extremely early in the season (Session 2), when much of the session was just watching a video about PlayPumps. It seems plausible that more learning may have occurred later in the season when sessions involved more hands-on building.

Figure 7. Site 7 Quality Ratings

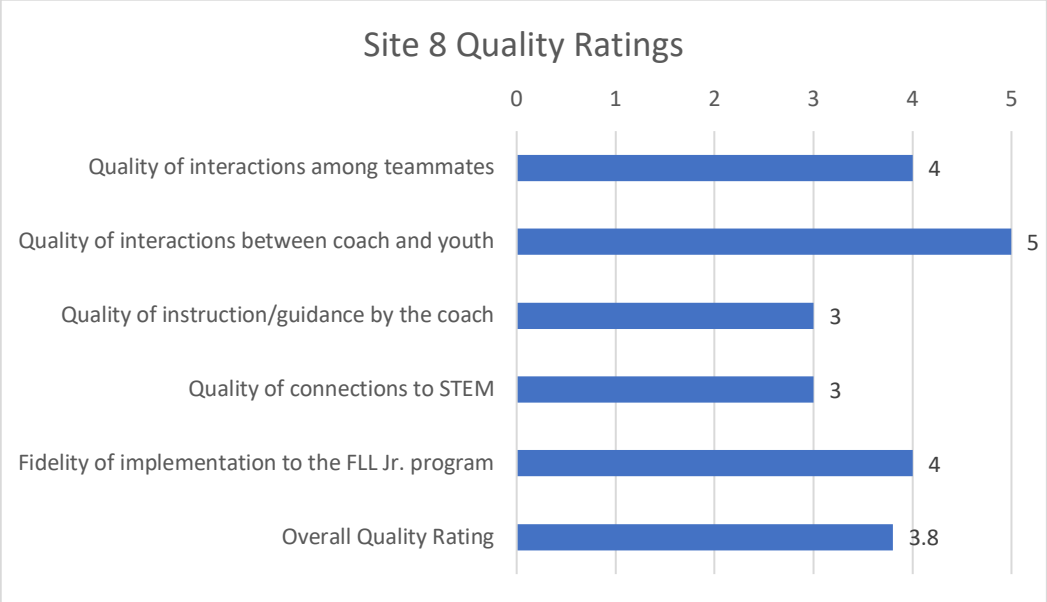


Site Visit #8

Site Visit #8 was conducted by Melissa Collins on March 22, 2018. The observation was conducted over an hour in one classroom after school. At this site, a nonprofit served as the site administrator and partnered with public schools to administer the program through a "seeding" model, in which the nonprofit provided initial support and then the schools were expected to continue the program on their own. The coach at this school was a fourth grade general education teacher, with her two daughters (high school students) also serving as coaches. Teams met once a week for an hour and fifteen minutes, including snack and dinner. The observation included two teams of third and fourth grade youth, who self-selected into single gender teams. They were on Session 6. Youth heard about what was coming up in Session 7 (designing their models) and viewed examples of models from last year's expo on the head coach's phone. Then they began to research their problems on Chromebooks. The research was extremely open-ended and student-driven, with coaches rarely checking in.

Student surveys were not submitted by this site. Quality ratings overall were average (6th out of 11).

