SPECS for PLAID Program Evaluation [2017-2018]

STEPHEN J. BAGNATO, Ed.D., NCSP
Professor of Psychology & Pediatrics
Coordinator, Children with Special Healthcare Needs specialization
Faculty, UCLID-LEND Disabilities Center at Children’s Hospital of Pittsburgh of UPMC
Principal Investigator, SPECS for PLAID
University of Pittsburgh
bagnatos@pitt.edu

TRACY LARSON, M.Ed., NCSP
Nationally Certified School Psychologist
Director, SPECS for PLAID
Office of Child Development
University of Pittsburgh
Tlk7@pitt.edu

SPECS Team for PLAID@Pace:
Cara Bliss, PhD; Project Coordinator
Antonio Fevola, PhD, MSW; Data Analyst/Statistician
Molly Matsik, PhD; Evaluation Implementer

Program Evaluation Executive Summary and Full Report for Year 1 (2017-18)
August 9, 2018
**EXECUTIVE SUMMARY**

This executive summary highlights the baseline results (Year 1-2017-2018) collected, analyzed, and interpreted by the SPECS program evaluation team of the University of Pittsburgh for Pace School’s PLAID (Positive Learning and Integrative Design) initiative. Pace received funding from private foundations in Pittsburgh (e.g. Hillman, Benedum, Grable) to design a new, holistic, positive, and growth-oriented intervention model (PLAID) and engaged the SPECS Research Team from the University of Pittsburgh to conduct the program evaluation and provide ongoing feedback.

As Pace was engaged in developing and operationalizing the new PLAID model during the 2017-2018 school year, the SPECS team collected and analyzed baseline data from Pace teachers, team members, and students using authentic assessment measures of student capabilities and needs, effective classroom practices, and teacher and staff beliefs. These data are focused upon Pace program-planning and continuous quality improvement (CQI) as a stage-setter for implementation of the new holistic PLAID model. This baseline analysis targets an essential evaluation research question:

“What are the distinguishing characteristics of the status and progress of Pace teachers and students in the current Pace model over one school year?”

These “baseline” or representative results will be used as a real-time field-validation comparison in evaluating the new PLAID model as it is implemented in Year 2. The following bullets highlight the major “take-home” points from this first year analysis of the Pace school’s programmatic practices and impact on teachers, team members, students, and the school climate.

**Participants:**

- The SPECS team conducted this baseline study with a “random” convenience sample of Pace students and teachers from 6 classrooms. Classroom enrollment ranged from 8 to 14 students per room over the course of the school year. The final sample included a total of 61 students, 6 teachers, 8 classroom aides, 4 behavior management consultants, and 2 coaches.

- The mean chronological age of students was 11.8 years (SD = 2.44), with student ages ranging from 8 to 17 years. The mean developmental age of participating students was 9.5 years, or an average of 2.3 years lower than their chronological age. The majority of students (77%) were male, and virtually all (98.2%) were classified systematically...
through clinical judgment procedures as requiring moderate to high, scope and intensity levels, of programmatic support.

**Frequency and Scope of Pace Coaching:**

- Pace coaches recorded, electronically and daily, their mentoring and consultation activities with classroom staff during a baseline period of 36 school days between February and May 2018 (SPECS Mentoring Monitor).
- Coaches engaged teachers and staff in the 6 participating classrooms in 102 coaching/mentoring sessions (averaging 2.8 sessions per day).
- Sessions on average lasted 53 minutes, and most were conducted face-to-face.
- Coaching strategies used most frequently were observation, verbal feedback, face-face, and combinations.
- Social-behavioral topics were emphasized in 84% of coaching sessions, and sessions addressing this topic lasted the longest (average duration = 110 minutes).
- Most common content was crisis management, preventive strategies, and tactics for responding to disruptive and other undesirable behaviors.
- Coaches prioritized their time to address those classrooms with the highest levels of student need, particularly in the social-emotional-behavioral domain.
- Classrooms with lower instructional quality ratings (CEPI) received more coaching sessions and more total coaching time.

**Impact on Pace Staffs’ Mindset:**

- A total of 76 Pace staff members engaged in the initial pilot of the Positive Growth Mindset Training by completing the Growth Mindset and Self-Efficacy surveys during the baseline year.
- Pace staff’s Growth Mindset can be categorized as average with levels remaining stable from fall to spring. No differences in mindset were associated with the type of classroom in which staff practiced.
- Teachers tended to endorse a more positive growth-mindset while classroom support staff (e.g. assistants) and behavior management staff tended to endorse a more neutral and fixed mindset, respectively, over the brief review period.
- Self-efficacy beliefs among Pace staff remained relatively stable over the course of the baseline year, and can best be described as moderate (neither weak nor strong).
- Pace staff generally rated themselves higher in perceived Personal Self-Efficacy than in School Self-Efficacy.
- Staff who reported a stronger positive growth mindset also reported a greater sense of self-efficacy.
Impact on Classroom Instructional Practices:

- The SPECS team trained Pace coaches in the use of an authentic and systematic observational assessment of effective classroom practices (CEPI).
- Results showed that the team (i.e., teachers, aides, related services staff) increased their overall use of research-based effective classroom practices from fall to spring.
- Increases in effective classroom practices were observed in these specific domains:
  - High-quality instructional and social-behavioral support practices
  - Expectations and interactions to foster positive engagement and age-appropriate capabilities
  - Supports to promote social interactions
  - Student engagement in personal goal-planning
  - Skills that facilitate positive post-school outcomes.
- Independent observations conducted by SPECS team members at least weekly in all participating Pace classrooms (DBR Connect) revealed the following trends:
  - High or increasing rates of academic engagement in 4/6 classrooms
  - Low or decreasing rates of disruptive behavior in all 6 classrooms
  - High or increasing rates of respectful behavior in 5/6 classrooms
- Academic engagement tended to be higher in classrooms where hands-on activities and instruction were occurring.
- Higher levels of effective classroom practices were positively and significantly associated with higher rates of academic engagement and respectful behavior ($p \leq .05$) and lower rates of disruptive behavior.

Impact on Student Academic and Social-Behavioral Progress:

- The SPECS team collected data to explore the functional capabilities and needs of students using a variety of authentic assessment methods (FOCAL, SEARS, DHCS).
- Pace students show overall functional competencies and assets that are below levels typically displayed by their same-age/grade peers. The general functional status (i.e., classification level) generally remained relatively stable during the baseline year.
- Most students made positive, observable, and educationally significant progress towards developing age/grade appropriate functional competencies in the following areas:
  - Acquiring and using knowledge and skills
  - Taking appropriate, effective actions to meet own needs
  - Acquiring and using academic skills for classroom learning
  - Acquiring and applying computer-assisted technology skills
- Areas of functional competency that are particularly challenging for Pace students include:
  - Positive social-emotional skills and social assets
Self-regulatory behavior for classroom learning

- Students in classrooms where the use of effective, best practices was increased or maintained over the course of the school year made more overall progress in developing age/grade-appropriate functional competencies than did students in classrooms where the use of best practices showed a decline.
- Students in most classrooms maintained and/or slightly improved their social-emotional competencies from fall to spring.
- However, in the view of their teachers, the differences between students exhibiting the most challenges with social-emotional skills and peer interactions became more pronounced over the course of the school year.
- Students with social-emotional skills categorized as At-Risk or High-Risk have significantly lower developmental ages than students with average to high levels of social-emotional skills and competencies.
- Relationships and interactions among younger students and their teachers are more positive than between adolescent students and their teachers.

Focus Group:
- A focus group was conducted in order to obtain direct, qualitative feedback from individuals involved in the development and piloting of the new PLAID model. The group was moderated by a SPECS team researcher and consisted of 3 teachers and 2 administrators.
- The group identified several ways in which their involvement with the “soft” roll out of PLAID had positively impacted the professional practice of participating teachers.
- The examples provided support for expansion of the PLAID model, as participants expressed the belief that incorporating PLAID components across the school day will lead to improved generalization of positive social skills and reduce the need for discipline or correction.

Conclusions and Recommendations:
- These conclusions and recommendations are derived from collaborative discussions among SPECS/Pitt and the Pace Leadership Team and the data.
- The data collected in year 1 provides important baseline understanding of the status and progress of Pace students and teachers over one school year under the current Pace model. This baseline information will provide important points of comparison that will be critical to addressing the evaluation research questions as Pace expands and implements the PLAID model during Year 2.
- SPECS/Pitt and the Pace Leadership team have collaboratively identified next steps of activities and deliverables (detailed in the full report) that will focus the work of the SPECS for PLAID team and the Pace Leadership team in Year 2 and build upon the clear success of the Year 1 developmental phase and initial implementation of the program evaluation and continuous quality improvement (CQI) plan.
Introduction & Background

Pace approached and received funding from private foundations in Pittsburgh (e.g., Hillman, Benedum, Grable) to design a holistic intervention model, PLAID (Positive Learning and Integrated Design), which is innovative in integrating new strategies for instruction and whole-student development, especially social-emotional competencies; positive behavioral support; wellness; technology supports for applied learning skills (STEAM); parent engagement; and generalization of learning. Pace has engaged the SPECS Research Team from the University of Pittsburgh to design and conduct the program evaluation to provide ongoing feedback on the implementation of PLAID for program improvement and eventually to document the outcomes of a pilot study of the potential effectiveness of the new holistic PLAID model elements for Pace. The logic model or “theory of change” schematic below illustrates the hypothesized interrelationships among variables.

Exhibit 1. Logic model schematic for the expected “active agents” for change in the NEW Holistic model—PLAID:

Exhibit 1 depicts the logic or “theory of change” model which denotes the “active agents” that could be responsible for positive progress and the dynamics of change in teaching and student performance for PLAID, Pace’s new holistic model. In this logic model, coaching/mentoring is assumed to play, arguably, the major role in promoting teaching practices including instructional, classroom management, and student support as well as
ensuring the fidelity of implementation of the PLAID model. Coaching/mentoring would have a direct effect on teacher effectiveness and an indirect effect on student performance (through improved teaching quality). Improved teaching and instruction is presumed to have a direct effect on student performance and functional gains as well as the content, type, and intensity of tiered and graduated (i.e., RTI) services and supports.

**BASELINE YEAR PROCESS & PRODUCTS OF SPECS FOR PLAID**

**Process**

- The SPECS program evaluation team supported Pace by reviewing the research literature and consulting on the PLAID holistic intervention model components, especially as they impacted the evaluation.
- The SPECS team supported Pace in presenting PLAID model components, as impact evaluation, to Pace administration and staff for input and collaborative decision making.
- Pace created and implemented components of the organic holistic model, now termed **PLAID (Positive Learning And Integrated Design)**, independently from the SPECS team to ensure non-biased processes.
- The SPECS team, in collaboration with Pace administrators, presented the program evaluation model to the Pace Board for approval.
- The SPECS team engaged the Pace team in consensus decision-making in planning the logic model and evaluation questions for Pace’s PLAID model.
- The SPECS team identified and/or significantly adapted, and field-validated appropriate assessment tools to address the evaluation questions for Pace’s PLAID model.
- SPECS engaged and trained the Pace staff and administrators to become semi-independent in ongoing authentic assessment procedures for continuous quality improvement, program planning and program outcomes evaluation.

