Journal of Early Intervention
2020, Vol. 42(2) 143–162

DOI: 10.1177/1053815119872451
journals.sagepub.com/home/jei
SAGE

The Effects of Email Performance-Based Feedback Delivered to Teaching Teams: A Systematic Replication

Erin E. Barton[1], Marina Velez[1], Elizabeth A. Pokorski[1], and Maddisen Domingo[1]

Abstract

A multiple-baseline across behaviors design was utilized to examine the relation between performance-based feedback (PF) and the frequency of teaching teams' use of target behaviors during observation sessions with six teachers in three inclusive preschool classrooms. Teachers increased their use of target behaviors when feedback was provided and maintained their use over time when feedback was withdrawn; however, there was considerable variability across teachers. Furthermore, teachers did not increase their use of target behaviors during covert observations—when they did not know they were being observed. The current study replicates and extends the research in this area by examining the provision of PF to teaching teams and with complex behaviors (e.g., play expansions, promoting social interactions).

Keywords

professional development, preschool children, performance feedback

Replication is a central tenet in educational research for both group (T. D. Cook, Campbell, & Shadish, 2002) and single-case research (Sidman, 1960). Replication is a particularly critical aspect of intervention research, as criteria for evidence-based practices typically require multiple studies (What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education, 2017). Furthermore, replications extend knowledge beyond just reproducing findings to establish evidence-based practices or uncover idiosyncratic, nuanced, or false findings (Banerjee, Movahedazarhouligh, Millen, & Luckner, 2018; Lemons et al., 2016; Makel et al., 2016). Systematic replications also identify new phenomena, which ensure comprehensive and balanced information is available to guide future research and practice (B. G. Cook & Therrien, 2017). Replications are critical for developing a comprehensive understanding of the assets and limitations of tested interventions (Horner et al., 2005). Even unsuccessful replications provide important information regarding external validity. Despite the importance of replication, special

[1]Vanderbilt University, Nashville, TN, USA

Corresponding Author:
Erin E. Barton, Vanderbilt University, 230 Appleton Place, Peabody 228, Nashville, TN 37204, USA.
Email: erin.e.barton@vanderbilt.edu

education intervention researchers rarely stipulate when their studies are replications, and the actual rates of replication are lower than expected (Lemons et al., 2016; Makel et al., 2016).

The research on professional development (PD) has been replicated such that we know providing information alone (e.g., traditional workshops) does not support long-term adult behavior change (Joyce & Showers, 2002). Conversely, adult behavior change *is likely* when ongoing support is provided (Casey & McWilliam, 2011; Fallon, Collier-Meek, Maggin, Sanetti, & Johnson, 2015). In fact, ongoing follow-up with performance-based feedback (PF) is necessary to improve and sustain implementers' use of effective practices and enhance child learning (Metz & Bartley, 2012; Sheridan, Edwards, Marvin, & Knoche, 2009). PD with PF is related to enhanced teacher behaviors, improved classroom quality, and positive child outcomes. Ongoing research on effective PD is critically important given the documented and significant gap between evidence-based intervention practices and Early Intervention/Early Childhood Special Education (EI/ECSE) policies and practices (Hebbeler, Spiker, & Kahn, 2012; Metz & Bartley, 2012; Odom, 2009).

Performance feedback is a ubiquitous component of effective PD approaches (Artman-Meeker, Fettig, Barton, Penney, & Songtian, 2015; Miltenberger, 2012; Snyder, Hemmeter, & Fox, 2015) and has been identified as an evidence-based practice for increasing teachers' implementation of school-based practices for students with disabilities (Fallon et al., 2015; Solomon, Klein, & Politylo, 2012; Stormont, Reinke, Newcomer, Marchese, & Lewis, 2015). For example, Fallon and colleagues found that PF was often delivered verbally and included one of the following: graphical feedback, student data, or problem-solving suggestions. Although they identified PF as an evidence-based practice, they noted considerable variability within their sample, and few of the studies were conducted in early childhood classrooms. Casey and McWilliam (2011) reviewed the research on PF in early childhood exclusively and found PF to be a promising practice. Similar to Fallon and colleagues, Casey and McWilliam noted significant variability across studies regarding PF characteristics and contexts. These authors call for additional research specifying optimal PF characteristics and contexts.

Since the time of these reviews, there has been an increase in the number of studies examining new technologies for providing PD in early childhood settings—often in an effort to improve efficiency (Artman-Meeker et al., 2015). For example, Coogle, Rahn, and Ottley (2015), Coogle, Rahn, Ottley, and Storie (2016), and Coogle, Ottley, Storie, Rahn, and Burt (2017) have conducted several studies successfully using bug-in-ear (BIE) technologies to deliver PF to increase preservice and in-service teachers' use of communication-promoting strategies with preschool children. Similarly, Barton and colleagues have conducted several systematic replications supporting the use of email PF to increase the use of target recommended practices by early childhood teachers (Barton, Chen, Pribble, Pomes, & Kim, 2013; Barton, Fuller, & Schnitz, 2016; Barton, Pribble, & Chen, 2013; Barton & Wolery, 2007). Other researchers also have replicated these findings (Hemmeter, Snyder, Kinder, & Artman, 2011). For example, Artman-Meeker and Hemmeter (2013) found a functional relation between email PF and teaching teams' use of strategies to prevent challenging behaviors. Rathel, Drasgow, Brown, and Marshall (2014) demonstrated that first year teachers provided a greater ratio of positive to negative student communication when provided with email PF. Overall, these replications support the use of email and other technologies as effective delivery methods for PF on discrete verbal behaviors, when sent on the same day as the intervention and when teachers know they are being observed.

Coyne, Cook, and Therrien (2016) recommend designing replications that intentionally vary essential components of previous research, specifically outline the aspects of a study that are being replicated, and report results within a framework of replication. Increased use of replications might ultimately increase the impact of research on practice. Although the aforementioned research clearly demonstrates that PF is effective for supporting early childhood teachers' increased use of discrete verbal behaviors, few studies have documented the relation between PF and the use of complex target behaviors, during covert observations, or with teaching teams

rather than individual teachers (Barton, Rigor, Pokorski, Velez, & Domingo, 2018). Sustained use of complex practices by teachers who are not aware they are being observed is critical for establishing evidence-based implementation practices (Snyder, Hemmeter, & McLaughlin, 2011). Also, early childhood classrooms often have multiple teachers working together within one classroom, yet few studies have examined the impact of email PF on teaching teams who are focused on the same target behaviors and the subsequent impact on child outcomes. Focusing on teaching teams might facilitate collaborative teaching practices, support ongoing use of recommended practices, and produce more robust child outcomes. Furthermore, a limited number of PF studies have examined the relation between early childhood teachers' use of recommended practices and child outcomes, although results from existing research were promising (Fullerton, Conroy, & Correa, 2009; Stormont, Smith, & Lewis, 2007).

