Using the Teaching Pyramid Observation Tool to Support Professional Development Experiences in Public 4 K Settings

Keith Johnston1 · Tillman Hall1 · Sandra Linder1 · Floyd Creech2 · Penny Danielson2

Accepted: 7 October 2021 / Published online: 1 November 2021 © The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract
This study examines how well school groups in the Coastal Plains region of a southeastern state align with the socio-emotional strategies for four-year-olds outlined by the Key Practices in the Pyramid Model Framework (Teaching Pyramid Observation Tool—Research Edition [Manual], Baltimore, MD: Brookes, 2014). Many schools in these groups serve children in poverty, so it is imperative that teachers engage in developmentally appropriate socioemotional practices. The authors used the TPOT observation tool to take baseline data and examine longitudinally after giving ongoing professional development experience sessions. The topic for each respective group’s professional development experience was not uniform but based on need. Therefore, the authors hypothesized that ongoing individualized professional development experiences would significantly increase the groups’ TPOT scores. The results show that the professional development experience treatment did, indeed, raise scores in most cases. Limitations and recommendations for continued research are discussed.

Keywords Early childhood · Socio-emotional learning · Pyramid model · TPOT · Professional development experience

Keith Johnston
keith8@g.clemson.edu
Sandra Linder
sandram@g.clemson.edu
Floyd Creech
fcreech@fsd1.org
Penny Danielson
danielsonpenny@gmail.com

1 Clemson University, Clemson, SC 29634, USA
2 Florence 1 School District, Florence, USA
Résumé
Cette étude examine dans quelle mesure les groupes scolaires de la région des plaines côtières d’un État du sudest s’alignent sur les stratégies socio-émotionnelles pour les enfants de quatre ans décrites par les pratiques clés dans le cadre du modèle pyramidal (Teaching Pyramid Observation Tool—Research Edition [Manuel] , Baltimore, MD : Brookes, 2014). De nombreuses écoles de ces groupes desservent les enfants en situation de pauvreté, il est donc impératif que les enseignants s’engagent dans des pratiques socio-émotionnelles adaptées au développement. Les auteurs ont utilisé l’outil d’observation TPOT pour prendre des données de base et les examiner longitudinalment après avoir donné des séances d’expérience de développement professionnel en cours. Le sujet de l’expérience de développement professionnel de chaque groupe respectif n’était pas uniforme mais basé sur les besoins. Par conséquent, les auteurs ont émis l’hypothèse que les expériences de développement professionnel individualisé en cours augmentereraient considérablement les scores TPOT des groupes. Les résultats montrent que le traitement d’expérience de développement professionnel a, en effet, augmenté les scores dans la plupart des cas. Les limites et les recommandations pour la poursuite de la recherche sont discutées.

Abstracto
Este estudio examina qué tan bien los grupos escolares de la región de las llanuras costeras de un estado del sureste se alinean con las estrategias socioemocionales para niños de cuatro años descritas en las Prácticas clave en el marco del modelo de pirámide (Herramienta de enseñanza de la observación de pirámides — Edición de investigación [Manual]) , Baltimore, MD: Brookes, 2014). Muchas escuelas en estos grupos atienden a niños en situación de pobreza, por lo que es imperativo que los maestros se involucren en prácticas socioemocionales apropiadas para el desarrollo. Los autores utilizaron la herramienta de observación TPOT para tomar datos de referencia y examinarlos longitudinalmente después de dar sesiones de experiencia de desarrollo profesional en curso. El tema de la experiencia de desarrollo profesional de cada grupo respectivo no fue uniforme sino que se basó en la necesidad. Por lo tanto, los autores plantearon la hipótesis de que las experiencias continuas de desarrollo profesional individualizado aumentarían significativamente las puntuaciones de TPOT de los grupos. Los resultados muestran que el tratamiento de la experiencia de desarrollo profesional, de hecho, elevó las puntuaciones en la mayoría de los casos. Se discuten las limitaciones y recomendaciones para la investigación continua.

The recognition of the positive outcomes associated with early socio-emotional competence has lent itself to a focus on pinpointing and executing effective environmental, interactional, and instructional strategies to foster children’s socio-emotional development (Hemmeter et al., 2018). On a broad scope, the study being reported on here focused on socio-emotional learning because of the positive impact it has on child outcomes. For example, Jones et al. (2015) put forth in their report that socio-emotional skills in the kinder year, children five to six years old, have a statistically significant relationship with positive adult outcomes such as higher levels of
education, less time in the justice system, and better mental health. It is essential to note the National Association for the Education of Young Children, the authority in research-based best practice for Early Childhood Education, affirms that socio-emotional skills are not innate. Rather, they are fostered and developed through contact and conversations with other children and adults. A principal feature of these interactions with children is their intentionality (Epstein, 2009; NAEYC, 2009). This is especially true for children with difficulties in the socio-emotional domain of development (Joseph & Strain, 2003). The Pyramid Model is a framework that recognizes the need for intentionality and guides preschool teachers, instructing children from 2–5 years old, in facilitating and teaching socio-emotional skill development. It behooves children to learn these skills early. Children with socio-emotional challenges should receive early intervention as the challenges may fester with time and result in negative outcomes (Bulotsky-Shearer & Fantuzzo, 2011; Hauser-Cram & Woodman, 2015; Brennan et al., 2012).