**Products**

- The SPECS team and Pace administration, through consensus decision-making, developed the logic model for Pace’s PLAID project (depicted in Exhibit 2).
**Exhibit 2: Logic Model**

<table>
<thead>
<tr>
<th>IMPACT</th>
<th>ACTIVITIES</th>
<th>ANTICIPATED OUTCOMES</th>
</tr>
</thead>
</table>
| Intervention model                 | • Staff work in professional learning communities (PLCs) to create unit plan structure that indicates cross-curricular connections and elements of PLAID, and to create resources for new staff/staff turnover  
• Unit plan structure drives lesson planning in pilot classrooms  
• Staff work in committees to address individual elements of PLAID | Instructional design framework (PLAID) will be developed, operationalized, and implemented |
| Staff practices and beliefs        | • Staff receive PD training on PLAID elements, best practices, growth mindset  
• Staff receive individualized and targeted coaching support | Staff consistently support student learning needs across instruction, settings, contexts  
Staff plan and deliver instruction that aligns with PLAID framework  
Staff report changes in beliefs relating to growth mindset and feelings of efficacy |
| Student functional and learning    | • Students receive services that align with and incorporate all elements of PLAID  
• Students receive high-quality instruction, classroom management, and support services that are informed by evidence-based best practices  
• Staff regularly use questioning and hands-on, product-based evaluation | Students demonstrate increases in academic, behavioral, functional, and social-emotional competencies that support improved post-school outcomes  
Students demonstrate increases in engagement and self-regulation, and decreases in aggressive behaviors |
| outcomes                           |                                                                           |                                                                                        |

- The SPECS team and the Pace administration collaboratively developed formative and summative evaluation questions. The evaluation questions focus on the functional interrelationships between teacher coaching/mentoring, quality of classroom instructional and management practices, and student performance and progress. The initial evaluation research questions for the SPECS for Pace program evaluation are outlined next:
- Does the PLAID model show greater overall effectiveness when compared to the current Pace model during a pilot evaluation period?

- Do students enrolled in classrooms implementing PLAID show significant differences in diverse domains of functional learning, social-behavioral competencies, and self-regulatory competencies?

- Do teachers in PLAID classrooms show significant improvements in the quality and content of their instructional and classroom management skills?

- Do teachers and support staff in PLAID classrooms show significant differences as compared to the current Pace model and compared to state and national normative data in overall instructional efficacy, including classroom quality, instructional formats, emotional responsiveness, and classroom organization?

- Is there a significant and predictive interrelationship or “path” among mentoring/coaching → teacher effectiveness → student participation, performance, and progress as a result of engagement in the holistic PLAID model?

- What are the characteristics of the status and progress of teachers and students in the current Pace model over one school year?

- The SPECS team and Pace administration selected appropriate assessment tools to address the evaluation questions. All evaluation tools/formative and summative assessment measures are authentic, functional, and collaborative. Selected tools are summarized in Exhibit 3.
### Exhibit 3. Evaluation Tools/ Authentic Assessments:

<table>
<thead>
<tr>
<th>Measure</th>
<th>Purpose</th>
<th>Description</th>
<th>Assessor</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classroom/Teaching</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classroom Effective Practices Inventory (CEPI)</td>
<td>Why?</td>
<td>Evidence-based classroom quality practices which promote engagement, normalization and progress</td>
<td>Observational assessment ratings of teaching quality</td>
<td>Pace SPECS</td>
</tr>
<tr>
<td><strong>Student</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional Outcomes Classification of Assets for Learners (FOCAL)</td>
<td>Why?</td>
<td>Functional skills classification of student status and progress</td>
<td>Observational assessment and consensus ratings by Pace team members</td>
<td>Pace</td>
</tr>
<tr>
<td>School Function Assessment (SFA)</td>
<td>Why?</td>
<td>Student status and progress on specific academic and self-regulatory curricular skills and support needs</td>
<td>Teacher and team member consensus observation of individual student curricular skills</td>
<td>Pace SPECS</td>
</tr>
<tr>
<td>SEARS-SF (Teacher, Child and Adolescent)</td>
<td>Why?</td>
<td>Authentic assessment of student social-emotional competencies and resiliency skills.</td>
<td>Student, teacher appraisals</td>
<td>Pace</td>
</tr>
<tr>
<td>Direct Behavior Rating (DBRConnect)</td>
<td>Why?</td>
<td>1-5 minute ratings of student prerequisite learning and self-regulatory behaviors</td>
<td>Observational computer ratings of behavior</td>
<td>Pace SPECS</td>
</tr>
<tr>
<td>Developmental Healthcare Complexity Scale (DHCS)</td>
<td>Why?</td>
<td>Ratings of individual students’ complexity of developmental and support needs to equate severity of disability</td>
<td>Consensus ratings based on extant records and expert judgment</td>
<td>Pace SPECS</td>
</tr>
<tr>
<td>Coaching/Mentoring</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPECS Mentoring Monitor</td>
<td>Why?</td>
<td>Ongoing profiling of the quality, content, scope, and intensity of coaching &amp; mentoring</td>
<td>Excel spreadsheet recording and analysis of coaching</td>
<td>Pace</td>
</tr>
</tbody>
</table>
- This report addresses the final evaluation question listed above (indicated with an asterisk, *). As Pace was engaged in developing and operationalizing the new PLAID model during the 2017-2018 school year, the SPECS team collected and analyzed data from Pace teachers and students using the selected assessment tools in order to establish a baseline understanding of status and progress under the current Pace model. This baseline will be essential in addressing the remaining evaluation questions as the new PLAID model is implemented, as it will provide a source for comparison.

- Results of this baseline evaluation are summarized in the following sections of this report, with recommendations for future evaluation efforts outlined at the conclusion.

**SPECS for Pace-PLAID: Baseline Data Summary (2017-2018)**

Data were collected throughout the 2017-2018 school year using the evaluation tools and authentic assessment measures outlined in Exhibit 3. Results and analyses are presented in the following sections of this report, organized by measure.

<table>
<thead>
<tr>
<th>Central Evaluation Question:</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are the characteristics of the status and progress of teachers and students in the current Pace model over one full school year?</td>
</tr>
</tbody>
</table>

**Participants**

The SPECS team conducted this baseline study with a “random” convenience sample of Pace students and teachers. The final sample included a total of 61 students enrolled in 6 classrooms. Classroom enrollment ranged from 8 to 14 students per room over the course of the school year. In Spring 2018, classroom size ranged from 8 to 12 students.

Classrooms were selected based on staff’s willingness and commitment to participate. Each classroom has 1 teacher; teachers are supported by classroom aide(s) and/or behavior management consultant(s), with the number of each varying by classroom. The total number of teachers in the sample is 6. The total number of classroom aides is 8, and the total number of behavior management consultants is 4. Additionally, the sample includes 2 coaches who work with staff and students in all 6 participating classrooms.

Of the 61 students included in the final baseline sample, 77% (n = 47) were male and 23% (n = 14) were female. The mean chronological age of students was 11.8 years (SD = 2.44), with student ages ranging from 8 to 17 years. The mean developmental age of participating students was 9.5 years, or an average of 2.3 years lower than their chronological age. (Note: Due to issues such as student attendance, measurement errors, etc., not every measurement tool was collected at each time point for all 61 participating students. Throughout this report, the numbers of students included in all analyses are reported accordingly.)
How Complex Are Students’ Functional and Programmatic Needs?

Developmental Healthcare Complexity Scale (DHCS)

The DHCS is a measure designed to enable research teams to quantify student capabilities and needs based on the concept of “complexity” rather than static diagnostic categories. It provides a systematic method of classifying the complexity of a student’s programmatic support needs in 4 domains:

- **Health:** Presence and extent of healthcare needs, level of impact on development, school adjustment, and/or family coping, and extent of support required
- **Behavior:** Presence and extent of behavioral dysfunction, impact on adjustment and coping, and extent of support required
- **Family:** Presence and extent of family problems, impact on child’s development and family’s ability to cope, and extent of support required
- **Developmental-Functional:** Presence and extent of developmental delays or functional limitations, and extent of support required

The DHCS consists of 8 items, each rated on a 4-point scale ranging from 0 (no needs and/or no support required) to 3 (severe problems exist/high levels of support required). The maximum possible sum score on the DHCS is 24, with higher scores indicating more complex programmatic support needs. Based on total sum score, the complexity of student needs can be classified as follows:

- **LOW:** Sum scores of 8 or less
- **MODERATE:** Sum scores of 9 – 16
- **HIGH:** Sum scores of 17 or greater

The DHCS was completed for each participating students in Fall 2018. Classroom teams used consensus decision-making procedures to assign ratings. This measure was administered at only one time point for classification purposes. Results indicated that:

- Pace students on average present moderate to high complexity of programmatic support needs, with an overall mean total complexity score of 13.98 (SD = 2.94; range = 8-23)
- 82% of students in this sample present moderate to severe levels of functional complexity
- Almost all (98.2%) of students require moderate to high levels of programmatic support
What Was the Scope, Content, and Intensity of the Pace Coaching of Classroom Staff?

SPECS Mentoring Monitor

As indicated in the Logic Model (Exhibits 1 and 2), a critical component of the PLAID model is coaching. The SPECS Mentoring Monitor was used to document the coaching/mentoring provided to staff in the 6 participating classrooms. This tool was created by SPECS team members, and adapted to reflect the work of coaches at Pace. Using this measure, Pace coaches logged the frequency and duration of coaching provided and team member(s) who received the coaching, as well as the following information for each interaction:

- **Communication Modes:** Amount of time spent in each mode
  - Face to face
  - Phone
  - Email
  - Written notes
  - Text messages

- **Strategies:** All strategies used during each coaching session
  - Observing
  - Demonstration/modeling
  - In vivo guidance
  - Formal workshop/training
  - Verbal feedback
  - Written feedback
  - Collecting resources
  - Collaboration on document creation
  - Pairing
  - Data review

- **Categories:** Topics emphasized during each coaching session
  - PLAID Components (wellness, knowledge, character, community membership, growth mindset)
  - Instructional Strategies (modifying curriculum/goals/tests, functional classroom routines, presentation methods, other strategies)
  - Environmental Modifications (sensory adjustments, physical arrangements, technology)
  - Social Behavioral Components (social skills instruction, preventative strategies, crisis management, whole classroom management strategies, strategies for responding to undesired behavior, CHAMPS)

Following collaborative adaptation of this measure, Pace coaches were trained in its use for data collection. To obtain baseline information about the coaching provided to pilot classrooms, Pace coaches recorded their activities using this tool over a span of 36 school days between February and May 2018. During this time, coaches provided services to staff and students in all Pace classrooms; however, for the purpose of baseline data collection, only coaching provided in the 6 participating classrooms was reported and analyzed.
Baseline Coaching/Mentoring Log Results

During the 36 school day data collection period, Pace coaches engaged teachers and staff in the 6 participating classrooms in a total of 102 coaching/mentoring sessions. On average, coaches completed about 2.8 sessions per day and 9 sessions per month in the 6 PLAID classrooms. The duration of coaching sessions ranged from a minimum of 10 minutes to a maximum of 180 minutes, with an average duration of 53 minutes. The vast majority of coaching sessions were conducted via face-to-face contact with the classroom teacher, staff, and/or the entire classroom team.