The current study replicates and extends early childhood PD research by examining the relation between email PF and teachers who are part of classroom teaching teams use of teaching behaviors during classroom and covert observations. We also measured teachers' use of teaching behaviors in generalization settings, after the termination of email PF, and child challenging behavior across conditions. The following research questions guided the study:

Research Question 1: Does email PF increase levels of teachers' use of team-selected target behaviors during observations in an inclusive preschool classroom?
Research Question 2: Is email PF associated with increased levels of teachers' use of team-selected target behaviors during generalization and covert observations?
Research Question 3: Do increases maintain when email PF ceases?
Research Question 4: Are increases in team-selected target behaviors associated with decreases in child challenging behaviors?

Method

Participants

After obtaining human subjects approval from the appropriate institutional review board, five teaching teams (four dyads and one triad) were recruited for participation. One team withdrew from the study prior to baseline data collection due to scheduling difficulties and staff turnover. Classroom 1 consisted of teaching dyad Courtney (26-year-old White female) and Carmen (24-year-old White female); pseudonyms are used for all participants. Classroom 1 had eight children aged 12 to 24 months including two children with disabilities. Classroom 2 consisted of Tea (22-year-old White female), Tessa (45-year-old Black female), and Tammy (43-year-old White female). Tammy was a first year teacher at the school and requested to be removed from the study due to scheduling issues prior to the start of intervention. Classroom 2 had eight children aged 24 to 36 months including three children with disabilities. Classroom 3 consisted of Rachel (30-year-old White female) and Roya (22-year-old White female). Classroom 3 had 10 children aged 36 to 54 months including three children with disabilities. Carmen, Tea, and Roya were pursuing a master's degree in ECSE at the time of the study; they were working as teaching fellows in the classroom. Courtney had a bachelor's degree in early childhood, Tessa had completed some college, and Rachel had a master's degree in music therapy; all had previous years of paid experience in preschool classrooms. Courtney was a lead teacher; Tessa and Rachel were assistant teachers. The coaches were two doctoral students (one White female; one White/Latina female) and the first author (White female faculty). The coaches had existing professional relationships with some of the teachers. Tessa and Rachel were assistant teachers for one of the coaches during the previous academic year, and Roya had worked with another coach to implement a behavior support plan for a child in her previous classroom; however, none of the coaches

was currently supervising the participants. Also, some of the participants were students of the first author although none were advisees.

Setting

The study was conducted in a university-based inclusive preschool in the southeastern region of the United States. Sessions were conducted in the participating teaching teams' classrooms during free choice time or small group activities (e.g., art, play-doh). Each classroom included a lead teacher, a coteacher, a graduate student assistant, and therapists who provided related services. Child-sized furniture and developmentally appropriate materials were available in each classroom. These included block materials, books, dramatic play materials, musical instruments, science and discovery materials, and water or sand play materials.

Response Definitions and Measurement System

Prior to baseline data collection, we sent participants an online survey via email that asked them to select verbal behaviors they would like to improve via email PF from a list of recommended practices (DEC, 2014). To generate the list, we selected verbal behaviors that might be conducive to email PF and related to improvements in child behaviors in previous research. The options for verbal behaviors included (a) behavior-specific praise, (b) statements to promote social interactions (PSI), (c) reminders of classroom expectations (RCE), (d) play expansions (PE), (e) language expansions (LE), and (f) emotion labels (EL). After the teachers individually selected target behaviors, the coaches selected three target behaviors for each teaching team that both members of the team had selected (see Table 1 for definitions). We identified challenging behavior as the primary child behavior for three reasons: it is (a) ubiquitous in early childhood classrooms given children's development status, (b) contextually bound, and (c) likely to be affected by changes in teachers' use of recommended practices.

The frequency of each teacher's use of the three target verbal behaviors was individually tallied in vivo using an interval-based event recording system using a paper-and-pencil data collection methods. The data collection form had 10 rows of 1-min intervals with separate columns to code each target behavior and the challenging behavior(s) of any children within 5 ft of the teacher. Teacher and child behaviors were measured simultaneously; however, we observed one teacher at a time per classroom. We often observed teachers within the same classroom consecutively.

Interobserver Agreement (IOA)

To estimate IOA on dependent variables, a second observer independently collected data in 40% of sessions across participants and conditions. IOA was calculated by dividing the number of agreements by the number of disagreements plus agreements for each behavior and then multiplying the quotient by 100. An agreement was scored when two observers coded the same frequency of behaviors in a 1-min interval; a disagreement was scored when one observer coded a behavior in an interval and the other observer did not. For example, if one coder scored three play expansions and the other coder scored four play expansions during Interval 2, three agreements and one disagreement were scored for that interval and behavior. IOA averaged 92% across participants, target behaviors, and conditions. Detailed results are provided in Table 2. IOA data were graphed to evaluate potential observer biases and none were identified. During the study, if IOA dropped below 80%, the primary researcher met with and retrained the coder(s); this occurred 4 times.

Table 1. Operational Definitions.