This study was state-funded with the goal of improving the quality of publicly funded 4 K classrooms in high poverty school Groups. The authors’ aim was to gain insight into the quality of the professional development experiences implemented to support the socio-emotional component of the Groups in the study. The study uses the Pyramid Model framework applying the assessment instrument the Teaching Pyramid Observation Tool (TPOT). The Pyramid Model parallels the Positive Behavior Interventions and Support (PBIS) framework that is already used in many K-12 schools (Dunlap et al., 2014). Both the Pyramid Model and PBIS have a primary tier in which all children are included in the socio-emotional strategy lessons. They also both include a second level for young children at risk and a third level for a more individualized intervention for those whose problem behavior is more persistent (Hemmeter et al., 2016). Behaviors are “challenging” when the behavior/s is persistent and interferes with the child’s relationships with peer and/or adults. The full definition of challenging behaviors in the Teaching Pyramid Observation Tool, TPOT for Preschool Classrooms Manual (Hemmeter et al., 2014) is as follows:

Challenging behavior is defined as behavior that includes (1) physical aggression, such as hitting, kicking, punching, spitting, throwing objects forcefully, pinching, pushing, and biting. (2) climbing on things in the classroom that are not permitted; (3) destroying property, destroying what another child is working on regardless of the other child’s response; (4) taking toys away from other children forcefully; running which poses a safety risk for the child or others or elopement from the classroom; (6) tantrum behaviors that include behaviors such as kicking, screaming, pushing an object or person, stomping feet or head banging; (7) verbal aggression including yelling, threats, screaming at another person, calling children bad names, and saying bad words; (8) ordering an adult to do something (e.g. “leave me alone”); (9) persistent or prolonged crying that is loud or disruptive or ongoing crying that interferes with the child’s engagement in activities; (10) inappropriate use of materials (e.g., jumping off chairs, slamming materials, throwing objects); (11) statements that are non-compliant (e.g.“I’m not going to do it!) or clear and explicit verbal or physical refusal to follow directions; or (12) inappropriate touching, stripping, and
other behaviors that are hurtful, disruptive, or dangerous to self or others. (p. 24)

By gathering data on teacher performance, the study tested the effectiveness of professional development interventions as it impacts overall teacher practice in terms of socio-emotional development, as measured by the TPOT assessment.

Review of Literature

Recent research in the area of socio-emotional development in early childhood education establishes the benefit of supporting socio-emotional skills as well as supporting teachers in instructing them. The Organization for Economic Co-Operation and Development’s comprehensive report “socio and Emotional Skills for Student Success and Well-being: A Conceptual Framework for the OECD Study on socio and Emotional Skills”, Chernyshenco et al. (2018) found that socio-emotional skills have a relationship with life satisfaction and well-being. A study by Jones et al. (2015) suggests that when teachers perceive students to have socio-emotional skills early on, they are likely to have positive outcomes as adults. This study also shows that teacher ratings of noncognitive skills predicted outcomes in the domains of education, employment, crime, mental health, public assistance, and substance abuse (2015).

The overwhelming trend in the literature on the connection between teacher–child relationships and social skills is that more positive relationships lend themselves to better social skills (Baker et al., 2008; Decker et al., 2007; Liew, 2012). It is therefore imperative that when we assess teacher effectiveness, we examine the quality of teacher–child relationships as well as classroom quality. The Teaching Pyramid Observation Tool (TPOT) is an instrument that assesses teacher–child interactions, and how well aligned a classroom’s practices are to the Pyramid Model. The Pyramid Model has been shown to foster practices that support socio-emotional development (Hemmeter et al., 2018). Hemmeter et al. (2018) highlight that environmental, interactional, and instructional practices come together to support children’s socio-emotional development. It is not enough to solely focus on academics.

Given the outcomes we see from the literature, educators need to be intentional about realizing students’ socio-emotional potential and intentionally teaching socio-emotional skills. One effective avenue to encourage this intentionality is through teacher professional development experiences in supporting children in their socio-emotional development in the preschool classroom. Snyder et al. (2012) state that quality early childhood training is imperative in order for children to receive quality instruction in socio-emotional competence. As cited by Hemmeter et al. (2014) in their efficacy trial of the Pyramid Model, teachers who participated in professional development experiences tended to see improvements in children’s socio-emotional skills.
**Methods**

The authors of the study along with their observation team used the TPOT instrument (described below) to observe prekindergarten teachers across 7 school Groups (only 5 school Groups are included in this study due to the inadequacy of sample size). The school Groups consisted of state-funded pre-k classrooms in public schools representing Groups 1—4 and Group 5 consisting of federally funded Head Start preschool classrooms. All classrooms in each group were observed with the TPOT in the Fall of 2018 to acquire baseline data for Pyramid Model Key Practices. We sought to investigate how professional development experiences (PDEs) brought each group of classrooms into alignment with Key Practices as described by the Pyramid Model (Table 1).

Group 1 included 8 observations, Group 2 included 5, Group 3 included 16, Group 4 included 6, and Group 5 included 22 for a total of 57 classrooms participating. PDE sessions were planned based on needs determined by baseline data. So, the professional development topics varied between groups as each group of teachers identified a unique set of needs-based upon their TPOT data. In the Spring of 2019, another iteration of TPOT observations was conducted following the PDEs in order to determine if there was an increase in Pyramid Model Key Practices. In the Fall of 2019, there was a third iteration of observations for the new school year to determine the longitudinal effects of the PDEs and determine needs for a second year of development. While the intention was to conduct another round of PDEs and TPOT observations in 2020 the Covid-19 pandemic prevented this from occurring. The 2018–2019 scores were analyzed by performing t-tests to determine the longitudinal, within-group relationship. ANOVA analyses were conducted to determine the between-group relationships.

**Participants**

The 5 Groups included in the study are all located in the Coastal Plains region of a southeastern state. The children/families attending the Pre-K programs represented in the project primarily live in a high poverty rural area of a region of a southeastern state. Approximately 85% of the children are African American and from families where English is their first language. In this state, we do not yet have Universal Pre-K. Students must qualify to attend the state-funded preschool program based upon family income, and be at 75% of the Poverty Index or higher. To qualify for the federally funded Head Start programs, students must be from

| Group   | Sample size |
|---------|-------------|
| Group 1 | 8           |
| Group 2 | 5           |
| Group 3 | 16          |
| Group 4 | 6           |
| Group 5 | 22          |
| Total   | 57          |
families living at 100% of Poverty. Approximately 75% of the teachers in the project were Caucasian, and 25% African American.