The strategies used most frequently by coaches included observation, providing verbal feedback, and pairing. Exhibit 24 shows the frequency distribution of the various strategies used during coaching sessions.

*Exhibit 24. Frequency of Coaching Strategies (Total Sessions and Percentage of Sessions Used)*

The category of coaching, or topic most frequently addressed during coaching sessions, was social-behavioral. Social-behavioral topics were emphasized in 84% of coaching sessions. During this baseline year PLAID components were not targeted; changes in this area in particular will be useful to monitor in future years as PLAID is implemented school-wide.
Exhibit 25 summarizes the frequency with which each major category was emphasized during coaching sessions, and provides both frequency counts (total number of sessions) and percentages (percent of total sessions) for each topic.

**Exhibit 25. Frequency of Categories Emphasized during Coaching Sessions.**

Among the 3 topics emphasized during coaching sessions, sessions addressing social-behavioral components lasted the longest with an average duration of 110 minutes (or almost 2 hours). The duration of these sessions ranged from an average minimum of 36 minutes to an average maximum of 90 minutes.

While environmental modifications were emphasized in only 5 of the total coaching sessions recorded, these sessions on average lasted 1 hour (61 minutes) in duration. Most (n = 4/5 sessions) involved coaching on modifying the physical arrangements of the classroom (furniture, seating, etc.) to maximize student success. The majority of sessions emphasizing instructional strategies (n = 9/14 sessions) focused on functional classroom routines.

Exhibit 26 summarizes the average duration (in minutes) of coaching sessions devoted to each major topic area, as well as the average minimum and average maximum length of sessions in each area.
Exhibit 26. Average Duration and Duration Range of Coaching Sessions by Topic Emphasized.

With 84% of coaching sessions emphasizing social-behavioral components, a more in-depth look at this category is warranted. Within this overarching category, coaches recorded 6 specific topics addressed in their mentoring session. The most common of these specific social-behavioral components were crisis management, preventative strategies, and strategies for responding to undesired behavior. Combined, these 3 social-behavioral components accounted for about 80% of all coaching/mentoring provided to Pace teachers and staff. Exhibit 27 provides a summary of the frequency with which each specific social-behavioral component was emphasized during coaching sessions.
Most coaching sessions were conducted with team members other than the teacher (n = 42), or with teachers and staff together (n = 38). Coaching sessions involving both teachers and staff tended to last the longest, with an average duration of 91 minutes (ranging from an average minimum of 35 minutes to an average maximum of 143 minutes). The mode, or most frequently reported average duration of these sessions was 43 minutes. Exhibit 28 summarizes the total numbers of coaching sessions conducted with various staff configurations, and data pertaining to duration of each type of session.
Exhibit 28. Total Number of Sessions, Average Duration, Duration Range (Minimum – Maximum) and Mode (Most Frequently-Reported Duration) for Coaching Sessions by School Personnel Involved.

During the baseline data collection period, Pace coaches reported 98 sessions specific to one of the 6 participating classrooms. Coaching efforts were not evenly distributed across classrooms with 2 classrooms accounting for 66% of all coaching sessions, as can be seen in Exhibit 29.
Appendix D of this report contains a detailed summary of differences in coaching duration, strategies, and topics between classrooms.

**Coaching Log Data in Context**

Considering the baseline coaching log data in the context of other student and classroom level data provides insight regarding variations in coaching duration, strategy, and focus. As summarized above and in Appendix D, Classrooms 2, 5, and 6 accounted for a significant proportion of total coaching time with most coaching emphasizing social-behavioral topics. This finding is more readily understood in conjunction with FOCAL data. Analyses of this data demonstrated that students in Classrooms 2, 5, and 6 displayed more challenges with controlling their own behavior, following classroom rules, and demonstrating age/grade appropriate functioning than did students in other classrooms.

SEARS data further explains coaches’ decisions to allot relatively more coaching time to Classrooms 4, 5, and 6. SEARS-T analyses indicated that a much higher proportion of students in these classrooms were seen as being At-Risk or High-Risk in terms of deficits in social-emotional skills and competencies.
Taken together, these results suggest that coaches targeted their time to address those classrooms with the highest levels of student need, specifically in the social-emotional-behavioral domain.

Additionally, considering coaching log data in conjunction with CEPI results, coaches’ decisions to allot relatively more time to Classrooms 4, 5, and 6 would seem to be supported and explained. These classrooms generally earned lower scores on this observational measure of effective practices, and exhibited a trend of declining scores from fall to spring. This would suggest a greater need for coaching/mentoring support in order to help staff in these classrooms exhibit and maintain effective classroom practices. Correlation analyses supported this possibility, showing a negative relation between classroom CEPI performance and total amount of coaching time \( (r = -0.18, p = 0.08) \), total coaching strategies used \( (r = -0.15) \), and total coaching sessions conducted \( (r = -0.13) \).

That is, classrooms with lower CEPI scores received higher numbers of coaching sessions and more total coaching time.

The coaching and mentoring support received by participating staff seems to have helped in maintaining or improving the use of effective classroom practices, as reflected in the overall increase in CEPI Total scores from fall to spring. Moreover, 5 of the 6 participating classrooms increased their scores in the CEPI Social Relationships domain, suggesting that the social-behavioral emphasis of most coaching sessions may have influenced these practices.

Additionally, the benefits of coaching and mentoring support are reflected in the small but positive progress made by students. FOCAL Progress results demonstrated that most students made positive and observable progress towards developing age/grade appropriate functional competencies in the Knowledge, Effective Actions, and Academics domains. The indirect benefit of coaching support provided to staff on student outcomes is further supported by the DBR Connect results, which demonstrated higher or increasing levels of Academic Engagement and Respectful Behavior and low or decreasing levels of Disruptive Behavior in most classrooms.

The benefits of coaching and mentoring on classroom practices, and the indirect effects of coaching on student outcomes, will need to be further studied and explored in Year 2. Nonetheless, these baseline exploratory analyses do seem to suggest that coaching and mentoring are having a positive influence on both classroom staff and students.

**What Was the Overall Impact of the Pilot Growth Mindset Training?**

**Staff Beliefs: Growth Mindset and Self-Efficacy**

As indicated in the Logic Model (Exhibits 1 and 2), a critical component of the PLAID model is its impact on staff beliefs. Staff beliefs, in turn, are thought to affect staff practices. Of particular relevance to PLAID are staff beliefs pertaining to growth mindset and self-efficacy;
therefore, establishing a baseline understanding of these beliefs is important. Baseline data in each area was collected using self-report survey instruments administered to staff electronically at two time points: October (Fall) 2017 and June (Spring) 2018.

Unlike most other measures summarized in this report, staff beliefs measures were issued to ALL Pace staff. In total, 76 staff members completed these surveys during the baseline period. Of this total, about 26% (n = 20) completed both the fall and spring surveys; the remaining 74% (n = 56) completed only the Fall 2017 (n = 35) or Spring 2018 (n = 21) survey.

The majority of staff beliefs survey participants (30%, n = 23) reported the position title of Classroom/Support Teacher; of this group, 26% (n = 6) completed surveys at both baseline time points. Exhibit 30 provides a summary of all respondents by position title, and indicates the number of respondents for each position title who completed surveys at 1 or 2 of the baseline time points.

Exhibit 30. Staff Belief Survey Sample Totals by Position Title and Number of Baselines Completed (N = 76).
Participants also identified the type of classroom in which they practiced. 25% (n = 19) of participants reported practicing in Partial Hospitalization (PHP) classrooms; 33% (n = 25) in Special Education (Therapeutic) classrooms; and 42% (n = 32) in “Other” classrooms. Of the 20 participants who completed both baseline surveys, 20% (n = 4) were from PHP classrooms; 25% (n = 5) were from Therapeutic classrooms; and 55% (n = 11) were from “Other” classrooms.

Growth Mindset

The concept of growth mindset refers to a person’s implicit beliefs about basic traits and abilities. According to the literature on this subject, these beliefs tend to vary along a mindset continuum ranging from “fixed” to “growth”-oriented. Those with a fixed mindset believe that a person’s basic traits and qualities are fixed and unchangeable, whereas those with a growth mindset believe that traits and qualities can be changed (or grow) with effort.

In an educational context, literature suggests that teachers/educators with a fixed mindset tend to believe that they cannot do much to improve students’ basic capacity to acquire and use knowledge. These educators tend to believe that praising or rewarding students for performance is the best strategy for building students’ confidence, motivation, and/or interest in learning. Educators with a growth mindset, on the other hand, tend to believe that with the right tools and dedication they can not only impact students’ desire to learn but also transform a student’s basic capacity to acquire and use knowledge and skills. With a growth mindset, errors become positive opportunities for learning and growth. These two opposing ends of the mindset continuum may in turn exert differing influences on educators’ sense of self-efficacy and on students’ beliefs about their own capabilities.

The Others Form for Adults Growth Mindset Questionnaire (Dweck, 1999) was used to assess staff beliefs, or implicit theories, in the following areas:

- **Intelligence**: Items include “People have a certain amount of intelligence, and they can’t really do much to change it” and “People can change even their basic intelligence levels considerably.”
- **Personality**: Items include “Someone’s personality is a part of them that they can’t change very much” and “Anybody can change their personality a lot.”
- **Kind of Person**: Items include “People can do things differently, but the important parts of who they are can’t be changed” and “People can always substantially change the kind of person they are.”
- **Morality**: Items include “There is not much that can be done to change a person’s moral traits such as conscientiousness, uprightness, and honesty.”
- **World View**: Items include “Though we can change some phenomena, it is unlikely that we can alter the core dispositions of our world.”
Pace staff self-reported beliefs in each area using a 6-point rating scale ranging “Strongly Agree” to “Strongly Disagree”. Average ratings reflect where along the mindset continuum an individual’s beliefs fall. For the purposes of this report, the following 3 categories were used:

- **Stable/Fixed Mindset**: Average rating scores below 3.98 reflect a tendency to view personal characteristics and traits as fixed and unchangeable, and thus are categorized as a stable mindset.
- **Average/Neutral Mindset**: Average rating scores of 3.99 – 4.99 reflect views within the average or middle range of the mindset continuum.
- **Growth Mindset**: Average rating scores of 5.0 and higher reflect a tendency to view personal characteristics and traits as malleable and believe that basic traits and abilities can be modified with effort.