Dependent variable	Definition	Examples	Nonexamples
Reminders of classroom expectations (RCE)	The teacher verbally restates the positively phrased classroom rules or expectations to one child or a group of children. Must be preventive (i.e., not a reaction to challenging behavior), a statement (i.e., not a question), and based on classroom rules.	Teacher says, "Remember we use our inside voices" right before going inside from the playground. Teacher says, "Taking turns on the slide keeps us safe," while transitioning onto the playground.	Teacher says, "We use our inside voices" when children are talking loudly in the hallway. Teacher says, "Stop running and walk."
Promoting social interactions (PSI)	The teacher uses an appropriate and specific verbal prompt to direct the attention or actions of one or more children to another child to elicit a social interaction. Verbal prompts should state exactly what the child can do and cannot be a question.	Teacher tells Brian, "Show Joseph your truck!" Teacher says, "Look how hard Sara is working. Go ask if you can help her."	Teacher says, "Joseph be nice." Teacher says, "Mona, give that toy back."
Play expansions (PE)	The teacher follows the child's lead in play, secures the child attention, and models a play behavior similar to the child's play actions and immediately after the child's play action. The teacher imitates exactly what the child did, adds to the play scheme, and maps language onto her action.	The child says, "tower" and stacks blocks. The teacher stacks blocks and moves a block around the tower saying, "Airplane on the tower!" The child holds a bottle up to the doll. The teacher imitates and rocks the doll back and forth, saying, "Doll is sleepy.	The child stacks two blocks and the teacher stacks two blocks and says "Stack." The child holds a bottle up to the doll. The teacher says "Let's have a picnic" and pretends to eat food.
Language expansions (LE)	The teacher responds to the child's previous utterance and current focus of attention by elaborating through information provided. The teacher repeats exactly (or a close approximation of) what the child said, maintains the child's original intent, and does not require a response from the child.	Child says, "fast!" while moving a plastic fish around the water table, and the teacher says, "Fish is fast!" Child says, "Drive car" while moving a toy car on the ground, and the teacher says, "Drive car to the store."	The child says, "truck" and the teacher says "truck" or "car." The child says, "I want turn" and the teacher says "You want a turn?"
Emotion labels (EL)	The teacher models an emotion-labeling statement or models an emotion in an appropriate situation.	The teacher brought out a new truck; Brian is visibly excited. The teacher says, "Brian, you look so excited to play with our new trucks!"	The child cries and the teacher says, "You're okay."
Child challenging behavior	Any purposeful verbal or nonverbal behavior that is characterized as aggressive, defiant, or disruptive. Includes, but is not limited to, physical aggression toward adults or peers, tantrums, screaming (>3 s), prolonged crying (>30 s), and noncompliance to teacher directive (with secondary indicator of noncompliance).	Child falls to floor, flailing, and crying for 45 s with no breaks in behavior. Teacher tells child to clean up toys from 3 ft away, child says, "No," and continues to play.	Teacher tells child to clean up toys from 6 ft away, but child does not look at teacher nor verbally refuse teacher (i.e., no secondary indicator of noncompliance).

Table 2. Average Percentage IOA and Procedural Fidelity.

IOA by dependent variable	Carmen *M* (range)	Courtney *M* (range)	Tea *M* (range)	Tessa *M* (range)	Rachel *M* (range)	Roya *M* (range)
Reminders of classroom expectations	90 (67-100)	93 (75-100)	93 (70-100)	94 (85-100)	—	—
Promoting social interactions	86 (73-100)	88 (63-100)	—	—	89 (63-100)	84 (67-100)
Play expansions	82 (71—100)	90 (80-100)	—	—	97 (67-100)	91 (64-100)
Language expansions	—	—	84 (70-100)	94 (82-100)	82 (70-100)	91 (63-100)
Emotion labels	—	—	95 (83-100)	92 (77-100)	—	—
Procedural fidelity by condition						
Baseline	100	100	100	100	100	100
Intervention	92.9 (0-100)	100	99.2 (80-100)	97.3 (80-100)	85.4 (0-100)	85.4 (0-100)
Generalization	97.8 (80-100)	83.3 (0-100)	100	100	100	100
Maintenance	100	100	90.9 (0-100)	87.5 (0-100)	100	100
Covert observations	100	100	100	100	100	100

Note. IOA = interobserver agreement.

Experimental Design and Analysis

A multiple-baseline across behaviors design was replicated across six teachers (i.e., three teaching teams) to examine the relation between PF delivered via email and teachers' target behaviors (Gast, Lloyd, & Ledford, 2014). Teachers' use of target behaviors was graphed daily, and visual analysis was used to examine behavior change, make condition change decisions, and identify functional relations. Condition change decisions were based on the performance of both teachers in each team; however, data were recorded and analyzed separately to identify functional relations. Email PF began in the first tier when baseline data demonstrated stability for both teachers; email PF began in subsequent tiers when both teachers demonstrated increased, stable levels of responding in the previous tier. Six data characteristics—level, trend, variability, immediacy, overlap, and consistency—were evaluated within and across conditions (Barton, Lloyd, Spriggs & Gast, 2018). Experimental control was established when an increase in teacher behaviors occurred after email PF commenced and no changes occurred in untreated tiers.

Procedures

Observations were conducted several times per week in each classroom for 10 min. For each observation, the coach confirmed with the participant teacher the time was appropriate to observe (i.e., the teacher was not going to be transitioning the children or going on a break in the next 10 min) and then began a timer. When the timer ended, the coach thanked the participant teacher and quietly left the classroom. A maximum of one classroom observation was conducted per day per teacher; however, generalization or covert observations often occurred on the same day as classroom observations (see Figures 3-8). During observations, the coaches did not provide feedback to the teachers, but did respond to overt social initiations by the children.

Baseline. Following baseline observations, the coach sent each teacher a separate email with a positive statement (e.g., "I really enjoyed watching you play with the children in dramatic play today.") and a request to respond (e.g., "Please let us know you received this email."). The response requests were included to increase our confidence that the participants read the emails.

Intervention. Once baseline stability was established for a dyad and target behavior, the coach sent a five-slide voice-over PowerPoint™ to each member of the participating teaching team