Group 1 was part of a rural school Group that serves 1,622 students as of 2018–2019 (SC Report Card, 2020). It consists of 3 schools (including elementary, middle, and high schools) that employ 91 teachers (SC Report Card, 2020). 45.8% of students enrolling in kindergarten were considered kindergarten-ready per the Kindergarten Readiness Assessment (KRA) used by the state. Group 2 is from one of two school Groups in their county. It enrolls about 4,116 students across 8 schools that employ 232 teachers (SC Report Card, 2020). Readiness assessments determined that 18% of kindergarten students entered school ready to start in the 2018–2019 school year. Group 3 is from a Group with 22 schools, 1,155 teachers and enrolls 15,856 students. It is the largest Group that participated in the study. The overall readiness of their kindergarten students upon entry was 32.1%. Group 4 is from a Group with 8 schools with 232 teachers that serve 3317 students (SC Report Card, 2020). According to their Kindergarten Readiness Assessment results, 20.6% of their kindergartners entered school ready to learn. Group 5 consists of 15 early childhood programs across 3 counties (PeeDee CAP Head Start, 2020). There are 742 children enrolled in the Head Start program, which also runs Early Head Start, serving children from birth to age 5.

**Professional development intervention**

The Pyramid Model is a framework with multiple tiers that outlines evidence-based Key Practices for fostering socio-emotional skills (Hemmeter et al., 2014). It works on three levels to foster the development of socio-emotional skills for all children including those who are at-risk and those who show persistent socio-emotional issues (Hemmeter et al., 2018). Pyramid Model practices are research-based, developmentally appropriate, and designed to support children instructionally in daycare, public Pre-K, Head Start, and other early childhood educational settings (Hemmeter et al., 2016). It is designed to be used with infants and young children. The framework fosters socio-emotional skills through implementing evidence-based strategies as well as intervening in situations where children show persistent behavioral challenges (Fox et al., 2003; Hemmeter et al., 2006). The Pyramid Model has built-in practices that have to do with classroom environments that support socio-emotional development and foster positive relationships.

The Pyramid Model tier that includes all children, the first tier, includes methods and strategies that are conducive to teacher–child interactions and relationships as well as learning environments (Hemmeter et al., 2014). These methods and strategies range from engaging in play with the children and conversing with them to providing them with positive feedback. Constructing the classroom environment so that it promotes engagement, reinforcing expectations, and facilitating transitions benefits all children as well (Burchinal et al., 2010; Chien et al., 2010; Hemmeter et al., 2014). The all-inclusive tier also incorporates strategies and methods that have to do with quality learning environments that foster student engagement and
developmentally appropriate activities. These strategies and methods include having proper and relevant materials at the learning centers, having a balance of whole group instruction and small group instruction that coincides with the needs of the children, and differentiating the communication of instructions so that they are clear to all children (Chien et al., 2010; Jolivette et al., 2001).

The second tier of the Pyramid Model consists of the intentional and purposeful teaching of socio-emotional skills to all children. This includes instruction in small groups or on an individual basis for those who are at risk of falling behind in their socio-emotional skills. Some topics may include recognizing and identifying emotions, expressing emotions, self-regulation, regulating difficult emotions, problem-solving skills, joining play, sharing, and working/playing with others (Hemmeter et al., 2018). The strategies, lessons, and methods that early childhood practitioners and family members use must be evidence-based as well as developmentally appropriate. They must be practical as well as be able to meet the diverse socio-emotional needs of an inclusive preschool classroom (Hemmeter et al., 2018).

At the third tier of the Pyramid Model framework, preschool teachers are part of a team that creates and executes an individualized behavior support plan that consists of practices that target the child’s specific needs (Dunlap and Fox, 2011; Dunlap et al., 2014) both at home and at school (Hemmeter et al., 2014). The third tier is reserved for individual students when the first tiers aren’t enough to make a change in their behavior. It includes targeting triggers of problematic behavior as well as teaching alternative acceptable behaviors to the inappropriate problematic behaviors. The third tier also calls to create consequence strategies to ensure that the negative behavior is not rewarded (Hemmeter et al., 2014). Teachers were trained to document challenging behaviors with the Behavior Incident Report form which we designed with our partner administrators as part of the grant activity. The purpose of this form was not just for documentation but to purposefully have teachers review the Pyramid Model strategies from the second and third tiers. The challenging behaviors which were most frequently reported were hurtful physical behaviors and temper tantrums. The TPOT assessment tool is designed to measure and score the implementation of this three-tiered system.

Groups 1, 2, and 3 received training on topics concerning children of poverty. It contained topics that are of particular relevance to that group. Poverty was defined as the absence of any resource that can impact a student’s school or life success. Absent resources may be financial, physical, social, emotional or cognitive. This definition allows educators to cast a broader net beyond the limitations of financial poverty, identifying students from any socio-economic levels who are in need of supports. This perspective also removes any stigmas associated with efforts that focus only on low-income children and families. The training topics included the following: Importance of relationships, Supporting socioemotional skill development, Managing stress in children, Understanding goals of behavior/misbehavior, Increasing hope, and Understanding growth mindsets.

The PDE topics on the Pyramid Model Key Practices are ample and several Groups received training in one or more of them. One school in Group 5 received training on socio-emotional skills, anger management, and problem-solving skills. A different school in the same Group took classes in managing challenging behaviors.
All Groups participated in a PDE session on teaching children problem-solving skills. All 4 K teachers in Group 5 took PDE in strategies for building friendships and for building community. Group 3 teachers participated in PDE on other topics that include understanding challenging behaviors, challenging behavior prevention, problem-solving skills as a response to conflicts, moving beyond time-out, and developing behavior support plans.