**Growth Mindset Baseline Results**

The Fall 2017 Growth Mindset survey (GMS) was completed by 55 staff members. The Spring 2018 Growth Mindset survey was completed by 41 staff members. A total of 20 respondents completed this survey at both time points. Mean scores, standard deviations, and skewness of response distributions at both time points were virtually identical (Fall 2017 Mean = 4.50, SD = 0.78; Spring 2018 Mean = 4.48, SD = 0.75). Therefore, fall and spring results were combined and averaged to provide an overall picture of the baseline results. A paired-samples comparison indicated that these combined, or composite, scores provide a reliable overall baseline assessment of staff beliefs. This analysis is summarized in Exhibit 31. As can be seen in this figure, the distribution of Growth Mindset scores among Pace staff follows a roughly normal distribution, with about half (53%) of scores falling below the mean and half (47%) falling above the mean.
Exhibit 31. *Histogram of Growth Mindset Score Distribution for Total Baseline Sample (Fall and Spring; N = 76).*

Exhibit 32 summarizes this overall baseline data in a different way, by indicating the percentage of total participants whose average rating scores place them in each of the previously defined mindset categories. Parenthetical values indicate the mean rating score for each category.
Results of the Pace Growth Mindset survey are very similar to those reported by Dweck and colleagues (1995; 2006), who indicated that on average about 80% of respondents tend to report a growth-oriented mindset. About 82% of the Pace sample had mean rating responses of 4 or higher, which Dweck et al. (1995) are indicative of an “incremental” growth mindset while scores of 3 or lower are typically indicative of a fixed or stable mindset. Dweck (1995) further recommends that average responses between 3 and 4 be considered as a separate group and classified as having a neutral mindset. 26% of the Pace sample falls into this category, which is slightly higher than average percentages reported in the literature (12% - 15% of all respondents). The mean for the Pace sample (Mean = 4.48, SD range = 0.84 – 1.01) is also very similar to average ratings reported in the literature (Mean = 4.62, SD range = 0.63 – 1.14). Based on these comparisons between the Pace sample and published research regarding Growth Mindset, the following average baseline thresholds and classifications are recommended for future analyses:

- Stable/Fixed Mindset: 3.49 and lower
- Neutral Mindset: 3.5 to 4.29
- Growth Mindset: 4.3 and higher

The mean scores at both Fall 2017 and Spring 2018 baseline assessments in each of the 5 implicit theory scales measured by the Growth Mindset measure are presented in Exhibit 33. As these comparisons show, the differences between fall and spring scores in all areas are minimal and insignificant. Staff beliefs in each area remained at similar levels across the
baseline year. Additional information about specific results for each scale can be found in Appendix E of this report.

Exhibit 33. Fall 2017 and Spring 2018 Mean Scores by Growth Mindset Scales.

Mean scores at each baseline time point, and the total mean reflecting averaged fall and spring scores, were compared across the 3 types of classrooms represented. Analyses found no statistically significant differences in scores across classroom types. Staff in all 3 classroom types had mean scores ranging from 4.2 to 4.8, with very little within-group variation. This indicates that there are no significant differences in implicit beliefs or mindset associated with the type of classroom in which staff practice. These comparisons are presented in Exhibit 34.
Mean scores on the overall Growth Mindset survey and in each of the 5 implicit theory scales were also compared by staff role. For this analysis, all staff positions were grouped into 3 main categories: Teachers (including those in therapeutic, PHP, or special area roles); Assistants (all staff in support roles, including coaches, therapists, etc.); and Behavior Management staff. Those in supervisory roles and personal care assistants were excluded from this analysis. Analyses found that the observed differences between these groups were marked and statistically significant in 2 implicit theory scales: Personality ($F = 3.607$, $p \leq .05$) and Kind of Person ($F = 3.136$, $p \leq .05$). In each of these areas, teachers reported relatively higher average ratings than staff in support or behavior management roles.

- In general terms, these results suggest that teachers tend more towards a growth mindset while classroom support staff (e.g. assistants) and behavior management staff tend more towards neutral and stable mindsets, respectively. The mean scores in each area for these major staff position categories are presented in Exhibit 35.
Self-Efficacy

Self-efficacy refers to beliefs regarding one’s ability to enact change or to be effective/successful with various tasks. In an educational context, self-efficacy refers to educators’ self-evaluative judgements of their capability to bring about change and positively influence student outcomes. Pace staff reported on their self-efficacy beliefs using a survey adapted from Bandura (1997), Justice, et al. (2008), and Pianta, et al. (2017). Staff self-reported beliefs in each of the following areas using a rating scale ranging from 0 (“no feelings of efficacy in this area”) to 4 (“very strong feelings of efficacy in this area”):

- **Efficacy to Influence Decision-Making**: Items include “How much can you influence the decisions that are made in your school?”
- **Instructional Self-Efficacy**: Items include “How much can you do to get through to the most difficult students?”
- **Disciplinary Self-Efficacy**: Items include “How much can you do to prevent problem behavior on school grounds?”
- **Efficacy to Create a Positive School Climate**: Items include “How much can you do to make the school a safe place?”
Self-ratings reported by staff were summed in or to produce a total score as well as scores in each of the above areas. The maximum possible total score was 72, with higher scores indicating stronger perceptions of self-efficacy.

**Self-Efficacy Baseline Results**

The Fall 2017 Self-Efficacy survey was completed by a total of 54 Pace staff members. The Spring 2018 Self-Efficacy survey was completed by 39 staff members. A total of 18 staff completed the Self-Efficacy survey at both the fall and spring time points. Exhibit 36 illustrates the comparison of mean scores at each time point. Compared to the fall, the average rating score at the spring baseline declined by 0.33 points. The one-sample t-test indicated that this difference approached statistical significance (\( p = .06 \)). However, paired-sample comparisons for the 18 staff members with both fall and spring data showed no significant differences in overall Self-Efficacy scores or in any of the specific scales across time points.

**Exhibit 36. Mean Self-Efficacy Score Comparisons (Fall 2017 N = 54, Spring 2018 N = 39).**

These baseline results suggest that in the spring staff report a slightly lower sense of self-efficacy, particularly in terms of their ability to influence instructional outcomes, create a positive school climate, and exercise control over student behaviors and compliance with classroom rules. In the spring, about 24% of the sample reported strong or very strong self-efficacy beliefs. In the fall, 36% of the sample reported strong to very strong self-efficacy.
Overall, however, levels of self-efficacy beliefs among Pace staff remained relatively stable over the course of the baseline year, and can be best described as moderate (neither weak or strong).

A regression analysis found that, of the 4 specific types of self-efficacy measured by this survey, Instructional Self-Efficacy was the strongest predictor of an individual’s overall self-efficacy score. The effect size ($d = 0.45$) indicates that 45% of the variance in overall self-efficacy scores can be explained by Instructional Self-Efficacy scores. Conversely, self-efficacy responses were generally lower regarding staff perceptions of their ability to influence and contribute to creating a positive school climate. About 40% of the total sample reported very low to no feelings of self-efficacy in this area, compared with 29% who reported strong or very strong feelings of self-efficacy.

Self-Efficacy survey results were also analyzed in terms of staff members’ beliefs regarding personal self-efficacy and school self-efficacy. Items from the Instructional and Disciplinary Self-Efficacy scales were grouped to reflect what is here termed Personal Self-Efficacy. These items may reflect a stronger internal locus of control regarding outcomes. Items from the Decision-Making and School Climate scales, on the other hand, may reflect more of an interaction between the individual and school-level factors outside of their control (e.g. influencing decisions about class size). Items from these scales were therefore grouped under what is termed School Self-Efficacy. Results from both fall and spring baseline assessments were averaged in each of these groupings to provide total baseline measures of Personal Self-Efficacy (Mean = 1.65, SD = 0.80, range = 0.20 – 3.95) and School Self-Efficacy (Mean = 1.34, SD = 0.77, range = 0.25 – 3.50).

The same analysis was then conducted with supervisors/leaders and unidentified positions omitted, as the self-efficacy beliefs of these individuals may be somewhat different from those of other staff. This yielded the following results: Personal Self-Efficacy (Mean = 1.67, SD = 0.76, range = 0.40 – 3.95); School Self-Efficacy (Mean = 1.28, SD = 0.72, range = 0.25 – 3.50). Exhibit 37 provides a comparison of mean Personal and School Self-Efficacy scores by staff role. (Note: The spread, or length of each bar, may be more reflective of the number of cases within each role rather than a true reflection of spread.)
Exhibit 37. Personal vs. School Self-Efficacy Mean Score Comparisons by Role.

As can be seen in the above figure, the Personal Self-Efficacy scores for staff in all roles tend to be higher than the School Self-Efficacy scores. Specifically, mean scores for Personal Self-Efficacy were 0.40 points higher than those for School Self-Efficacy. Moreover, results of a paired-samples t-test found that this difference was statistically significant (t(65) = 10.36, p ≤ .0001).

Exhibit 38 illustrates the percentage of total staff falling into each of 4 broad self-efficacy categories. As this exhibit shows, staff generally rated themselves as higher in perceived Personal Self-Efficacy than in School Self-Efficacy.
Results of an exploratory regression analysis found that ratings on several items were inversely associated with Personal Self-Efficacy. Of statistical significance were the following:

- “How much can you do to influence the class sizes in your school?” ($R^2 = -0.53$)
- “How much can you do to control disruptive behavior in the classroom?” ($R^2 = -0.26$)
- “How much can you do to get students to trust teachers?” ($R^2 = -0.20$)
- “How much can you do to help other teachers with their teacher skills?” ($R^2 = -0.15$)

On the other hand, Personal Self-Efficacy was positively and significantly associated with the following items:

- “How freely can you express your views on important school matters?” ($R^2 = 0.40$)
- “How much can you do to increase students’ memories of what they have learned?” ($R^2 = 0.47$)
- “How much can you do to get children to follow classroom rules?” ($R^2 = -0.27$)
What is the Interrelationship Among Growth Mindset and Self-Efficacy Beliefs?

Finally, analyses explored the relation between staff beliefs in each area (growth mindset and self-efficacy). The composite scores from the Growth Mindset and Self-Efficacy surveys, created by combining and averaging fall and spring results, were used in all analyses as these were found to be a reliable overall baseline assessment of staff beliefs. Correlation analysis found significant and positive associations between Personal and School Self-Efficacy and implicit beliefs about Personality ($R = 0.39$), Kind of Person ($R = 0.36$), and World View ($R = 0.30$). Each of these average correlations was statistically significant at $p \leq .05$. Staff beliefs about Intelligence or Morality were not significantly associated with Self-Efficacy.