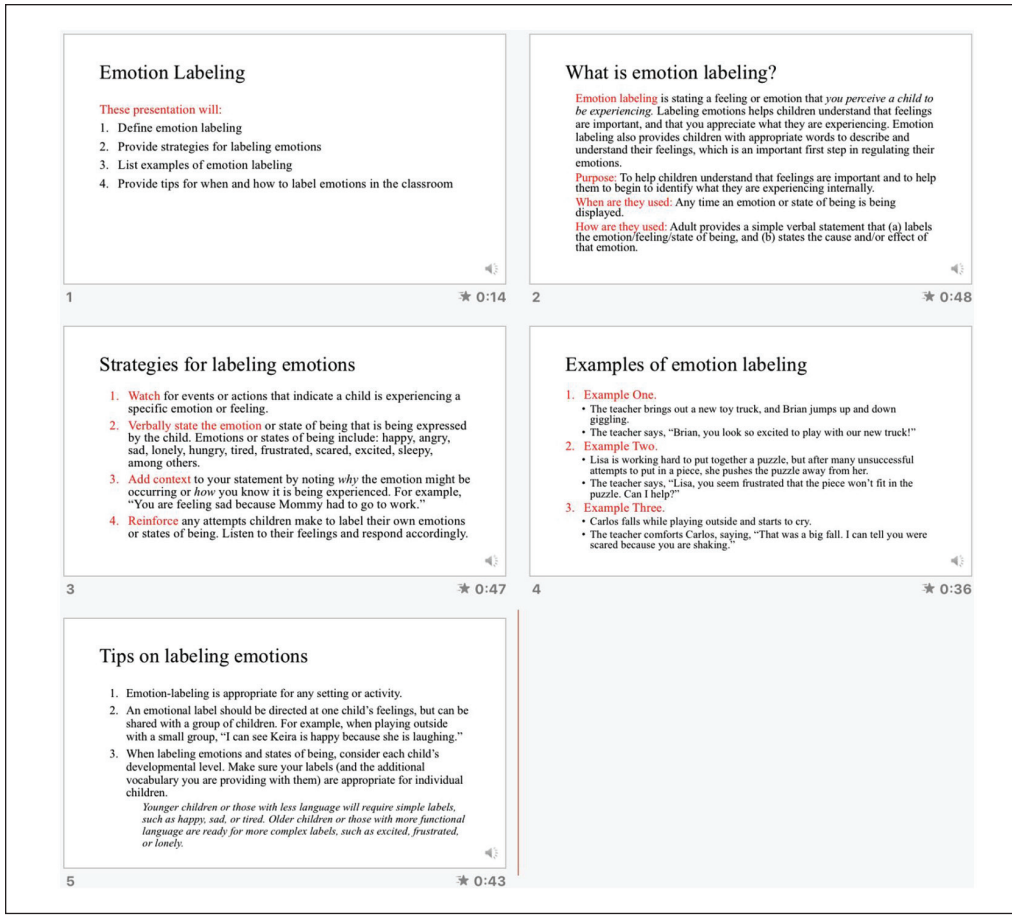


Figure 1. Example presentation slides sent to participants via email.
Note. These slides were designed for the emotion-labeling target behavior.

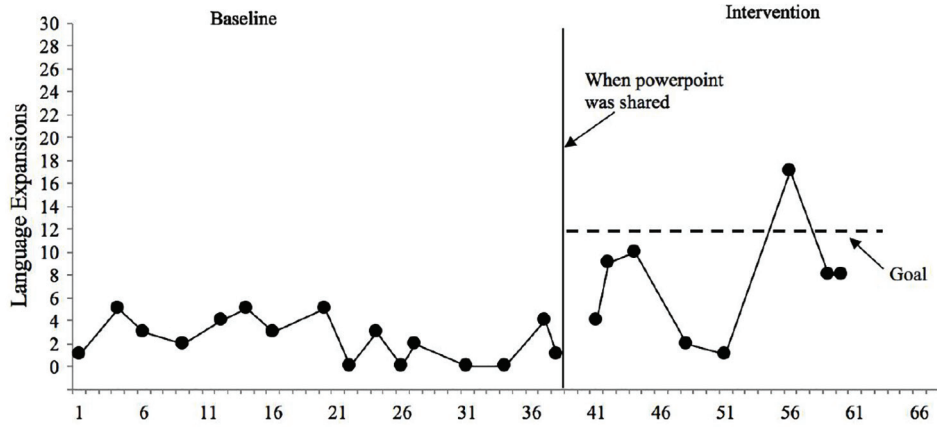
introducing the target behavior (see Figure 1 for an example of these slides for the target behavior EL). The slides consisted of (a) a definition of the target behavior, (b) examples of what the target behavior might look like in the teacher's classroom, and (c) suggestions for specific ways to use the target behavior with the children in their classrooms. Although these slides included similar content across teaching teams, modifications were made as needed based on the developmental level of children in their classroom. We asked the participants to email us after reviewing the slides and commenced PF immediately after receiving this confirmation from both teachers in the dyad. The PF emails were identical to emails sent during baseline conditions with the addition of (a) the frequency count of the teacher's use of the target behavior (and the target behaviors of the previous tiers, when applicable) and (b) two to four verbatim examples of the teacher's use of the target behavior and (c) suggested ways to use the target behavior, which were identified based on observed missed opportunities from each observation (see Figure 2).

Generalization and maintenance. Generalization sessions were conducted on the preschool's outdoor playground or indoor gymnasium. The teachers went to the indoor gym when weather conditions did not permit extended outdoor play. For participants in Classroom 3, generalization sessions were identical to baseline in that the emails did not include PF. For participants in

Hi XXX:

Today you were able to use **8 language expansions** during our observation! Great job! I attached your graph again so you can see how close you are to our goal of 10 per session (~1 to 2 per minute).
Some great examples from today are:

A child said "I wanna build on it" and you said " I wanna build it ON THE SHELF"
A child said "Look" and you said "Look AT MY HEART"



You also did a great job noticing and responding to all the child communication and we understand that having a child with an AAC in a center may require more support from you. Something you can consider is reducing your questions and using language expansions as the response. For example, today when he said "block" you asked "tower?" or "building?" Next time try modeling "block tower" and see if you get the same communicative feel. That shows him you heard block and also gives him a way to expand for next time he wants to build a block structure.

You also were able to promote **10 social interactions** and incorporate **2 play expansions**. I really enjoy coming in your room and watching you interact with the children. Keep up the good work, and I look forward to my next observation!!

Let me know if I can observe you on Thursday morning at the same time.

Thank you!
--Coach A

Figure 2. Example email sent to Roya with graphical feedback.

Classrooms 1 and 2, PF was delivered in generalization emails with the onset of intervention in the classroom setting due to low levels of observed target behaviors. Maintenance sessions were conducted in the classroom following the completion of intervention conditions; emails sent during the maintenance conditions were identical to those sent during baseline (i.e., did not include PF).

Covert observations. For all teaching teams, covert observations were conducted from small observation rooms adjacent to each classroom, each containing a one-way mirror and speaker system; coaches could hear everything said by adults and most of what children said. This allowed coaches to covertly watch classroom activities, such that teachers and children could not see they were being observed. No emails were sent following covert observations. The participants were aware of covert observations but did not know the days or times when they occurred as they changed weekly (see Figures 3-8).