**Instrumentation**

The TPOT observations session consists of a 2-h observation and an interview that are scored against a 14 Key Practices List (Hemmeter et al., 2014). The 14 Key Items include indicators in the categories of Schedules, Routines, and Activities, Transitions, Teachers Engage in Supportive Conversations, Promoting Children’s Engagement, Providing Directions, Collaborative Teaming, Teaching Behavior Expectations, Teaching Social Skills, and Emotional Competencies, Teaching Friendship Skills, Teaching Children to Express Emotions, Teaching Problem Solving, Interventions for Children with Persistent Challenging Behavior, Connecting With Families, and Supporting Family Use of Pyramid Model Practice.

The interview questions which follow the TPOT observation give the observer an opportunity to ask the teacher about specific teaching strategies which may or may not have been observed. It also gives the teacher an opportunity to articulate her beliefs about her teaching practices. The trained observer is to follow the script of the interview questions, so there is reliability of the data collected. Some of the questions include:

- Tell me how you teach or help children learn to solve common social problems in the classroom (e.g. when one child has a toy that another child wants or when a child wants a turn on the computer but another child is there).
- How do you individualize instruction around problem-solving for specific children (children who may need extra help)? Please give me examples.
- Tell me how you teach or help children recognize and deal with emotions? Give me some examples of the range of emotions you teach or help children learn. If clarification is needed the interviewer will ask,
  1. What strategies do you use?
  2. What materials do you use?
- What do you do when children have severe and persistent challenging behavior?
  1. What steps do you go through to get support for these children?
  2. What is your role in establishing a behavior plan for these children?

There is also a subscale for Red Flags. The Red Flags subscale includes items for both structural and procedural shortcomings in the observed classroom. Some of the structural items include poorly defined center boundaries and having open
spaces where children have the opportunity to run (Hemmeter et al., 2014). Some of the procedural items included in the Red Flag list include chaotic transitions, scarce interactions between students during activities, and the use of threats with punishments if challenging behavior continues (Hemmeter et al., 2014).

The TPOT measures the fidelity to which educators implement the Pyramid Model in Early Childhood classrooms (Hemmeter et al., 2014). As cited by Hemmeter et al. in 2014, the instrument went under an efficacy trial by Hemmeter et al. in 2011 as well as undergone measures to increase validity and reliability. For this efficacy trial, 40 teachers were either assigned to a business as usual control group or an experimental group. The experimental group received coaching and attended workshops on how to implement the Pyramid Model in their classrooms. The experimental group showed statistically significant improvements. Also cited by Hemmeter, Fox, and Snyder in 2014 are other studies that support the efficacy of the TPOT. In these studies, the TPOT was used to determine if coaching sessions or professional developments made significant differences in teacher performance as it related to Pyramid Model practices (Artman-Meeker and Hemmeter, 2013; Steed and Durand, 2013).

To establish the validity of the TPOT, the creators of the instrument searched for evidence of positive relationships with other instruments, evidence of a shift in scores after a PDE, as well as a link between Pyramid Model practices and child outcomes. The instruments used to investigate whether there was a relationship between variables was the ECERS-R and the CLASS. The ECERS-R had significant relationships with 10 of the 14 TPOT Key Practices. When analyzed with the CLASS instrument, all TPOT Key Practices showed a substantial relationship with CLASS domains (Hemmeter et al., 2014). The researchers used data from the efficacy trial to determine whether PDEs made an impact on Pyramid Model implementation. The analysis showed that the PDEs made a positive difference in all Key Practices except promoting children’s engagement, connecting with families, and supporting family use of the Pyramid Model (Hemmeter et al., 2014). The data from the efficacy trial as well as SSIS scores were also used to investigate the relationship between teacher implementation of Pyramid Model Key Practices and children’s socio skills. The statistical analysis revealed that Pyramid Model Key Practices scores accounted for 12% of the variance in class average socio skills scores (Hemmeter et al., 2014).

The researchers investigated the reliability of the TPOT by examining the amount of random error from raters, the occasions of measurement, and items (Hemmeter et al., 2014). To reduce error in inter-rater scoring, raters undergo a training process that includes videos from master scorers that inform the proper way to score the TPOT. The creators of the TPOT also measured generalizability coefficients across 3 occasions. The scores showed that raters who scored the observed teacher highly on the first occasion also tended to score them highly on the second and third occasions showing a high degree of score stability (Hemmeter et al., 2014). To establish the reliability of the items on the TPOT, the authors investigated the correlations between the key practice items and the Key Practices subscale (Hemmeter et al., 2014). Items and Key Practices tended to be significantly correlated.
Results

The first analysis of the TPOT data was done to determine whether there are significant differences in scores prior to (TPOT1), during (TPOT2), and following (TPOT3) a year of ongoing PDEs. The means for TPOT2 and TPOT1, TPOT3 and TPOT2, and TPOT3 and TPOT1 of each group were compared. To make the comparison a paired-samples t-test was used for each group and for each set of TPOT score differences. Only the sample of teachers that had scores for all three TPOT waves was used (Table 2).

For Group 1, with an N=8, paired-samples t-tests were run to investigate the longitudinal effects of the TPOT PDE intervention. There was a significant difference between the TPOT2 score (M=107.29, SD=7.50) and the TPOT1 score (M=87.57, SD=13.18); t(6) = 8.06, p < 0.001. The data supports the notion that the PDE between TPOT1 and TPOT2 improved scores and teacher performance and the teachers in the Group performed better on the Key Practices. To determine if the PDE program had a significant effect on TPOT outcomes overall, another paired-sample t-test was run on Group 1 teacher’s TPOT3 (M=107.86, SD=3.44) and their TPOT1 score (M=87.57, SD=13.18) which resulted in a statistically significant difference t(6)=4.16, p=0.006. The data suggests that in Group 1 the PDE interventions between TPOT observations had an impact on instrument scores (Table 3).