Regression analysis further supported and clarified these relations. Results indicated that staff beliefs about Personality ($r^2 = 0.37$, $p \leq .002$) and World View ($r^2 = 0.22$, $p = .08$) best predict Personal Self-Efficacy. Exhibit 39 illustrates the mean score for each Growth Mindset implicit theory scale for staff with various levels of Personal Self-Efficacy.

Several trends can be observed in the above exhibit. First, staff with moderate to strong levels of Personal Self-Efficacy report fairly similar Growth Mindset beliefs; those with minimal reported Personal Self-Efficacy differ most starkly from other staff in their Growth Mindset beliefs. Additionally, the area of largest difference in staff beliefs can be seen in the Personality scale. Here, staff with the strongest self-reported Personal Self-Efficacy report much more growth-oriented mindsets. Differences between these beliefs and those of staff with moderate to minimal Personal Self-Efficacy are statistically significant \((F = 4.21, \ p \leq .01)\).

**What was the Impact on Classroom Instructional Practices?**

**Classroom Effective Practices Inventory (CEPI)**

The CEPI is an authentic observational assessment adapted by SPECS team members with permission from research-based criteria as outlined in *Essential Best Practices in Inclusive Schools* (Jorgensen, McSheehan, & Sonnenmeier, 2013). The CEPI consists of 39 items, each rated on a 3-point Likert-type scale reflecting the degree to which the practice is evident (1 = Partially: some evidence that the practice exists, but needs improvement or strengthening; 2 = Usually: much evidence, but practice could be improved or implemented more; 3 = Fully: much evidence that the practice exists, and it would be difficult to find ways to improve). These ratings result in a Total score, reflecting overall use of best practices, as well as best practice in 6 specific domains:

- **Instruction**: Best instructional practices for students with disabilities are utilized (e.g. research-based strategies, universal design, multiple modalities)
- **Supports**: High quality instructional and behavioral support strategies are evident (e.g. positive reinforcements, data-based decisions, plans promoting generalization of skills)
- **Expectations**: Staff behavior conveys expectations of positive engagement and age-appropriate capabilities (e.g. ‘People First’ language used, staff speaks directly to student, age-appropriate vocabulary is used)
- **Social Relationships**: Support for social interactions is provided (e.g. opportunity for peer interaction is provided, development of social relationships is facilitated, students are supported in building social networks)
- **Communication**: Meaningful student communication is facilitated (e.g. respectful communication is used and reinforced, students have access to appropriate modes of communication at all times)
- **Self Determination and Futures Planning**: Students are engaged in goal planning and develop skills that will facilitate positive post-school outcomes (e.g. students are taught self-advocacy and problem-solving skills, opportunities for meaningful choice are provided, strengths-based planning is supported)

CEPI data can be interpreted in several ways: as a rating score, reflecting the average rating assigned using the 3-point scale summarized above; as a sum score, reflecting the total amount of rating points assigned (with a maximum of 117 possible points); or as a percentage
score, reflecting the percentage of total points assigned (rating points / 117 possible points). Higher scores on this instrument are desirable, with mean scores closer to 3, sum scores closer to 117, and percentage scores closer to 100% indicating higher quality classroom practices.

The CEPI was completed for each of the six participating PLAID classrooms by two Pace coaches, trained in data collection procedures by the SPECS team. CEPI data was collected at two time points during the 2017-2018 school year: Fall 2017 (baseline) and Spring 2018 (end-of-year). Spring 2018 CEPI ratings were completed approximately 5 months (22 weeks) after the baseline ratings.

Baseline CEPI Results

Exhibit 4 provides a comparison of CEPI total scores at each timepoint across all 6 participating Pace classrooms. The score types reflect various ways of interpreting the same data, as summarized above. Mean (or average) values for each interpretation are presented, as well as the difference between Spring 2018 and Fall 2017 total scores.

- The increase in CEPI total scores indicates an overall increase in the use of research-based “best practices” across all participating classrooms.

Exhibit 4. CEPI Total Mean and Difference Scores, Fall 2017 to Spring 2018 (N = 6 Classrooms)

<table>
<thead>
<tr>
<th>Score Type</th>
<th>Fall 2017</th>
<th>Spring 2018</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>1.75 (SD = 0.31)</td>
<td>1.90 (SD = 0.61)</td>
<td>0.15 increase</td>
</tr>
<tr>
<td>Sum</td>
<td>67.30 (SD = 15.60)</td>
<td>71.50 (SD = 21.40)</td>
<td>4.20 increase</td>
</tr>
<tr>
<td>Percentage</td>
<td>58%</td>
<td>62%</td>
<td>4% increase</td>
</tr>
</tbody>
</table>

Exhibit 5 provides a visual comparison of CEPI mean rating scores, including total scores as well as scores in each of the 6 domains, at both time points. The plotted points in this figure represent the difference between Spring 2018 and Fall 2017 mean scores, with positive values indicating areas in which mean scores increased.
Exhibit 5. CEPI Total and Domain Mean Rating Score Comparisons, Fall 2017 and Spring 2018 (N = 6 Classrooms)

As can be seen in Exhibit 5, the greatest increase across all classrooms occurred in the Social Relationships domain. The average rating score in this area increased by 0.43 points. Other areas of effective practice in which the average CEPI score increased from fall to spring were the Supports, Expectations, and Self-Determination and Futures Planning domains.

For purposes of establishing a baseline understanding of changes in classroom practices, paired samples t-tests were conducted to compare CEPI mean scores at each time point. Results indicated that the observed differences between Fall 2017 and Spring 2018 CEPI mean scores were not statistically significant at $p \leq .05$.

The information summarized above represents the average CEPI performance across all 6 participating classrooms. Exhibit 6 provides a comparison of CEPI Total sum scores in Fall 2017 and Spring 2018 for each classroom.
Comparisons of CEPI domain scores for each participating Pace classroom were also conducted, and are summarized in detail in Appendix A of this report. These comparisons showed that at least 3 of the 6 classrooms maintained or improved scores in 5 of the 6 CEPI domains, providing more detailed information as to areas of increased use of effective classroom practices. From Fall 2017 to Spring 2018:

- 3/6 classrooms maintained or improved scores in the Supports domain
- 4/6 classrooms improved scores in the Expectations domain
- 5/6 classrooms improved scores in the Social Relationships domain
- 3/6 classrooms improved scores in the Communications domain
- 3/6 classrooms maintained or improved scores in the Self-Determination and Futures Planning domain

Similarities and differences in CEPI scores for the individual classrooms at each time point were also analyzed in order to identify any trends. Results of this analysis indicated that the 6 classrooms seemed to divide into two performance groups:
The first group exhibited a pattern of maintaining or improving upon CEPI scores from fall to spring. This group consisted of Classrooms 1, 2, and 3.

The second group exhibited the reverse pattern; that is, these classrooms showed a general decline in CEPI scores from fall to spring. This group consisted of Classrooms 4, 5, and 6.

**What was the Impact on Student Self-Regulatory Behaviors?**

**Direct Behavior Rating Connect (DBR)**

The Direct Behavior Rating (DBR) Connect is an electronic screening and progress-monitoring tool. It was used to collect and evaluate data on student and teacher behaviors. The DBR Connect system allows an observer to capture a picture of the frequency of various behaviors as they occur in the classroom through brief, repeated assessments of specific behaviors across various observation periods.

**DBR Core Behaviors**

Three “core behaviors” were measured in Pace PLAID classrooms using the DBR Connect. These core behaviors were selected by the PLAID steering committee as being most relevant to the goals and focus of the new PLAID model and therefore important to measure:

- Academic engagement
- Disruptive behavior
- Respectful behavior

**Academic engagement.** Academic engagement was defined as “actively or passively participating in the classroom activity” (Chafouleas & Riley-Tillman, 2015). Active engagement included behaviors such as hand raising, answering aloud, reading aloud, writing, commenting on the topic, and/or using technology in a way that was related to the topic. Passive engagement included listening to the teacher, reading silently, and/or looking at instructional materials. By nature, passive engagement is less overt and can be difficult to report accurately. As such, behavioral definitions were written to align with each form of passive engagement. Listening to the teacher was defined as: student sitting in a designated area, mouth closed, head and eye gaze frequently oriented toward the speaker. Reading silently was defined as: student sitting in a designated area with a text visible and in close proximity to the reader. Finally, looking at instructional materials was defined as: student sitting in a designated area with instructional materials in close proximity, eye gaze frequently oriented toward the instructional materials.

**Disruptive behavior.** Disruptive behavior was defined as a student engaging in an “action that interrupts regular school or classroom activity” (Chafouleas & Riley-Tillman, 2015). Disruptive behaviors included: being out of seat, excessive fidgeting, playing with objects not
related to the task, acting aggressively, and/or talking out/yelling about topics unrelated to the task. Behaviorally, “acting aggressively” was defined as: student throwing classroom objects (e.g., desks, chairs, writing/cutting utensils, etc.), student threatening or making contact with another person with their open hand or fist, and/or student standing in close proximity to another person while using a raised tone of voice.

**Respectful behavior.** Respectful behavior was defined as a student engaging in “compliant and polite behavior in response to adult direction or interactions with peers and adults” (Chafouleas & Riley-Tillman, 2015). Respectful behaviors included: following teacher direction, prosocial interaction with peers, positive response to adult request, and/or verbal or physical disruption without a negative tone/connotation. Prosocial behaviors were defined as: student acting or moving in a way that would benefit another person without the expectation of receiving an immediate reward (i.e., helping a peer in need).

**DBR Observation Procedures**

DBR observations were conducted in Pace PLAID classrooms between March 20, 2018 and May 29, 2018 by a SPECS team member and a Pace staff member trained in DBR procedures. Observations typically took place once per week. Observation periods varied between 5 and 15 minutes. Due to scheduling and other school activities, the precise number of observations conducted in each classroom varied. However, a minimum of 8 observations were conducted in each of the 6 PLAID classrooms.

During each observation period, the observer sat in the classroom away from students so as not to interrupt instruction. The observer recorded the occurrence (or nonoccurrence) of each core behavior over a 5-second interval, one student at a time, for all students in the class. For example, the observer would set a timer for 5 seconds. At the end of the interval, she would look at one student and mark a tally if the behavior was occurring and an “x” if it was not. Each observation covered a minimum of 36 intervals. The time that each observation took place, the activity occurring, and whether the activity was “hands-on” were also recorded.