Figure 8. Site 8 Quality Ratings

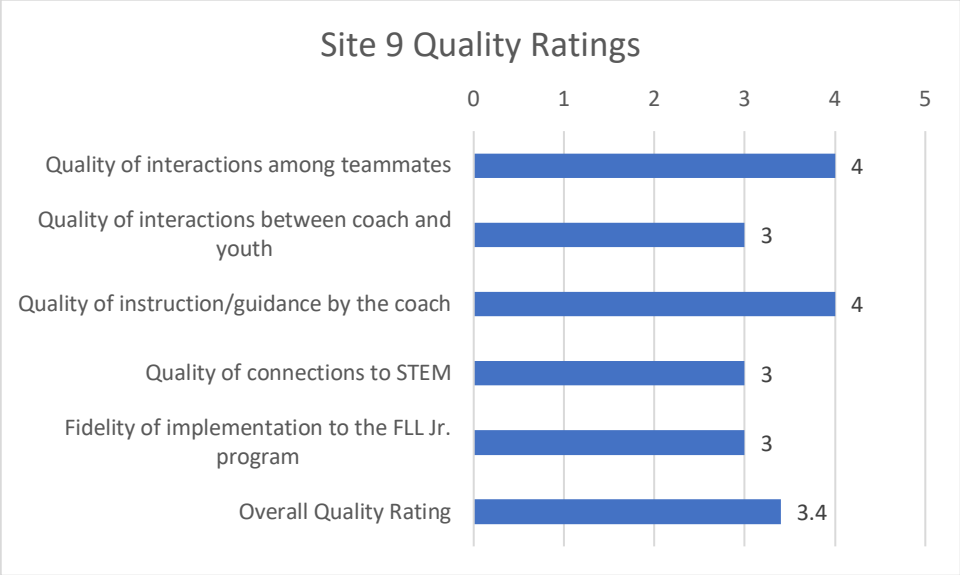


Site Visit #9

Site #9 was conducted by Rosalinda Nava on April 16, 2018. The observation was conducted over two hours with two different classes during the school day. At this site, the school's STEM specialist was the site administrator and lead coach for every team. She implemented the program once a week, for one hour, with every Art class in the school. The STEM specialist delivered the program independently for the most part, but received some support from the youth's Art teacher. The observed session was only the second day of the program. Youth talked about the different uses of water and discussed some learning goals as a group. Then, youth were randomly assigned to their teams and read the PlayPump story with a partner in their team. Afterwards, the teams built a LEGO model that represented something they could do with water. Finally, some teams were selected to present their work to the rest of the class.

Student surveys were not submitted by this site. Quality ratings were quite low (9th out of 11).

Figure 9. Site 9 Quality Ratings



Site Visit #10

Site Visit #10 was conducted by Rosalinda Nava on April 21, 2018. The observation was conducted over 2 hours at a science center. At this site, a nonprofit was the site administrator for the schools in the region. The nonprofit provided sites with the FIRST LEGO League Jr. materials, conducted a training for coaches, and provided some ongoing support when necessary. The observed session was of a mini expo that took place during the science center's grand opening. The expo took place in a section of a much larger room but there wasn't an effort made to invite visitors to listen to the presentations. There were 3 teams in attendance and each only presented their Team Models to the site administrator. The presentations took roughly 3 minutes, including some time for questions from the site administrator. After the brief presentations, the youth and their coaches explored the rest of the science center.

Youth at this site showed relatively low Emerging Activation (9th at pre and 8th at post)and average gains (6th out of 10).

Table 11. Site 10 Emerging Activation Growth

Average Scores (min=1, max=5)		**Change**	**Comparison with other sites**		
Pre Mean (SD)[b]	**Post Mean (SD)[c]**		**Starting Rank**	**Ending Rank**	**Change Rank**
3.50 (0.81)	3.68 (0.86)	0.18	7th	9th	6th

Because this visit observed an Expo rather than a session, a full observation protocol was not completed. Thus, quality ratings were not recorded.

Site Visit #11

Site Visit #11 was conducted by Rosalinda Nava on May 8, 2018. The observation was conducted over 3 hours in three classrooms, 2 during the school day and 1 after school. At this site, a nonprofit was the site administrator for the schools in the region. The nonprofit provided sites with the materials, and either implemented the program with their own educators (afterschool programs) or provided some training and co-taught the program with classroom teachers (during the school day programs). Teams met between 1-4 times a week for an hour. The observed teams that met during the school day were working on finishing up their Team Models and were encouraged to work as a team. In both cases, the classroom teacher was the only coach. The observed teams in the afterschool program were creating their Show Me poster. Each student was told which specific question to address on the poster. There were two coaches for this program, one was an educator from the nonprofit organization, and the other was a student teacher.