Similarly, for Group 2, with N=5, there was a significant difference between the TPOT2 score (M=84.6, SD=19.03) and the TPOT1 score (M=65.2, SD=30.13); t(4) = 2.78, p=0.05. These results suggest that the PDE between TPOT scores improved the teachers’ Key Practices scores in Group 2. This was the only significant difference in TPOT scores found in this round of t-tests. The results suggest that the TPOT professional development programming in the Group 2 had little longitudinal effect on Key Practices (Table 4).

The same analysis was conducted with Group 3 with an N=16. A paired-samples t-test was used to compare TPOT scores. There was a statistically significant difference between Group 3’s TPOT2 numbers (M=87.40, SD=5.13) and their TPOT1 numbers (M=75.00, SD=8.75); t(4)=2.78, p=0.05. These results support the premise that

| Group 1 |
|---------|
| TPOT 1 | TPOT 2 | t     | p    |
| M      | SD     | M      | SD   |      |
| 87.57  | 13.18  | 107.29 | 7.50 | 8.06  | <.001 |
| TPOT 2 | TPOT 3 | t     | p    |
| M      | SD     | M      | SD   |      |
| 107.29 | 7.50   | 107.86 | 3.44 | 4.16  | .006  |
| TPOT 1 | TPOT 3 | t     | p    |
| M      | SD     | M      | SD   |      |
| 87.57  | 13.8   | 107.86 | 3.44 | 0.193 | .853  |

Springer
the PDE between the first two TPOT observations had a positive effect. There was a significant difference between Group 3’s TPOT3 (M = 88.2, SD = 8.07) and TPOT1 (M = 75.00, SD = 8.75); t(4) = 5.82, p = 0.004. While there was no significant impact between TPOT3 and TPOT2, the overall impact of the professional development interventions did have a significant longitudinal effect (Table 5).

The last Group that has data for all three waves of TPOT testing is Group 5, with an N = 22, and a paired-samples t-test was conducted to explore significant differences. There was a significant difference between this Group’s TPOT2 scores (M = 88.64, SD = 12.19) and their TPOT1 scores (M = 71.18, SD = 15.66); t(21) = 5.00, p < 0.00. These numbers are evidence that the intervention between the first two TPOTS had a positive effect. The paired sample t-test for TPOT3 (M = 86.00, SD = 16.45) and TPOT1 (M = 71.18, SD = 15.66) resulted in a statistically significant difference; t(21) = 3.30, p = 0.003. The results support the evidence that the professional development program had an effect on TPOT3 scores the following school year (Table 6).

For the next set of analysis paired-samples t-tests were used on each Group once again, this time on the sample of teachers that had completed the first two waves of
TPOT testing. In these analyses the sample size is generally greater as there were more teachers who met that criteria in each group than those who completed all 3 waves. There was a significant difference between Group 1’s TPOT 2 score (M = 106.63, SD = 7.19) and TPOT 1 score (M = 87.75, SD = 12.21); t(7) = 8.28, p < 0.00. These results are evidence to support that the PDE for Group 1 between the TPOT observations had a positive effect on the number of Key Practices observed (Table 7).

For Group 2 there was a significant difference between the TPOT2 score (M = 84.6, SD = 19.03) and the TPOT1 score (M = 65.2, SD = 30.13); t(4) = 2.78, p = 0.05. These results suggest that the PDE between TPOT scores improved the teachers’ scores in Group 2 (Table 8).

The difference in Group 3’s TPOT2 score (M = 90.44, SD = 7.18) and TPOT1 score (M = 75.44, SD = 9.30); t(15) = 6.36, p < 0.00 is significant. It supports the notion that the PDE intervention after the TPOT1 wave impacted positively the Key Practices score on the TPOT2 (Table 9).

Once again, a paired-samples t-test was run to determine the significance in the mean difference between Group 4’s TPOT2 and TPOT1 scores. The analysis shows that there is no significant difference between the Group’s TPOT2 (M = 98.33, SD = 9.00) and TPOT1 (M = 87.33, SD = 7.17); t(5) = 2.011, p = 0.101. These numbers suggest that the PDE intervention had no statistically significant impact on TPOT scores (Table 10).

Lastly, Group 5 showed a significant difference in the means between TPOT2 (M = 86.87, SD = 13.31) and TPOT1 (M = 70.43, SD = 15.80); t(29) = 5.16, p < 0.00.

Next, we conducted a one-way Anova in order to make comparisons between Groups on each iteration of TPOT scores. For TPOT1, an analysis of variance

| Table 5 Group 5 |
|---------------|
| TPOT 1       | TPOT 2       | t       | p   |
| M           | SD           | M       | SD   |
| 71.18       | 15.66        | 88.64   | 12.19 | 5.0   | <0.00 |
| TPOT 2      | TPOT 3       | t       | p   |
| M           | SD           | M       | SD   |
| 88.64       | 12.19        | 86.00   | 16.45 | 0.65  | .521  |
| TPOT 1      | TPOT 3       | t       | p   |
| M           | SD           | M       | SD   |
| 71.18       | 15.66        | 86.00   | 16.45 | 3.30  | .003  |

| Table 6 Group 1 |
|---------------|
| TPOT 1       | TPOT 2       | t       | p   |
| M           | SD           | M       | SD   |
| 87.75       | 12.21        | 106.63  | 7.19  | 8.28  | <0.00 |
showed that the effect of the Group on TPOT1 score was significantly different F(4, 52) = 3.40, \( p = 0.015 \). However, Levene’s Test indicated unequal variance F(4,52) = 3.40, \( p = 0.011 \), the \( p \) value showing that Levene’s null hypothesis is rejected and assumption of homogeneity of variances has been violated. We continued by doing a more robust test of homogeneity of variance, the Brown-Forsythe test, where F(4, 8.41) = 2.45, \( p = 0.127 \) whose \( p \) value is high enough to support the null-hypothesis and fails to reject it, maintaining the assumption that variance is homogenous. Post hoc Tukey HD analysis revealed that there were no significant differences between Groups. Descriptives from the TPOT1 ANOVA analysis are available in Table 11.