Following each observation period, raw scores were calculated and converted to scores on a scale from 0 - 10. For positive core behaviors (Academic Engagement and Respectful Behavior), **higher scores** are more desirable as they indicate that these behaviors are occurring **more** frequently. In the case of Disruptive Behavior, interpretation of scores is reversed. That is, because disruptive behavior is unwanted, **lower scores** are more desirable as they indicate that these behaviors are occurring **less** often.

**DBR Core Behavior Data Summary**

Overall, the average DBR Academic Engagement score across all 6 classrooms was 7.8/10 (SD = 0.67). The average Respectful Behavior score was 9.0/10 (SD = 0.44), and the average Disruptive Behavior score was 0.80/10 (SD = 0.39). These overall mean scores are depicted in Exhibit 7.
Exhibit 7. Overall DBR Core Behaviors Mean Scores (N = 6 Classrooms)

Note: Higher scores indicate behaviors are occurring more frequently and therefore are desirable for Academic Engagement and Respectful Behavior; conversely, lower scores are desirable for Disruptive Behavior as they indicate less frequent occurrences of these behaviors.

Exhibit 8 provides a more detailed summary of the DBR data by illustrating the variations in mean DBR scores across the 6 participating classrooms.
Exhibit 8. DBR Core Behavior Mean Scores by Classroom.

Note: Higher scores indicate behaviors are occurring more frequently and therefore are desirable for Academic Engagement and Respectful Behavior; conversely, lower scores are desirable for Disruptive Behavior as they indicate less frequent occurrences of these behaviors.

In addition to the rating scores summarized above, cumulative DBR ratings allowed for categorization of data patterns or trajectories in each classroom over time. These categories reflected the overall trend in the data (“increasing,” “decreasing,” “stable,” or “variable”).

- **Results of this cumulative data (summarized in Exhibit 9) indicate that 4/6 classrooms had high or increasing rates of Academic Engagement; all 6 classrooms had low or decreasing rates of Disruptive Behavior; and 5/6 classrooms had high or increasing rates of Respectful Behavior.**
Exhibit 9. DBR Data Pattern/Trajectory by Classroom

<table>
<thead>
<tr>
<th>Classroom</th>
<th>Academic Engagement</th>
<th>Disruptive Behavior</th>
<th>Respectful Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>High, stable</td>
<td>Low, stable</td>
<td>High, stable</td>
</tr>
<tr>
<td>2</td>
<td>High, some variability</td>
<td>Low, mostly stable</td>
<td>Mostly high, some variability</td>
</tr>
<tr>
<td>3</td>
<td>Increasing</td>
<td>Decreasing</td>
<td>Increasing</td>
</tr>
<tr>
<td>4</td>
<td>Decreasing</td>
<td>Decreasing</td>
<td>High, stable</td>
</tr>
<tr>
<td>5</td>
<td>Decreasing</td>
<td>Low, stable</td>
<td>Increasing</td>
</tr>
<tr>
<td>6</td>
<td>Increasing</td>
<td>Low, stable</td>
<td>Decreasing</td>
</tr>
</tbody>
</table>

Finally, DBR data included the activity occurring during each observation and the nature of that activity (hands-on or not). Hands-on activities were observed in science, technology, and language arts. Lessons centered on science content appeared to have the highest rates of engagement and participation for all students.

- In classrooms where hands-on activities were occurring, Academic Engagement scores tended to be higher.

What Are the Interrelationships Among Classroom Practices and Student Behaviors (CEPI & DBR)?

Data collected using the CEPI and DBR were also analyzed in relation to one another in order to provide a fuller picture of the complex processes occurring within classrooms. Regression analysis was used to assess the association between classroom practices (as measured by the CEPI in Spring 2018) and student behaviors (as measured by the DBR). Results of these analyses demonstrate, as may be expected, significant relations between the two.

- Specifically, higher CEPI scores are associated with higher rates of positive student behaviors (engagement and respectful behaviors) and lower rates of disruptive behavior.

Exhibit 10 provides partial correlation coefficients that represent the strength of the association between each of the DBR core behaviors and the 6 CEPI domains. These values demonstrate that all 3 core behaviors relate to higher CEPI scores (with positive values indicating that higher levels of the core behavior are associated with higher CEPI scores, and negative values indicating that lower levels of the core behavior are associated with higher CEPI scores).

Of note are those values indicating a statistically significant relationship, meaning that they are very unlikely to have occurred by chance alone. As illustrated in Exhibit 10, Academic
Engagement and Respectful Behavior are positively and significantly related various domains of effective classroom practice.

*Exhibit 10. Strength of Association between CEPI Domains and DBR Core Behaviors*

<table>
<thead>
<tr>
<th>DBR Core Behaviors</th>
<th>CEPI Domains</th>
<th>Engaged</th>
<th>Disruptive</th>
<th>Respectful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction</td>
<td>0.75*</td>
<td>-0.12</td>
<td>-0.26</td>
<td></td>
</tr>
<tr>
<td>Supports</td>
<td>0.57*</td>
<td>-0.10</td>
<td>0.34*</td>
<td></td>
</tr>
<tr>
<td>Expectations</td>
<td>0.56*</td>
<td>0.21</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>Relationships</td>
<td>0.69*</td>
<td>0.08</td>
<td>0.32*</td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>0.27*</td>
<td>-0.14</td>
<td>0.43*</td>
<td></td>
</tr>
<tr>
<td>Futures</td>
<td>0.51*</td>
<td>-0.15</td>
<td>0.34*</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Asterisks (*) indicate that the association is statistically significant at p ≤ .05*

**What was the Impact on Students’ Progress in Academic and Functional Capabilities?**

**Functional Outcomes Classification of Assets for Learners (FOCAL)**

The FOCAL is an authentic assessment measure adapted from the US Department of Education’s Office of Special Education Programs (OSEP) framework for mandated documentation and reporting of status and progress data for students with special education needs. This framework was originally developed to promote accountability for early intervention programs. The FOCAL was adapted and extended for use with students in Pre-K through 12th grade in both inclusive and specialized education programs.

The FOCAL is completed collaboratively by classroom teams, and provides a functional classification of the assets and competencies displayed by a student. Classroom teams engage in discussion and consensus decision-making to assign ratings based on observations and assessments of each student in a variety of natural settings. The completion of these ratings by individuals who are familiar with and knowledgeable about the student’s daily functioning makes this a socially valid assessment tool. Two versions of the FOCAL were used in the Pace baseline evaluation: **FOCAL Status** and **FOCAL Progress**.

**FOCAL Status**

The FOCAL Status scale was completed by classroom teams twice during the baseline year, Fall 2017 and Spring 2018. This scale provides an assessment of student functioning in 6 overarching domains representing expected functional competency outcomes for students (PreK – High School), as well as several specific subskill areas within each domain. The 6 FOCAL domains are summarized below:
• **Social Emotional**: Positive social-emotional skills and social engagement
• **Knowledge**: Acquiring and using knowledge and skills
• **Effective Actions**: Taking appropriate actions to meet own needs
• **Self-Regulation**: Demonstrating skills in self-regulatory behavior for classroom learning
• **Academics**: Acquiring and using academic skills for classroom learning
• **Technology**: Acquiring and applying technology skills

All FOCAL Status ratings utilize the following scale, reflecting team judgements regarding the extent to which each student demonstrates *age/grade appropriate functional competencies*:

<table>
<thead>
<tr>
<th>Not Yet</th>
<th>Emerging</th>
<th>Somewhat</th>
<th>Completely</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

**FOCAL Progress**

In addition to the FOCAL Status items, classroom teams completed the FOCAL Progress scale in Spring 2018. This scale provides an assessment of the extent to which each student has shown progress towards acquiring age/grade appropriate functional competencies in each domain area over the course of the school year, and thus can be sensitive to student progress that may not be evident in FOCAL Status ratings alone. The FOCAL Progress items utilize the following rating scale:

<table>
<thead>
<tr>
<th>No Observable Progress</th>
<th>Made Observable Progress</th>
<th>Reached Level of Same Age/Grade Peers</th>
<th>Maintained Level of Same Age/Grade Peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

**Scoring and Interpretation**

Scores for FOCAL Status and Progress scales can be interpreted in two ways. **Rating** scores reflect the average rating assigned using the 7-point scales summarized above. **Sum** scores reflect the total amount of rating points assigned (with a maximum of 147 total points possible).
**Baseline FOCAL Results**

The mean FOCAL Status sum score for all participating Pace students in Fall 2017 was 88.5 (SD = 23.5; range = 35-134). These results place the functional competencies of students within this sample in the “Emerging” range as compared to their same-age/grade peers. In total, about 73% (n = 40) of students in the sample were rated as being below their same-age/grade peers in terms of functional competencies, while 27% (n = 15) of students were rated as having functional competencies or skills that were at or very close to age/grade level.

Students’ Spring 2018 scores remained at about the same level as Fall 2017 scores. The mean FOCAL Status sum score in Spring 2018 was 87.2 (SD = 24.4; range = 40 – 136). Differences between Fall 2017 and Spring 2018 were not statistically significant at $p \leq .05$. This indicates that students’ overall functional competency status remained relatively stable from fall to spring, as illustrated in Exhibit 11.

*Exhibit 11. FOCAL Status Mean Sum Score Comparisons by Functional Skill Categories, Fall 2017 and Spring 2018.*

![FOCAL Status Mean Sum Score Comparisons by Functional Skill Categories](image-url)
Exhibit 12 provides a comparison of the overall mean FOCAL Status rating scores for each participating Pace classroom in Fall 2017 and Spring 2018. The overall mean scores reported here represent the average rating score across all 6 FOCAL domains. Comparisons of pre-post difference scores found that for students in 4/6 classrooms, the overall mean FOCAL Status score increased by an average of 0.37 rating points from Fall 2017 to Spring 2018. In the remaining 2/6 classrooms, the overall mean FOCAL Status score decreased by an average of 0.65 rating points.

Exhibit 12. FOCAL Status Overall Mean Rating Score Comparisons by Classroom, Fall 2017 and Spring 2018

Statistical comparisons of these pre-post differences between classrooms using analysis of variance (ANOVA) and post-hoc ANOVA procedures further confirmed the overall trend summarized above. That is, despite some variation between individual classrooms and a very slight decline in overall mean scores, students in this sample generally maintained relatively stable levels of functional performance across the baseline school year.

Appendix B of this report provides a detailed summary of FOCAL Status and Progress results in each of the 6 domains addressed by this measure. Analysis of this data provides some insights as to the strengths and challenges observed in this population of students.
• The majority of students in this sample made positive, observable progress towards developing age/grade appropriate functional assets and competencies in the following areas:
  o Acquiring and using knowledge and skills
  o Taking appropriate, effective actions to meet own needs
  o Acquiring and using academic skills for classroom learning
  o Acquiring and applying computer-assisted technology skills

• Areas of functional competency that are particularly challenging for Pace students include:
  o Positive social-emotional skills and social assets
  o Skills in self-regulatory behavior for classroom learning

Finally, exploratory regression analysis procedures were used to examine the relations between FOCAL Status and Progress ratings. Results of this analysis indicated that students’ skills in the self-regulation, knowledge, and academic domains were particularly important contributors to the observed progress made by students in the spring.