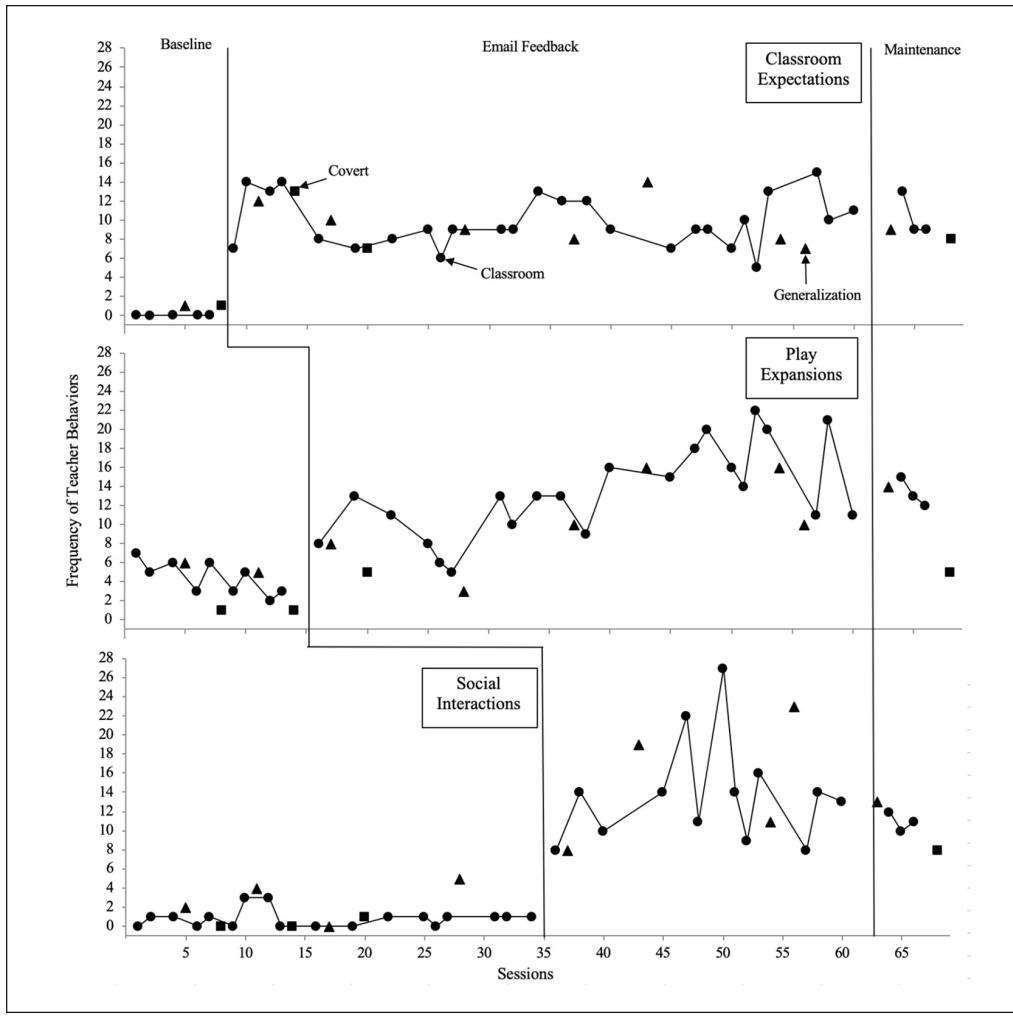


Figure 3. Courtney's use of target behaviors across conditions in Classroom 1.

Procedural adaptations. Three adaptations were made to the intervention. Due to low rates of responding during generalization sessions by teachers in Classrooms 1 and 2, emails with PF were sent following generalization sessions. In addition, due to low rates of responding during classroom observations in Classroom 3, the coaches provided graphical feedback and set a frequency goal for each target behavior in the PF email (PF + G). The PF + G commenced during intervention with LE for Rachel and Roya and included a marked line graph depicting the teacher's frequency of the target behavior and a target goal (see Figure 2).

Procedural Fidelity

Procedural fidelity was measured in 100% of sessions across conditions and participants. This allowed for documentation of both adherence to experimental conditions (i.e., baseline, intervention, maintenance, generalization, and covert observations) and differentiation across conditions (Barton, Meadan-Kaplansky, & Ledford, 2018). A graduate student scored emails sent to participants, verifying that each email (a) was sent on the day of the observation, (b) began with a

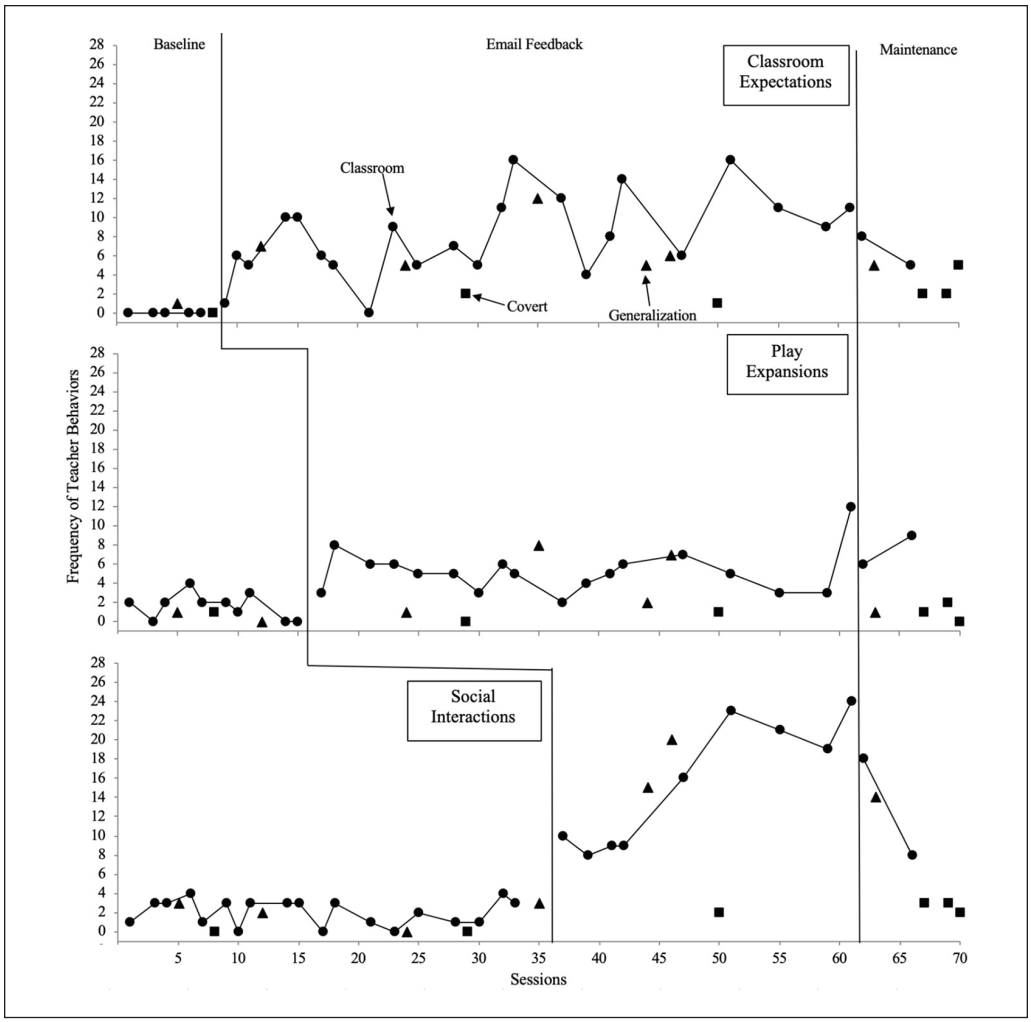


Figure 4. Carmen's use of target behaviors across conditions in Classroom 1.