In the same way, we conducted a one-way ANOVA analysis to compare TPOT2 scores between Groups. The analysis of variance showed that in the omnibus test there was at least one significant difference between Groups F(4,52) = 5.34, \( p = 0.001 \). Levene’s Test of homogeneity of variance resulted in F(4,52) = 3.06, \( p = 0.024 \). The significant \( p \) value indicated a violation of the assumption of homogeneity of variance. The more robust Brown-Forsythe test also indicated a violation of homogeneity of variance with F(4, 11.4) = 3.061, \( p = 0.023 \). Since the assumption of homogeneity of variance only applies to parametric ANOVA we ran a non-parametric ANOVA test, The Kruskal–Wallis H test, in order to compare TPOT2 scores across Groups. The results indicated that there were 3 pairs of the Groups that had significant results. Group 1 (M = 106.63, SD = 7.20) and Group 2 (M = 84.6, SD = 8.51) showed a significant difference in TPOT2 score with F(4,52) = 26.64

| Table 7 Group 2 | TPOT 1 | TPOT 2 | t    | \( p \) |
|-----------------|--------|--------|------|--------|
| M               | SD     | M      | SD   |        |
| 65.2            | 30.13  | 84.6   | 19.03| 2.78   | 0.05  |

| Table 8 Group 3 | TPOT 1 | TPOT 2 | t    | \( p \) |
|-----------------|--------|--------|------|--------|
| M               | SD     | M      | SD   |        |
| 75.44           | 9.30   | 90.44  | 7.18 | 6.36   | <0.00 |

| Table 9 Group 4 | TPOT 1 | TPOT 2 | t    | \( p \) |
|-----------------|--------|--------|------|--------|
| M               | SD     | M      | SD   |        |
| 87.33           | 7.17   | 98.33  | 9.00 | 2.01   | 0.10  |

| Table 10 Group 5 | TPOT 1 | TPOT 2 | t    | \( p \) |
|------------------|--------|--------|------|--------|
| M                | SD     | M      | SD   |        |
| 70.43            | 15.80  | 86.87  | 13.31| 5.16   | <0.00 |
and the adjusted significance of the non-parametric test of $p = 0.048$. The second significant difference between TPOT2 scores is between the Group 5 schools ($M=88.64$, $SD=2.60$) and Group 1 ($M=106.63$, $SD=7.20$) with $F(4,52) = 24.69$, $p = 0.003$. The last significant difference found in the non-parametric ANOVA was between Group 3 ($M=90.44$, $SD=7.18$) and Group 1 ($M=106.63$, $SD=7.20$) with $F(4,52) = 24.44$, $p = 0.007$. The descriptives from the ANOVA analysis on TPOT2 is available in Table 12.

Lastly, we conducted a One-way ANOVA to compare TPOT3 scores between groups. The omnibus analysis of variance revealed that there was at least one significant difference between the Groups in TPOT3 scores $F(4,52) = 2.96$, $p = 0.028$. Levene’s test showed that the data did not violate the homogeneity of variance assumption $F(4,52) = 1.76$, $p = 0.152$. Tukey HD post hoc analysis showed that there were 2 significant differences between Groups. Group 1 ($M=108.5$, $SD=5.36$) and Group 4 ($M=84$, $SD=12.82$) were also significantly different $F(4,52) = 2.96$, $p = 0.045$. The second statistically significant difference in TPOT3 scores between Groups was between Group 1 ($M=108.5$, $SD=5.36$) and the Group 5 schools ($M=85.86$, $SD=16.41$); $F(4,52) = 2.96$, $p = 0.02$. Descriptive statistics for the TPOT3 are available in Table 13.

**Discussion**

The analysis of the data provides evidence for several conclusions. One, the PDEs between the first two waves of TPOT made a significant difference in teacher outcomes with little exception. Two, despite the lack of significant difference between the TPOT3 and TPOT2 scores, the PDE program produced significant results longitudinally. Lastly, there were few significant differences between Groups for any given wave of TPOT assessment. There were few cases in which any given Group’s teachers showed significantly more improvement in scores on Key Teaching Practices than another Group for each iteration of the TPOT.

The first two statistical analyses that were conducted included paired samples t-tests between TPOT2 and TPOT1 as part of a longitudinal analysis as well as
another paired sample t-test that analyzed the TPOT2 and TPOT1. The sample sizes for the t-tests as part of the longitudinal analysis were lower as only the teachers who had scores for all three waves were included. There was more data for those teachers who only had TPOT2 and TPOT1 scores and not TPOT3. For the longitudinal portion, all Groups: Group 1, Group 2, Group 3, and Group 5 showed significant gains on their TPOT2 scores from their TPOT1. This suggests that the PDE interventions between the assessments had a positive effect on teacher compliance with Pyramid Model Key Practices. The cross-sectional analysis with the data from just teachers who had both TPOT2 and TPOT1 also showed that teachers benefited from the PDE intervention except in one Group. Group 4 did not show a significant increase in TPOT score. However, Group 1, Group 2, Group 3, and Group 5 did. This supports the notion that the vast majority of Groups benefited from the professional development intervention between the TPOT2 and TPOT1 observations.