These results provide valuable baseline information about the functional competencies of Pace students and how these competencies develop over the course of one school year under the current Pace model. This baseline assessment will provide a useful point for comparison as the new holistic PLAID model is implemented.

**What Were the Interrelationships Among Classroom Practices and Student Functional Competencies (CEPI & FOCAL)?**

Regression analysis was used to explore associations between classroom practices and student progress in developing age/grade-appropriate functional competencies. Specifically, associations between the Spring 2018 CEPI and Spring 2018 FOCAL Progress were examined to determine whether or how these two things may be related. Results of the overall analysis (summarized in Exhibit 13) were mixed and at times contrasting, with both positive and negative associations. Positive associations indicate that higher CEPI scores are associated with higher FOCAL Progress scores, while negative associations suggest an inverse relationship.

*Exhibit 13. Relations between Spring 2018 CEPI Total and Spring 2018 FOCAL Progress (Overall and Domain Scores).*

<table>
<thead>
<tr>
<th>FOCAL PROGRESS DOMAIN</th>
<th>R</th>
<th>SIG. LEVEL</th>
<th>EFFECT-SIZE R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOCIAL EMOTIONAL</td>
<td>-0.04</td>
<td>n/s</td>
<td>0.16</td>
</tr>
<tr>
<td>KNOWLEDGE</td>
<td>-0.007</td>
<td>n/s</td>
<td>-0.02</td>
</tr>
<tr>
<td>EFFECTIVE ACTIONS</td>
<td>-0.20</td>
<td>0.068</td>
<td>-0.28</td>
</tr>
</tbody>
</table>
In order to further explore these results, additional analyses were completed in which classrooms were divided into two groups based on overall pattern of CEPI performance (as described in the section of this report summarizing CEPI results). The first group, designated “high-CEPI”, showed a pattern of maintaining or improving on CEPI scores from fall to spring; the second group, designated “low-CEPI,” showed a general decline in CEPI scores from fall to spring. Exhibit 14 provides a comparison of the mean FOCAL Progress scores for students in high-CEPI vs. low-CEPI classrooms.

Exhibit 14. Student FOCAL Progress Scores by Classroom CEPI Performance Categories
Considering the data in this manner is useful in understanding the overall relations between the CEPI and FOCAL Progress. As shown in Exhibit 14, students in high-CEPI vs. low-CEPI classrooms seem to have made progress in different areas of functional competency. It is possible that these differences indicate different areas of instructional focus; this question may be worth exploring further in Year 2. The result of these differences is that when the 6 classrooms are considered together, the at times inverted progress scores cancel one another out and thus lead to the non-significant or negative correlations noted above.

When such effects are accounted for in analyses, independent t-tests indicate that effective classroom practices positively impact student progress in acquiring and using academic skills for classroom learning. Additionally, when general progress trends of students in the two classroom categories are compared, it is evident that students in high-CEPI classrooms made somewhat more overall progress in developing age/grade-appropriate functional competencies ($R^2 = 0.33$) than did students in low-CEPI classrooms ($R^2 = 0.11$).

What was the Impact on Students’ Strengths and Resiliency Skills?

Social Emotional Assets and Resilience Scales (SEARS)

The SEARS (Merrell, 2011) is a multi-informant, strengths-based measure of students’ social-emotional attributes, including constructs such as self-regulation, responsibility, social competence, and empathy. The SEARS Short Form was used in this evaluation. This measure consists of 12 items that reflect students’ social resiliency and assets, with items worded as positive and desirable traits. Multiple versions of this measure are available, each designed to be completed by a different rater. The following versions of the SEARS Short Form were used to assess the social emotional strengths and assets of students at Pace:

- **SEARS-Teacher (SEARS-T):** Classroom teachers in each of the 6 participating classrooms completed this form for all participating students.
- **SEARS-Child (SEARS-C):** Participating students aged 8-12 years, or in grades 3-6, completed this self-report measure.
- **SEARS-Adolescent (SEARS-A):** Participating students aged 13-18 years, or in grades 7-12, completed this self-report measure.

The SEARS-T and SEARS-C/A were each completed at two time points, Fall 2017 and Spring 2018. Items on all forms are rated using a 4-point scale indicating how often the positive statement is true for that student (0 = Never, 1 = Sometimes, 2 = Often, 3 = Always). These ratings produce 2 primary forms of score: **sum score** (reflecting the sum total of all ratings), and **T-score** (a form of standard score reflecting performance in relation to a normative group). In addition, scores can be categorized by Tiers associated with risk and/or functional levels. Higher score values on all SEARS forms are indicative of higher levels of student social-emotional competence. Exhibit 15 provides a guide to interpreting each type of SEARS score.
Exhibit 15. SEARS Score Interpretation.

<table>
<thead>
<tr>
<th>Tier</th>
<th>Raw Score</th>
<th>T score range</th>
<th>Functioning</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0 thru 7</td>
<td>23 - 32</td>
<td>High Risk</td>
</tr>
<tr>
<td>2</td>
<td>8 thru 13</td>
<td>33 - 40</td>
<td>At Risk</td>
</tr>
<tr>
<td>1</td>
<td>14 thru 36</td>
<td>41 - 68</td>
<td>Average to High Functioning</td>
</tr>
</tbody>
</table>

Baseline SEARS Results

SEARS-T. At the first time point (Fall 2017), teachers completed a total of 57 SEARS-T forms. The average SEARS-T sum score in Fall 2017 was 11 (SD = 6.0, range = 0 – 24). This corresponds to a T-score of 37.3 (SD = 7.4, range = 24 – 53), which places the average score within Tier 2 or the At-Risk category. In total, almost 70% of students (n = 39) had SEARS-T scores within the At-Risk or High-Risk categories at the Fall 2017 time point. Specifically, At-Risk students comprised 42% (n = 24) of the total sample and High-Risk students 26% (n = 15).

At the second time point (Spring 2018), the average SEARS-T sum score increased by 1.6 points to a mean of 12.6 (SD = 7.5, range = 0 – 30). The associated T-score increased by 2.1 points to a mean of 39.4 (SD = 9.6, range = 23 – 61). Paired samples t-test comparisons between Fall and Spring SEARS-T mean scores found these changes not to be statistically significant at $p \leq .05$.

A post-hoc analysis contrasting students’ chronological age, developmental age, and Fall 2017 Tier category showed statistically significant differences. As previously discussed, the developmental age can be used as a proxy indicator of students’ overall functioning. As illustrated in Exhibit 16, students whose SEARS-T scores fall into the At-Risk or High-Risk categories have significantly lower developmental ages than do students with SEARS-T scores in the average to high functioning category (Tier 1).
Exhibit 16. Comparisons of Student Chronological and Developmental Ages by SEARS-T Tier Category.

As shown in Exhibit 16, students in the average to high functioning category exhibited chronological and developmental ages that were roughly equivalent. Students in the At-Risk and High-Risk categories displayed an average difference of approximately -2.7 years between their chronological and developmental ages. Students in the At-Risk and High-Risk categories displayed developmental competencies rated at about 8.8 years, whereas students in the average to high functioning category displayed developmental functioning at about 11.6 years.

In Exhibit 17, pre/post differences in SEARS-T scores are considered by student risk category. As can be seen in this exhibit, the mean T-score for students in the Tier 1 (average to high functioning) and Tier 2 (At-Risk) categories increased from fall to spring. For those students exhibiting the most challenges with social-emotional skills and competency (those categorized as Tier 3 or High-Risk), however, the mean T-score decreased. In Spring 2018, the differences between students in the lower risk categories and those in the High-Risk category were statistically significant ($p \leq .001$). These trends suggest that over the course of this baseline year, the differences between these groups of students became more pronounced in the view of their teachers.
The 6 participating classrooms also exhibited differences in how students were distributed across the SEARS-T Tier categories. Exhibit 18 shows the frequency (total number) of students in each classroom categorized as At-Risk or High-Risk vs. those categorized as No-Risk (Tier 1, average or high functioning) as well as the percent of students in the At-Risk or High-Risk group in each classroom. Approximately 65% (n = 24) of all students categorized as At-Risk or High-Risk on the SEARS-T were in classrooms 4, 5, and 6.
Exhibit 18. Students’ Risk Status and Percent of Students Categorized as At-Risk or High-Risk by Classroom.

An analysis of variance (ANOVA) was conducted to compare pre/post (Fall 2017/Spring 2018) T-scores as well as pre/post mean T-score differences across classrooms. These comparisons are illustrated in Exhibit 19.
Exhibit 19. Pre/Post T-Score Comparisons and Pre/Post Mean T-Score Differences by Classroom.

The post-hoc ANOVA results showed that observed differences between Fall/Spring T-scores were generally not significant at $p \leq .05$. As can be seen in Exhibit 19,

- **Students in most classrooms were reported to have maintained and/or slightly improved upon their Fall 2017 baseline levels of social-emotional skills and competencies.**

Finally, SEARS-T results were analyzed in relation to 4 specific social-emotional asset domain areas. These analyses are summarized in Appendix C of this report.

**SEARS-A.** This measure was completed by students aged 13 and older, who comprised 49% ($n = 30$) of the total student sample. Of these students, 30% ($n = 9$) were female. The mean chronological age for this group of students at the Spring 2018 baseline time point was 14.4 years ($SD = 1.6$ years, range = 12-17 years). The mean developmental age for this group at the Spring 2018 baseline was 11.6 years ($SD$ = years, range = 7.7 – 16 years). The average difference between students’ chronological and developmental ages was -3.7 years.
The response rate for the SEARS-A was 90% in Fall 2017 and 83% in Spring 2018. At the spring time point, several students were reported as chronically absent while several others refused to complete this measure.

At the Fall 2017 baseline, the mean SEARS-A sum score was 18.6 (SD = 6.3, range = 3 – 31). This corresponds to a mean T-score of 42.5 (SD = 10.6, range = 16 – 63). At the Spring 2018 baseline, the average SEARS-A score increased by 1.9 sum points to a mean of 20.4 (SD = 5.6, range = 8-34) which is equivalent to a mean T-score of 45.3 (SD = 10.2, range = 24 – 68). The paired samples statistic showed that this slight increase in student self-ratings from fall to spring was not statistically significant.

In total, at the Spring 2018 baseline 20% (n = 18) of the 25 adolescent students who completed this self-rating reported T-scores of 41 or higher which are categorized as Tier 1 or indicative of average to high levels of social-emotional functioning. Students' self-report was compared with teacher ratings for both fall and spring baselines, using the standardized T-scores for each SEARS measure. This comparison is illustrated in Exhibit 20.