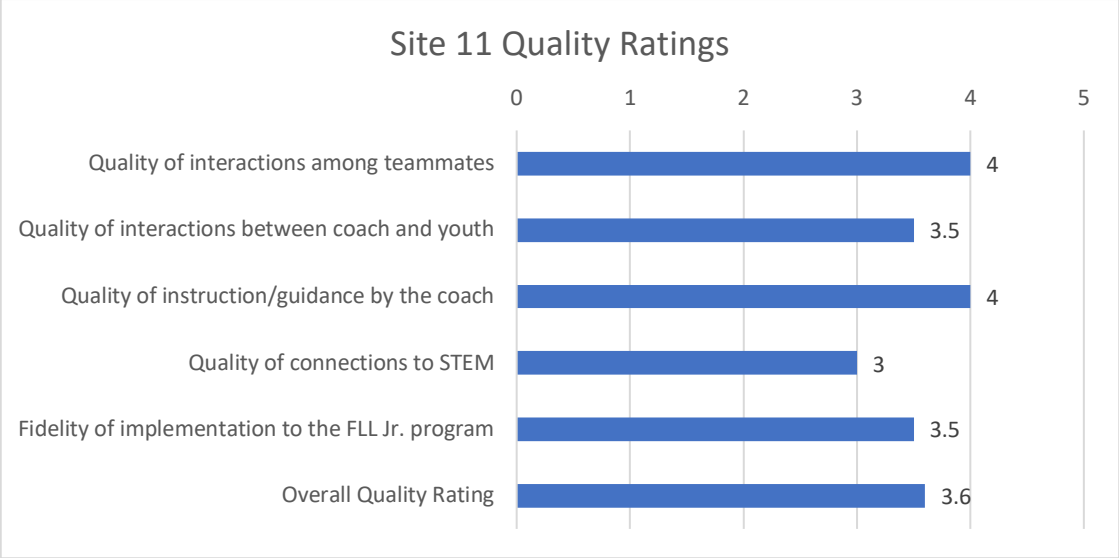
Youth at this site reported average Emerging Activation scores, but showed slightly above average growth (4th out of 10).

Table 12. Site 11 Emerging Activation Growth

Average Scores (min=1, max=5)		Change	Comparison with other sites		
Pre Mean (SD)[b]	**Post Mean (SD)[c]**		**Starting Rank**	**Ending Rank**	**Change Rank**
3.51 (0.61)	3.76 (0.60)	0.25	6th	6th	4th

Quality ratings were slightly below average (7th out of 11).

Figure 10. Site 11 Quality Ratings



Site Visit #12

Site Visit #12 was conducted by Melissa Collins on May 29, 2018. The observation was conducted over two hours in one classroom during the school day. At the observed school, the STEM specialty teacher served as coach. Teams met once a week for 45 minutes. Two classes were observed. Each class started by going through the Core Values aloud together, then they worked through Session 5, then they did a gallery walk to see each other's work. The season pass at this site was held at the district level, and kits were distributed to multiple schools within the district. Because of this arrangement, this school only received two kits originally. The coach was able to convince her district to purchase an additional three kits, so each group within each

class could have its own kit. However, at the end of each session, the groups needed to disassemble their designs so that the next class could build. Despite this constraint, this teacher was an excellent facilitator of STEM-rich learning that was a perfect balance of instructor scaffolding and student autonomy.

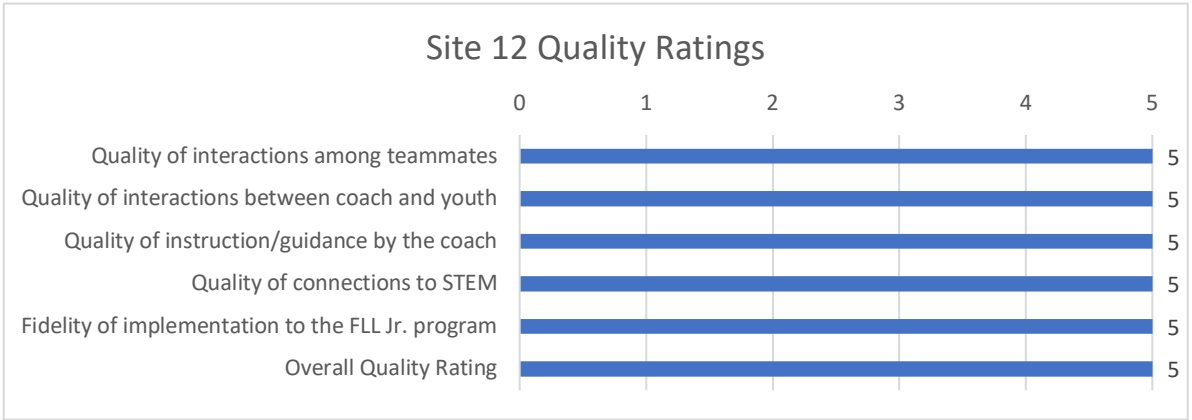
Youth at this site started out quite high on their Emerging Activation scores (2nd out of 10), but showed low growth (9th out of 10), reported average scores at post-test (5th out of 10).