positive statement, (c) included a frequency count for the target behavior (and the behaviors of previous tiers), (d) included examples of the target behaviors observed, and (e) included a request for response. Covert observations were scored as 100% correct if no email was sent. Average procedural fidelity across participants and conditions was 98% (see Table 2). Fidelity errors occurred in emails following four observations, in that coaches did not send an email, did not include a request for response, or did not include examples of target behaviors. A coach inadvertently sent one email without PF after a covert observation.

Results

Classroom Observations

Target teacher data across all classroom, generalization, and covert observations are presented by individual teacher participant in Figures 3 to 8. A functional relation was identified between email PF and target teacher behaviors, which was replicated across four teachers (Carmen,

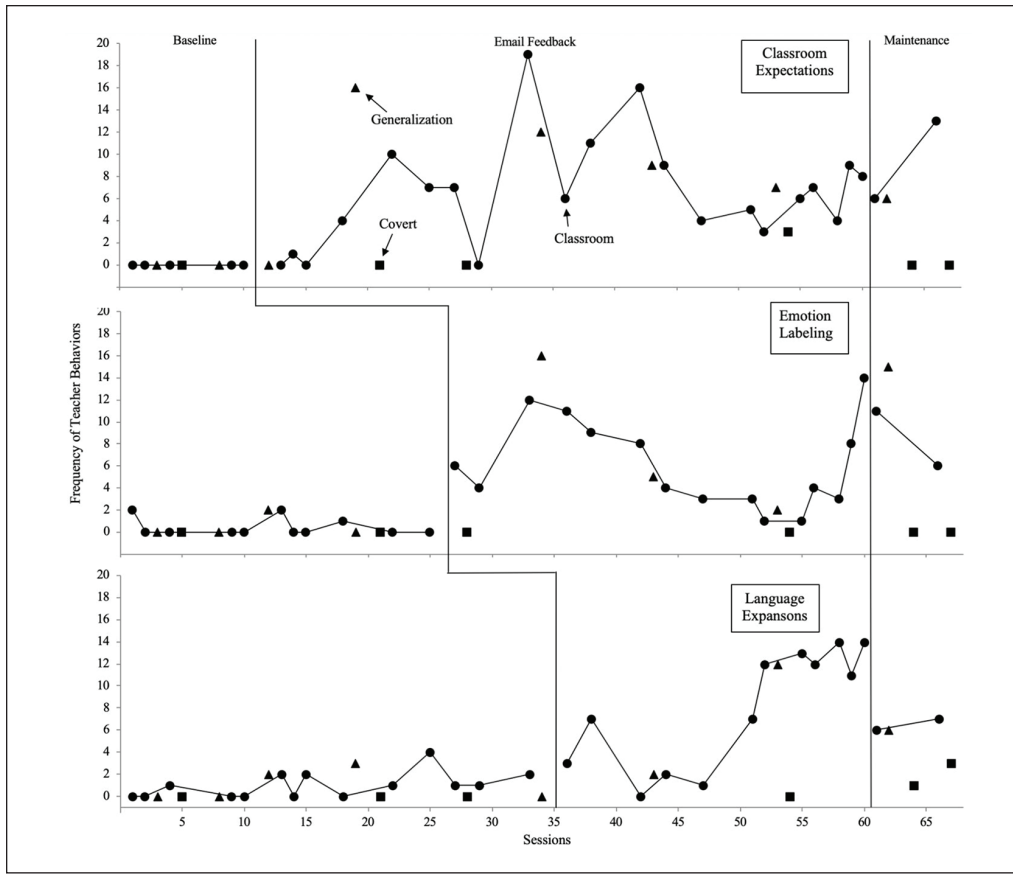


Figure 5. Tea's use of target behaviors across conditions in Classroom 2.

Courtney, Tessa, and Tea). A functional relation was identified between email PF + G and target behaviors for Rachel. Roya demonstrated behavior change with email PF, but considerable variability in her data limits confidence in the results. Target teacher behaviors maintained at or slightly below intervention levels for five teachers; maintenance was not measured for Roya.

Classroom 1. Carmen and Courtney demonstrated low and stable use of the three target behaviors during baseline (range = 0-7). Neither teacher used RCE during baseline. When email PF was provided for RCE, Carmen had an increasing trend for the initial four sessions; however, her level of RCE decreased when PF started with the second tier (PE). RCE remained at levels higher than baseline with some variability for the remainder of intervention (range = 0-16). Courtney's RCE had an immediate increase in level with minimal variability (range = 7-15) and no overlap with baseline. During baseline for PE, Carmen and Courtney demonstrated low levels of the behavior (range = 0-4 and range = 2-7, respectively). Both teachers increased their use of PE immediately following the start of email PF. Although Carmen's level of PE was low and stable for the remainder of intervention (range = 2-12), Courtney's had some variability with an increasing trend (range = 5-20). Courtney had minimal overlap with baseline condition (n = 2, 9%); Carmen had a moderate amount of overlap (n = 6, 33%). PSI were low and stable during baseline (range = 0-4) for both Courtney and Carmen. They increased their use of PSI immediately following commencement of email PF. Carmen had an increasing