Despite the fact that there was no lasting effect of the PDE sessions, most groups experienced a significant increase in score longitudinally, across TPOT1, TPOT2, and TPOT3 scores. The support for this claim comes from the paired t-test analysis of the difference between TPOT3 and TPOT1 scores. All Groups had statistically significant results between those waves except Group 2. All Groups, except the aforementioned Group 2, benefitted from the PDE sessions.

| Table 12  | TPOT2 Descriptives |
|-----------|-------------------|
|           | N | Mean | Std dev | Std error | 95% Confidence interval for mean |
|           | Lower bound | Upper bound | Min | Max |
| Group 1   | 8 | 106.63 | 7.19 | 2.54 | 100.61 | 112.64 | 91.00 | 113.00 |
| Group 2   | 5 | 84.6 | 19.03 | 8.51 | 60.97 | 108.23 | 62.00 | 109.00 |
| Group 3   | 16 | 90.44 | 7.18 | 1.80 | 86.61 | 94.27 | 77.00 | 104.00 |
| Group 4   | 6 | 98.33 | 9.00 | 3.68 | 88.88 | 107.78 | 81.00 | 106.00 |
| Group 5   | 22 | 88.64 | 12.19 | 2.60 | 83.23 | 94.04 | 61.00 | 107.00 |
| Total     | 57 | 92.33 | 12.40 | 1.64 | 89.04 | 95.62 | 61.00 | 113.00 |

| Table 13  | TPOT3 Descriptives |
|-----------|-------------------|
|           | N | Mean | Std Dev | Std Error | 95% Confidence Interval for Mean |
|           | Lower Bound | Upper Bound | Min | Max |
| Group 1   | 6 | 108.50 | 5.36 | 2.19 | 102.88 | 114.12 | 99.00 | 114.00 |
| Group 2   | 5 | 82.20 | 26.68 | 11.93 | 49.08 | 115.32 | 37.00 | 104.00 |
| Group 3   | 10 | 87.20 | 13.16 | 4.16 | 77.79 | 96.61 | 66.00 | 108.00 |
| Group 4   | 8 | 84.00 | 12.82 | 4.53 | 73.28 | 94.72 | 55.00 | 96.00 |
| Group 5   | 28 | 85.86 | 16.41 | 3.10 | 79.49 | 92.22 | 49.00 | 110.00 |
| Total     | 57 | 87.89 | 16.86 | 2.23 | 83.42 | 92.37 | 37.00 | 114.00 |
The ANOVA analysis of the TPOT scores revealed that there were few differences between groups. For TPOT1 the post hoc Tukey HD pairwise comparisons showed that there were no significant differences between Groups. Baselines scores were statistically similar. We see from the TPOT2 scores that some Groups had significant differences. Group 1 scored higher than the other Groups. While all the Groups scored significantly higher than their TPOT1 score except Group 4, Dillon is the only Group to perform significantly higher than all the other Groups on the TPOT2. For the TPOT3 test Group 1 scored significantly higher than Group 2 as well as Group 4 and Group 5. There were no other significant differences between Group relationships.

The results of the current study align well with and support the existing research on PDEs and the Pyramid Model. Snyder et al. (2012) stated the importance of PDEs so that teachers can continue their training and give quality socio-emotional instruction. The results of the PDE interventions in this study have shown to improve teaching strategies that support socio-emotional skills. School coaches and administrators used the TPOT data to set goals with teachers, which provided a guide for teachers to continue to make improvements in their socio skill teaching practices. Administrators from all Groups reported that the number of children needing referrals for challenging behaviors had been greatly reduced due to the Pyramid training. Hemmeter et al. (2014) in their efficacy trial for the Pyramid Model showed that PDEs tended to improve young children’s socio-emotional skills. Administrators defined “greatly reduced” in the following ways: some reported there were no referrals to the office at all, others reported that the number of referrals was reduced from seeing one or two daily to one or two each month. One district reported that at one point their referrals were reduced by 60 to 70%. Additionally, administrators reported that they had updated their program policies to move away from practices of suspension or expulsion of preschool children with challenging behaviors. This was an early outcome of our project. All of our partner districts adopted policies to focus on fostering socio-emotional development and responding to challenging behaviors by incorporating positive discipline practices. These findings are especially interesting, not only for teachers who want to use an effective framework for their own classrooms, but also for principals, Head Start administrators, and school Group leaders who seek to implement effective interventions on a broader scale.

Limitations and Recommendations for Future Research

There are several limitations to the current study. The low sample size limits the generalizability of the results and poses problems for statistical power. All three ANOVA analyses failed Levene’s test of homogeneity and needed a more robust Brown-Forsythe test. Even so, the TPOT2 ANOVA failed the more robust Brown-Forsythe and Welch tests and a non-parametric ANOVA was conducted as it does not assume homogeneity. While the pairwise significance levels are acceptable using the non-parametric ANOVA they are not as robust as the standard, parametric ANOVA analysis. The study also presents itself with some threats to validity, mainly
maturation validity threat, history threat, and instrumentation threat. The first two waves of the TPOT, the baseline and the one after the first professional development, take place in the same school year, 2018–2019. The third wave of TPOT observations takes place the following school year, 2019–2020. However unlikely, it could be that the onset of a new school year affected the TPOT scores. Furthermore, there were multiple observers used in this study and while there were measures taken to increase inter-rater reliability there still lies the possibility of inter-rater discrepancies.

The current study gives direction for future research. The next logical step is to replicate the study with a larger sample size. This would provide a broader base to support generalizing the results and would provide more statistical power. The larger sample could also include private early childhood centers making it more representative of the population. Another natural next step would be to include other geographical regions. The sample in the current study allows us to draw conclusions about the sample only. Broadening the sample’s geographical region would, again, make the results more generalizable. The outcome that most PDE sessions had in the current study makes considering these next steps that much more worthwhile.