Table 13. Site 12 Emerging Activation Growth

Average Scores (min=1, max=5)		**Change**	**Comparison with other sites**		
Pre Mean (SD)[b]	**Post Mean (SD)[c]**		**Starting Rank**	**Ending Rank**	**Change Rank**
3.73 (0.64)	3.78 (0.57)	0.05	2nd	5th	9th

These relatively small changes were surprising given the extremely high quality ratings given for the session observed. One potential explanation for this lack of change was the duration of the program. This site received the materials extremely late in the year (April), had only 45 minutes per week, and did not have enough materials for all the classes (teams had to break down models at the end of each session). Nevertheless, the coach's instruction and facilitation of the activities, within these constraints, was fantastic.

Figure 11. Site 12 Quality Ratings



Appendix C – Case Study Comparison – Two Domestic Sites

To explore potential factors influencing outcomes for youth participating in *FIRST* LEGO League Jr. under the Season Pass Model, we will be focusing on comparing two deep dive sites, Site 1 and Site 2. These sites were selected because they had the lowest growth and the highest growth, respectively, yet they were using similar implementation models. Additionally, their implementation model is fairly common among sites in general, so comparing them provides an opportunity to consider best practices for many sites. Site 2 and Site 1 were both implemented during the school day by classroom teachers who were the only coaches for their class(es). We will explore how these sites differed and potential reasons behind the variation in outcomes.

Basic Site Details

Site 1: At Site 1, one person was the site administrator and sole coach for *FIRST* LEGO League Jr. We therefore have responses from all of Site 1. While, we did not observe all classes at Site 1, we will be talking about them in general because we observed the sole coach implementing the program and can therefore speak to how the program was implemented across classes. Site 1 met once a week for one hour, for 15 weeks. Across 6 classes, 48 boys and 50 girls attended regularly; 31 from 2nd grade, 40 from 3rd grade, and 27 from 4th grade.

Site 2: At Site 2, the Interim Director of Career Technical Education was the site administrator. He recruited four schools within his school district to participate in *FIRST* LEGO League Jr. While we have some data from two schools, we will be discussing **one classroom in one school** in this case site write up. We have chosen to focus on the classroom we observed at Site 2 because we have the most complete data for them. Throughout this section, we will refer to this classroom as Site 2, to be consistent with the rest of the report. Please note, however, that

some values reported here will differ from those in other sections of the report, due to a change in sampling. Site 2 met 2-3 times per week for about an hour, for 14 weeks. The coach at this site noted that it took longer to do some of the sessions with her youth because they were 1st graders. 13 boys and 12 girls regularly attended the program at Site 2.

Table 1. Basic Site Details

	Site 1	Site 2
# Youth	98	25
# Classes	6	1
# Teams	19	4
Grades served	2nd/3rd/4th	1st
Meeting schedule	1 time per week, 1 hour, 15 weeks	2-3 times per week, 1 hour, 14 weeks

Table 2. Outcomes by Site

Emerging Activation	Site 1[a]	Site 2[b]
Pre score Mean (SD)	2.80 (0.56)	2.83 (0.45)
Post score Mean (SD)	2.84 (0.82)	3.82 (0.51)
Average change	0.04	0.99

a. n = 37
b. n = 17

Key Differences Observed

Student-led OR teacher-led

Site 1 ran a mostly teacher-led classroom. During our visit, we observed youth putting together their water pump. The instructor projected slides that drove all water pump assembly, and led the build in a very structured manner. She had youth go around in a circle and take turns with doing one of the steps in the assembly. The coach controlled the pace and checked team progress before proceeding to the next step.

Site 2 ran a very student-led classroom. During our visit, we observed youth working on their Inspire Sets by finishing their water pump build, coding and troubleshooting their water pump, and/or adding to the surrounding landscape. During observations, we noted that the coach told youth what they would be working on for the day, and left them to work on it. She did check

in on each group once or twice, but did not micromanage their build. The coach stayed at groups longer if there were behavior management or teamwork issues, but would otherwise leave groups to work on their own. This allowed youth to troubleshoot any issues that came up, from working as a team to difficulties building their water pump.