As can be seen in Exhibit 20, adolescent students generally reported a much more positive view of their social-emotional skills and assets than did their teachers. Paired-sample comparisons indicated that, at the Fall 2017 baseline, these differences were not statistically significant. At the Spring 2018 baseline, however, the difference between student and teacher mean T-scores was statistically significant (p ≤ .05), with students rating their social-emotional
assets on average 6.5 points higher than teachers. In general, these trends suggest that students’ confidence in their own social-emotional competencies tended to increase from fall to spring while teachers’ perceptions of student competencies and skills remained relatively stable.

Finally, Exhibit 21 compares the SEARS-A and SEARS-T mean T-scores at fall and spring baselines for each of the 3 participating classrooms with students of adolescent age. As seen here, and verified in statistical comparisons, the classrooms show some variability in the degree of consonance between teacher and student perceptions of students’ social-emotional strengths and competencies.


**SEARS-C.** This measure was completed by students aged 12 and younger, who comprised 51% (n = 31) of the total student sample. Of these students, just 13% (n = 4) were female. At the Spring 2018 baseline the average chronological age of this group was 10 years (SD = 1.1 year, range = 8 – 12 years), with most students (n = 21 or 70%) being between the ages of 10 and 12 years. The mean developmental age for this group at the Spring 2018
baseline was 8.8 years (SD = 1.7 years, range = 5.9 – 12.5 years). The average difference between chronological and developmental age for this group was -3.1 years.

The response rate for the SEARS-C was 100% at both time points. At the Fall 2017 baseline, the mean SEARS-C sum score was 17.4 (SD = 6.1, range = 5 – 29). This is equivalent to a mean T-score of 40.1 (SD = 10.3, range = 19 – 59). At the Spring 2018 baseline, the mean SEARS-C sum score increased slightly to 19.6 (SD = 7.5, range = 7 – 35) which corresponds to a mean T-score of 43.7 (SD = 12.4, range = 22 – 66). In the fall, 48% of students in this group (n = 15) reported scores within the At-Risk or High-Risk range. By spring, this number decreased slightly with 39% (n = 12) students reporting scores in the At-Risk or High-Risk range.

Paired-samples t-tests indicated that changes in SEARS-C mean scores and risk categories from fall to spring were not statistically significant. In general, the self-reported ratings of social-emotional competency for students 12 and under remained fairly consistent across the two time points.

Student self-reports were again compared with teacher ratings using the mean T-score for each measure at both fall and spring time points. These comparisons are illustrated in Exhibit 22.

Exhibit 22. Comparison of SEARS-C and SEARS-T Mean T-Scores.
This group displayed a trend that is roughly opposite to that displayed by the adolescent group. That is, student self-report scores remained fairly consistent from fall to spring whereas teacher ratings of student social-emotional competencies increased by an average of about 7 points over this time. In the fall, the difference between teacher and student ratings was statistically significant ($p \leq .001$) but by spring the scores were almost equivalent. This suggests that, contrary to the trend observed in the adolescent group, children aged 12 and under maintained relatively stable perceptions of their social-emotional competencies over time while teachers’ perceptions of student social-emotional assets and skills increased. Exhibit 23 illustrates the mean SEARS-C and SEARS-T T-score comparisons for each classroom with students aged 12 or younger.

Exhibit 23. Comparison of SEARS-C and SEARS-T Mean T-Scores by Classroom.

In summary, there seems to be much more positive concurrence in the appraisal of social-emotional strengths and competencies between younger students and teachers than between adolescent students and teachers. These contrasting baseline trends should be closely monitored and explored in future analyses.
PLAID Focus Group Summary

At the conclusion of the 2017-2018 school year, a focus group was conducted in order to obtain direct, qualitative feedback from individuals involved in the development and piloting of the new PLAID model. The group was moderated by a SPECS team researcher, and consisted of 3 teachers and 2 administrators. The participating teachers had been involved in the PLAID project throughout the school year in two ways: operationalizing the PLAID framework by developing lesson/unit plans, and piloting these plans in their own classrooms. The participating administrators were members of the PLAID steering committee.

Positive Impacts

The group identified several ways in which their involvement with PLAID had positively impacted the professional practice of participating teachers. These included:

- “More thoughtful planning,” including thinking and planning in terms of the technology framework (even though it was not yet “up and running”) and incorporating more “character-building aspects” into academic lessons.
- Getting “outside of (their) comfort zone,” specifically by increasing students’ independent work in the classroom rather than relying primarily on group-led instruction.
- Making increased connections between academic lessons and other activities, such as “Fun Fridays.”

Positive Examples of PLAID “In Action”

Participants were asked to further illustrate positive impacts of PLAID by providing positive examples, drawn from their own experience, of PLAID “in action.” Examples provided by participants focused primarily on increased connections between various aspects of the curriculum and school day. These included:

- Staff simultaneously addressing both interpersonal skills and de-escalation strategies when interacting with a student in the hallway.
- Teachers purposely drawing connections between subject matter (such as a novel), character traits (such as honesty), interpersonal skills (such as eye contact), and life events (such as changes, tradition, or family routines).
- Helping a student to reframe problematic remarks to a peer by drawing upon academic work (poetry) produced by the student.
Examples Showing Change is Still Needed

While the positive examples summarized above are encouraging, the group had little difficulty providing examples showing that changes associated with PLAID are still needed. Examples provided in this regard predominantly emphasized students’ difficulty generalizing social skills or positive behaviors beyond the lesson or context in which these skills are taught. They included:

- Difficulty waiting patiently in lines
- Disrespectful treatment of cafeteria staff
- Use of language inappropriate for school
- Problematic social interactions with peers

These examples offer further support for expansion of the PLAID model, as participants expressed the belief that incorporating PLAID components across the school day will lead to improved generalization of positive social skills and reduce the need for discipline or correction.

Implementation and Expansion

As Pace leadership looks ahead to the 2018-2019 school year and plans to expand PLAID, the group was asked to provide input on this process based on their experiences piloting the new model. They first discussed general considerations regarding implementation of PLAID, then more specifically addressed expansion.

Regarding general implementation of PLAID, participants were first asked to identify basic classroom practices that they believed to provide a necessary foundation. These included:

- Behavior management
- Clear structure (expectations, consequences, etc.)
- Well-established classroom procedures

In addition to these basic practices, participants also more specifically addressed operationalization of the PLAID framework. Based on their experience, they identified the following as the most important considerations in developing lesson plans that operationalize PLAID:

- Appropriate accommodations and modifications to lessons that account for differing ability levels of students in a given classroom
- Flexibility, allowing for teachers to respond to and adjust for student interest and attention levels (or “let them (students) lead”).
Finally, participants provided specific suggestions regarding the planned expansion or “roll-out” of PLAID during the upcoming school year. All participants agreed that selecting 1-2 subject areas as a focus for incorporating PLAID during expansion would help to ease the process. Additionally, all agreed that English Language Arts (ELA) and Social Studies were ideal subjects to begin expansion efforts.

Participants also identified several considerations that they believed would be important in planning for expansion. These included:

- **Training** to ensure that staff knew how to use or access available resources and materials; this was thought to be especially important in making new staff aware of the materials available.
- **Functional technology** for unit and lesson plans, and adequate training to increase staff comfort level with using this technology.
- **Staff characteristics**, particularly comfort levels with addressing personal, social, or emotional issues with students.

**Conclusions and Recommendations**

The information summarized in this report provides an important baseline understanding of the status and progress of Pace students and teachers over one school year under the current Pace model. This baseline information will provide important points of comparison that will be critical to addressing the evaluation research questions outlined in earlier portions of this report as Pace expands and implements the PLAID model during Year 2.

**These conclusions and recommendations are derived from collaborative discussions among SPECS/Pitt and the Pace Leadership Team and the data.** The following next steps of activities and deliverables will focus the work of the SPECS for PLAID team and the Pace Leadership team in Year 2 and build upon the clear success of the Year 1 developmental phase and initial implementation of the program evaluation and continuous quality improvement (CQI) plan:

- Continue to revisit the logic model to ensure a match between the PLAID goals, interventions and evaluation plan.
- Consider input obtained from the PLAID focus group when planning PLAID implementation and expansion.
- Construct operational definitions of the PLAID components in the matrix.
- Subject the operational definitions to focus group analysis and feedback for finalization.
- Design rubrics to evaluate the implementation and effect of the PLAID operational features or “active agents”.
- Continue with data collection (i.e., CEPI, FOCAL, SEARS, Direct Behavior Ratings, DCHS, & SPECS Monitor) as stated in year 1 evaluation plan.
• Conduct trainings for continued, independent use of the DBR by Pace staff supplemented by independent observations by SPECS team members. Construct and use “social validity surveys” to supplement the focus group sessions in order to examine the spread of effect of the Growth Mindset philosophy and related PLAID components and operational features across the Pace school as well as in the classrooms and with parent/families and perhaps any community partners (e.g., placement sites for older students).

• Conduct evaluation data analysis to examine the complex interrelationships among coaching/mentoring; instructional practices; student progress and outcomes, including individual goal attainment scaling and provide regular feedback to Pace administrators and staff for CQI.

• Conduct within and between group analyses of the Pace (control) and PLAID (pilot) classrooms.

• Conduct analyses to compare Pace data with state normative data from Include Me data from the PDE/Bureau of Special Education & the Arc of Pa.

• Students in high-CEPI vs. low-CEPI classrooms seem to have made progress in different areas of functional competency. Further explore these differences to determine how they may be related to differences in instructional focus. Contrasting baseline trends were observed in student and teacher assessments of social-emotional competencies. These trends should be closely monitored and explored in future analyses.

• Baseline differences were observed in students’ functional capabilities. Consider adopting the School Function Assessment (SFA) or another similar measure to further explore these differences. Teachers and team members could also use the SFA as a uniform curricular for both individualized education program planning and assessment of student status and progress.

• The SPECS Monitor provided a clear picture of the scope, intensity and content of coaching under the current Pace model. Continue to utilize the SPECS Monitor to look for changes in the content and to ensure a match with PLAID.

• Conduct an on-going analysis of the coaching/mentoring dimensions recorded daily/weekly/monthly on the SPECS Mentoring Monitor (Bagnato, 2010) to determine the most effective mentoring/coaching features and strategies which are determined to promote improved instruction and management.

• Disseminate / co-present the year 1 outcomes from the developmental phase of PLAID to the Pace Board and other partners.
Positive Growth Mindset References


Strosher, H. L. W. (2003). Prospective and Practicing Teachers' Beliefs: A Study of Implicit Theories of Intelligence and Teacher Efficacy. (MA Research), University of Victoria, Canada.

Yeager, D. S., Fweck, Carol S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. Educational Psychologist, 47(4), 302-314.