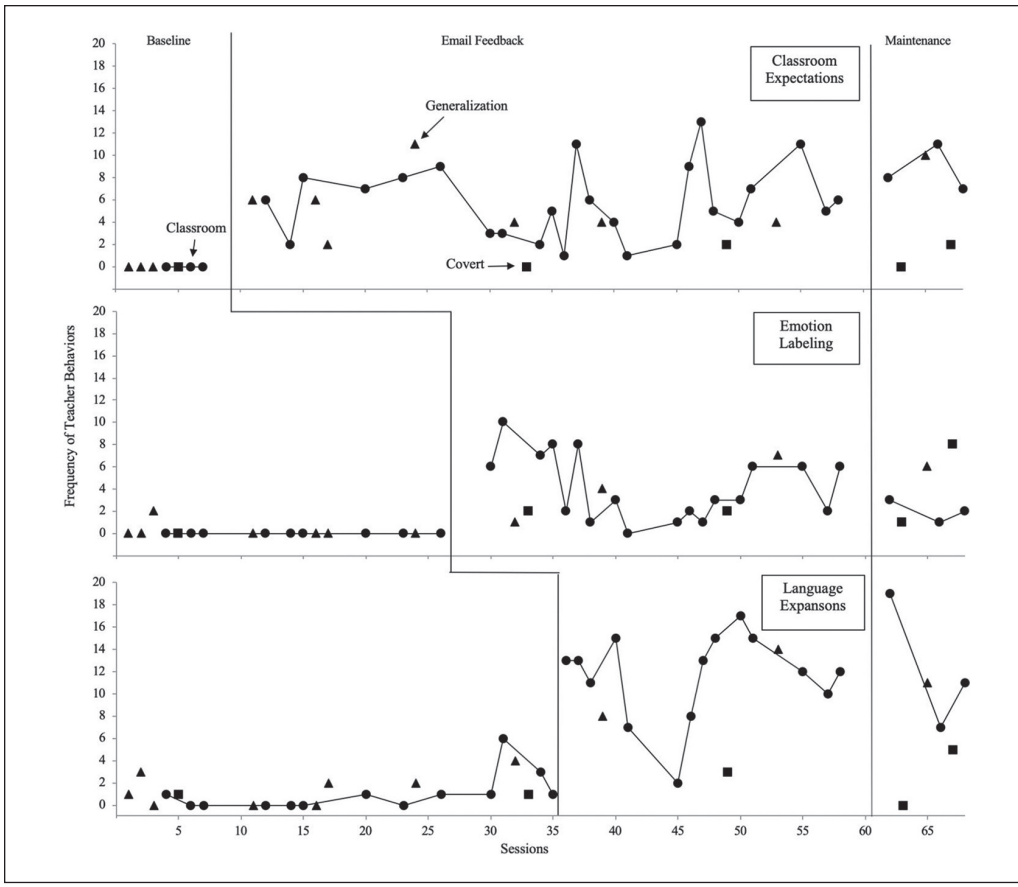


Figure 6. Tessa's use of target behaviors across conditions in Classroom 2.

trend (PSI range = 8-24), and Courtney had some variability during intervention (PSI range = 8-27). Neither teacher had overlapping data between baseline and intervention conditions. Courtney maintained intervention levels for all three target behaviors when PF ceased; Carmen had a decreasing trend.

Classroom 2. Neither Tea nor Tessa used RCE during baseline. Upon the introduction of email PF, Tessa demonstrated an immediate increase in level, but considerable variability for the remainder of intervention and no overlapping data with baseline. Tea minimally increased her use of RCE after three email PF sessions with considerable variability (range = 1-19), with minimal overlap with baseline (n = 3, 14%). During baseline, Tea and Tessa used few EL (range = 0-2). Both teachers demonstrated an immediate increase in EL with email PF, with some variability. Tea had a range of 1 to 14 EL and minimal overlap with baseline (n = 2, 13%). Tessa had a range of 0 to 8 EL and also had minimal overlap with baseline (n = 1, 5%). During baseline, Tea used few LE (range = 0-4). Tessa's use of LE was low and stable except for an increase when email PF started in the previous tier. Her use of LE decreased immediately prior to the introduction of email PF (range = 0-6). Both teachers had an immediate increase in LE with email PF followed by considerable variability (Tea's range = 0-14; Tessa's range = 1-17) and minimal overlap (Tea: n = 4, 33%; Tessa: n = 1, 7%). With the removal of PF, Tessa's use of LE