Mixed-gender classroom OR Gender-segregated classroom

Site 1 classrooms were gender-segregated. At this site, *FIRST* LEGO League Jr. was run during youth's computer class, which was split by girls and boys. Interestingly, boys and girls varied in their change on the Emerging Activation scale at this site. While both started in similar places (Boys: *mean* = 2.87, *SD* = 0.57; Girls: *mean* = 2.84, *SD* = 0.50), boys showed a decrease (*mean* = 2.46, *SD* = 0.80) in Emerging Activation, while girls showed an increase (*mean* = 3.25, *SD* = 0.70). This difference in change can be due to any number of reasons, including but not limited to: differences in previous STEM experiences, differences in the how the program was implemented, or participating in a gender-segregated class. Even though we don't know what caused this change with the current data, it raises interesting questions on whether single or mixed gender can have an impact on the success of the program.

FIRST LEGO League Jr. was implemented at Site 2 in a regular mixed-gender 1st grade classroom. Not only were boys and girls in the same classroom, but in most cases they were also on the same team. During our visit, we focused on observing 2 teams: one with 3 girls and 2 boys, and one with 3 girls and 3 boys. Both observed teams worked well together. The evaluator observed that the girls and boys had varying levels of engagement, though it didn't appear that any gender was on average more or less engaged than the other. Based on youth survey responses, boys showed greater increases than girls, but both genders grew a lot in terms of Emerging Activation (Table 3).

Table 3. Emerging Activation by Gender

	Site 1			Site 2		
	Pre	Post	Change	Pre	Post	Change
Boys	2.87 *(0.57)*	2.46 *(0.80)*	-0.41	2.76 *(0.42)*	3.95 *(0.28)*	1.19
Girls	2.84 *(0.50)*	3.25 *(0.70)*	0.41	2.89 *(0.50)*	3.70 *(0.65)*	0.82

Site 1 Boys: n = 18; Girls: n = 17
Site 2 Boys: n = 8; Girls: n = 9

Class details: grade, number of classes, classroom setting, youth's previous experiences

At Site 1, the program was implemented with 2nd, 3rd, and 4th graders, across 6 classes. The program was taught by youth's computer teacher and was held in a separate classroom. At Site 2, the program was implemented in one 1st grade classroom. The program was taught by youth's classroom teacher and took place in their regular classroom.

Grade level: The youth themselves were different ages. It is possible that the program is better equipped to support learning in younger kids. This fits with some data we have from youth: in terms of Emerging Activation, Kindergarteners and 1st graders were the highest both before and after participating in the program, 2nd and 3rd graders were in the middle, and 4th and 5th graders were the lowest. However, this seems to be in direct contrast with other data we have. In surveys, some coaches and site administrators reported that the AQUA ADVENTURE materials required more modifications for younger audiences.

Number of Classes: While the coach at Site 2 only had to implement the program with her class of 1st grade youth, the coach at Site 1 worked with 6 classes from 3 different grade levels. Based on feedback from other coaches and site administrators, we know that different grade levels required different scaffolding. This suggests that it could have been less of a burden for the coach at Site 2 to tailor *FIRST* LEGO League Jr. implementation to one grade level, as opposed to 3 grade levels like the coach at Site 1.

Classroom Setting: The program at Site 2 took place in youth's regular classroom, with their regular classroom teacher as their *FIRST* LEGO League Jr. coach. This means that the coach had the opportunity to be flexible with scheduling, make connections to *FIRST* LEGO League Jr. during class time (including when youth weren't doing *FIRST* LEGO League Jr. activities), and work with youth she knew well. The program at Site 1 took place in youth's computer classroom, with a STEM specialist as their coach. This would not allow the coach the same opportunities as the Site 2 coach, though it could mean that the Site 1 coach had helpful background knowledge to support her in implementing *FIRST* LEGO League Jr.

Youth's Previous Experience: Youth at Site 2 were all fairly new to building. None of them had ever participated in *FIRST* LEGO League Jr. before, and several hadn't worked much with LEGO bricks prior to participation in the program. Youth at Site 1, however, had previously participated in *FIRST* LEGO League Jr. The STEM specialist implements the program with 2nd-4th graders, and only administered the youth survey to 3rd graders. Therefore, we have survey responses from a group of youth who participated in the program last year as well. This could account for some of the lack of change we saw this year.