If future research indicates that these kinds of PDEs get desirable results in a representative sample of classrooms, the potential impact that it could have on young children’s socio-emotional skills nationwide could be considerable.

References

Artman-Meeker, K. M., & Hemmeter, M. L. (2013). Effects of training and feedback on teachers’ use of classroompreventive practices. *Topics in early childhood special education, 33*(2), 112–123.

Baker, J. A., Grant, S., & Morlock, L. (2008). The teacher-student relationship as a developmental context for children with internalizing or externalizing behavior problems. *School Psychology Quarterly, 23*(1), 3–15. https://doi.org/10.1037/1045-3830.23.1.3

Brennan, L. M., Shaw, D. S., Dishion, T. J., & Wilson, M. (2012). Longitudinal predictors of school-age academic achievement: Unique contributions of toddler-age aggression, oppositionality, inattention, and hyperactivity. *Journal of Abnormal Child Psychology, 8*, 1289–1300. https://doi.org/10.1007/s10802-012-9639-2

Bulotsky-Shearer, R. J., & Fantuzzo, J. W. (2011). Preschool behavior problems in classroom learning situations and literacy outcomes in kindergarten and first grade. *Early Childhood Research Quarterly, 26*, 61–73.

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low income children in pre-kindergarten programs. *Early Childhood Research Quarterly, 25*, 166–176.

Chernyshenko, O., Kankaraš, M., Drasgow, F. (2018). Socio and emotional skills for student success and well-being: Conceptual framework for the OECD study on socio and emotional skills, *OECD Education Working Papers*. OECD Publishing, Paris. 173. https://doi.org/10.1787/db1d8e59-en

Chien, N. C., Howes, C., Burchinal, M., Pianta, R. C., Ritchie, S., Bryant, D. M., Clifford, R. M., Early, D. M., & Barbarin, O. A. (2010). Children’s classroom engagement and school readiness gains in pre-kindergarten. *Child Development, 81*, 1534–1549.

Decker, D. M., Dona, D. P., & Christenson, S. L. (2007). Behaviorally at-risk African American students: The importance of student–teacher relationships for student outcomes. *Journal of School Psychology, 45*(1), 83–109. https://doi.org/10.1016/j.jsp.2006.09.004

Dunlap, G., & Fox, L. (2011). Function-based interventions for children with challenging behavior. *Journal of EarlyIntervention, 33*(4), 333–343.
Pee Dee Community Action Partnership (2020). Program overview. https://www.peedeecap.org/menus/program-overview-mission-vision-statement.html

Dunlap, G., Kincaid, D., Horner, R. H., Knoster, T., & Bradshaw, C. P. (2014). A comment on the term “positive behavior support.” Journal of Positive Behavior Interventions, 16, 133–136.

Epstein, A. S. (2009). Me, you, us: Socio-emotional learning in preschool. MI: HighScope Press.

Fox, L., Dunlap, G., Hemmeter, M. L., Joseph, G. E., & Strain, P. S. (2003). The teaching pyramid: A model for supporting socio competence and preventing challenging behavior in young children. Young Children, 58, 48–52.

Hauser-Cram, P., & Woodman, A. C. (2015). Trajectories of internalizing and externalizing behavior problems in children with developmental disabilities. Journal of Abnormal Child Psychology, 44(4), 811–821. https://doi.org/10.1007/s10802-015-0055-2

Hemmeter, M. L., Fox, L., & Snyder, P. (2014). Teaching Pyramid Observation Tool—Research Edition [Manual]. Baltimore, MD: Brookes.

Hemmeter, M. L., Ostrosky, M., & Fox, L. (2006). Socio and emotional foundations for early learning: A conceptual model for intervention. School Psychology Review, 35, 583–601.

Hemmeter, M., Snyder, P., & Fox, L. (2018). Using the Teaching pyramid observation tool (TPOT) to support implementation of socio–emotional teaching practices. School Mental Health, 10(3), 202–213. https://doi.org/10.1007/s12310-017-9239-y

Hemmeter, M. L., Snyder, P. A., Fox, L., & Algina, J. (2016). Evaluating the implementation of the Pyramid Model for promoting socio-emotional competence in early childhood classrooms. Topics in Early Childhood Special Education, 36(3), 133–146. https://doi.org/10.1177/0271121416653386

Jolivette, K., Wehby, J. H., Canale, J., & Massey, N. G. (2001). Effects of choice making opportunities on the behaviors of students with emotional and behavioral disorders. Behavioral Disorders, 26, 131–145.

Jones, D. E., Greenberg, M., & Crowley, M. (2015). Early socio–emotional functioning and public health: The relationship between kindergarten socio competence and future wellness. American Journal of Public Health, 105, 2283–2290.

Joseph, G. E., & Strain, P. S. (2003). Comprehensive evidence-based socio-emotional curricula for young children: An analysis of efficacious adoption potential. Topics in Early Childhood Special Education, 23, 65–76. https://doi.org/10.1177/02711214030230020201

Liew, J. (2012). Effortful control, executive functions, and education: Bringing self-regulatory and socio-emotional competencies to the table. Child Development Perspectives, 6(2), 105–111. https://doi.org/10.1111/j.1750-8606.2011.00196.x

National Association for the Education of Young Children. (2009). Developmentally appropriate practices in early childhood programs serving children from birth through age 8 (Position statement). Retrieved from http://www.naeyc.org/files/naeyc/file/positions/position%20statement%20Web.pdf.

SC School Report Card (2020). https://screportcards.ed.sc.gov/

Snyder, P., Hemmeter, M., Meeker, K., Kinder, K., Pasia, C., & McLaughlin, T. (2012). Characterizing key features of the early childhood professional development literature. Infants & Young Children, 25, 188–212. https://doi.org/10.1097/IYC.0b013e31825a1e7f

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.