Coach training

The coach from Site 1 participated in *FIRST* LEGO League Jr. as a coach and site administrator. She did not have someone provide training for her, or a group of teachers to turn to for feedback and support.

The coach from Site 2 participated in *FIRST* LEGO League Jr. solely as a coach. She participated in the program along with other teachers from 4 schools. In terms of support, the site administrator held a training for all coaches prior to the start of implementation, and the Site 2

coach worked with 3 other teachers at her school to implement the program. While she was the only coach for her class, she had people to turn to for support or feedback.

Table 4. Summary of Program Differences

	Site 1	Site 2
Student-led or Instructor-led	Instructor-led	Student-led
Mixed-gender or Gender-segregated	Gender segregated	Mixed gender
Grade level	2nd/3rd/4th grade	1st grade
Classroom Setting	Computer Class	Regular Classroom
Number of Classes	6	1
Previous participation in *FIRST* LEGO League Jr.	Yes	No
Coach training	No	Yes
Average change in Emerging Activation	0.04[a]	0.99[b]

a. n = 37
b. n = 17

Overall, Site 1 and Site 2 differed in a number of ways, and it is likely that several of the differences contributed to the difference in student outcomes. It does seem that whether a site was more student-led or teacher-led had a particularly strong effect. This is supported as contributing to program success by other data sources. During observations, evaluators provided ratings for the balance between student-led activities and instructor-led activities. When we use this rating to separate sites into mostly student-led, mostly instructor-led, or a mix of student and instructor-led, we can see that there was more growth in Emerging Activation for youth at mostly student-led sites (Table 5). Additionally, student-led classrooms showed the greatest evidence of *learning through play*, according to observation data, particularly when compared with teacher-led classes. When looking at the components of learning through play, both mostly student-led and mixed classrooms scored higher across all *learning through play* components than mostly instructor-led classrooms (Table 6).

Table 5. Emerging Activation by Student-led vs. Instructor-led

	Pre score	Post score	Change
Mostly student-led [a]	3.37 *(0.59)*	3.79 *(0.60)*	0.42
Mostly instructor-led [b]	3.24 *(0.71)*	3.43 *(0.88)*	0.19
Mix of student and instructor-led [c]	3.62 *(0.60)*	3.80 *(0.58)*	0.18

a. n = 113 youth
b. n = 77 youth
c. n = 218 youth

Table 6. Learning through play by Student-led vs. Instructor-led[a]

	Mostly student-led [b]	Mostly instructor-led [c]	Mix of student and instructor-led [d]
Joyfulness	4.75 *(0.71)*	4.20 *(0.45)*	4.67 *(0.58)*
Meaningful play/participation	4.13 *(0.64)*	3.80 *(0.45)*	4.00 *(1.00)*
Iterative play/participation	4.00 *(0.93)*	3.20 *(0.84)*	4.33 *(1.16)*
Socially interactive play/participation	4.38 *(0.52)*	3.60 *(0.55)*	4.33 *(0.58)*
Teamwork and communication	4.25 *(0.46)*	3.60 *(0.89)*	4.33 *(0.58)*
Active engagement	4.75 *(0.46)*	4.60 *(0.55)*	4.33 *(0.58)*
Learning through play [e]	4.38 *(0.39)*	3.83 *(0.42)*	4.33 *(0.67)*

a. Scores are at the team level because observations focused on one team at a time
b. n = 8 teams
c. n = 5 teams
d. n = 3 teams
e. Overall learning through play was calculated by averaging the other 6 factors.

In observation notes, the evaluation team noted that youth at Site 2 "made observations and tested out different solutions" on their own. Meanwhile, at Site 1 "when one teammate placed a piece incorrectly, her teammates helped her figure out where to replace it. Also, the teacher helped the youth identify the problem." While it is important for coaches to provide support when needed, so as to prevent youth from getting discouraged, there is room for youth to problem-solve on their own.

While most data supporting the idea that student-led classrooms are more successful than instructor-led classrooms is anecdotal, this is something sites should consider in subsequent implementations.

Appendix D – *FIRST* LEGO League Jr. Season Pass Model Logic Model