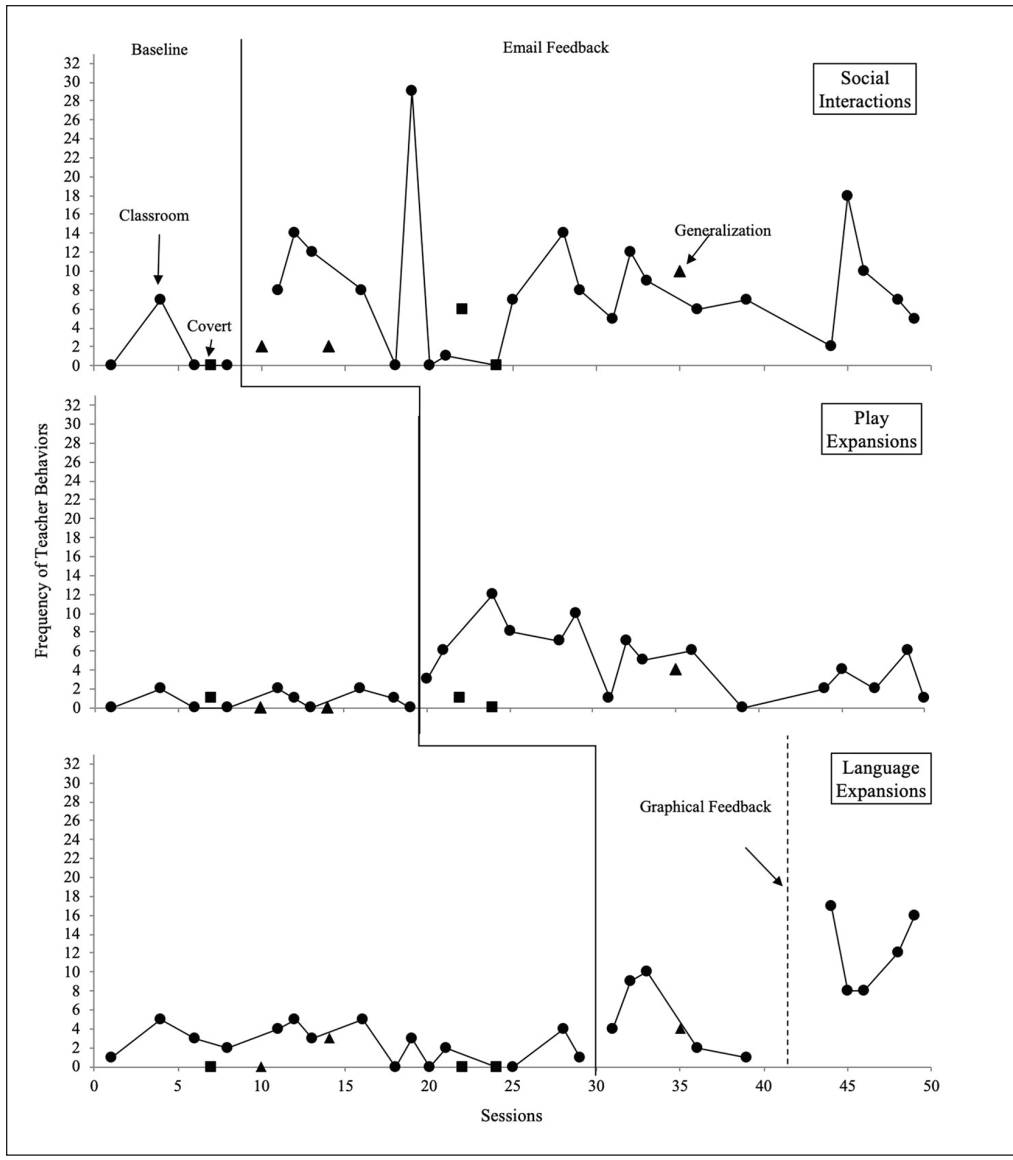


Figure 7. Rachel's use of target behaviors across conditions in Classroom 3.

immediately increased and eventually returned to intervention levels (range = 7-19). Tea's use of LE decreased to 6 and 7, respectively, for two maintenance observations.

Classroom 3. Both Rachel and Roya used few PSI during baseline, but had an immediate increase once email PF was introduced. Following this initial increase, PSI were variable for both teachers with some overlap with baseline (range = 0-29 for Roya; range = 0-20 for Rachel). Similarly, Rachel and Roya had low levels of PE during baseline (range = 0-2 for Rachel and 0-7 for Roya). Both teachers increased their use of PE immediately with email PF. However, during subsequent observations they had considerable variability (range = 0-12). Roya and Rachel also had low levels of LE during baseline (range = 0-6). When intervention began, both teachers increased

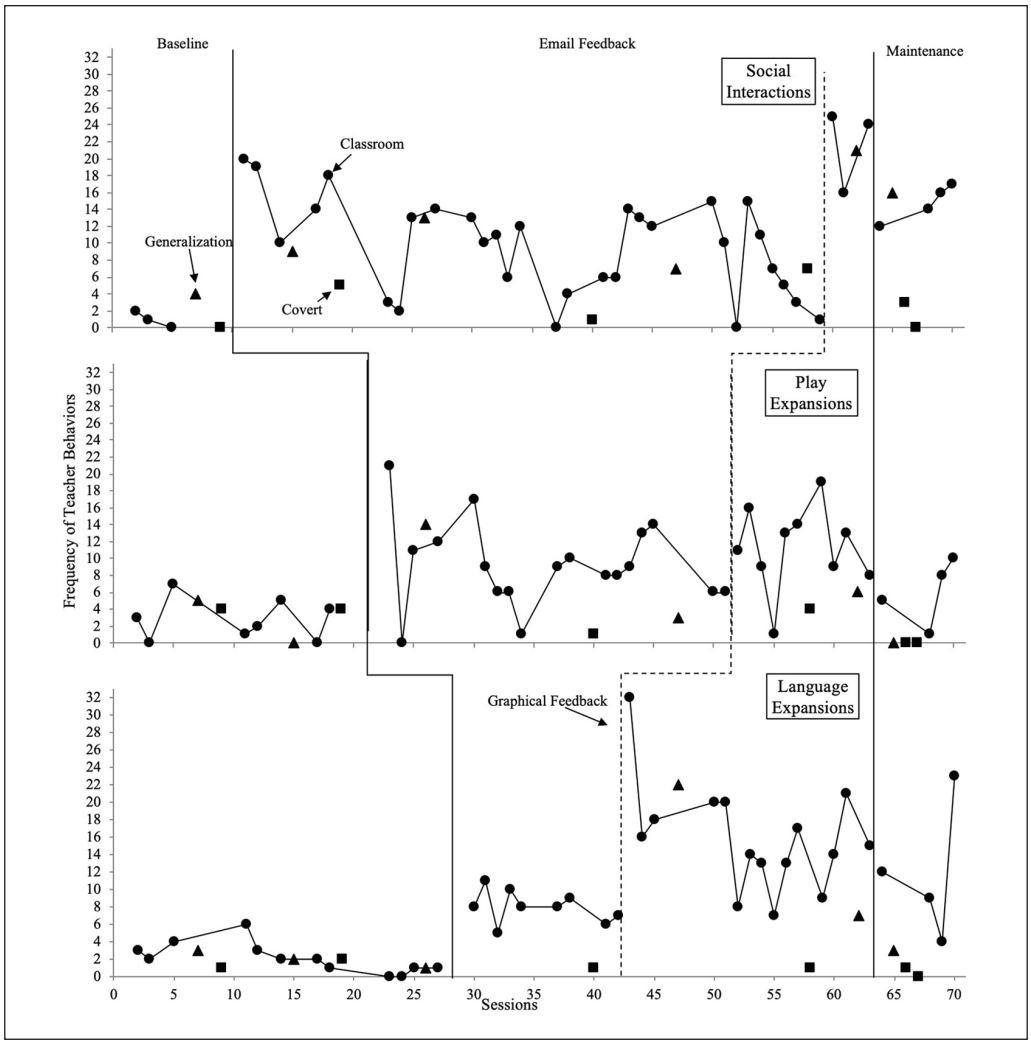


Figure 8. Roya's use of target behaviors across conditions in Classroom 3.

Table 3. Average Frequency of Child Challenging Behavior per Session.

Setting	Carmen *M* (range)	Courtney *M* (range)	Tea *M* (range)	Tessa *M* (range)	Rachel *M* (range)	Roya *M* (range)
Classroom	.7 (0-1)	0	.2 (0-1)	.1 (0-1)	.5 (0-4)	.1
Generalization	0	.2 (0-1)	0	.4 (0-3)	.3 (0-1)	.2 (0-1)
Covert observations	0	0	0	0	0	.3 (0-2)

their use of LE. However, levels remained low for both teachers, and Rachel had some overlap with the baseline condition. When graphical feedback was added to the email PF, both teachers increased their use of LE to a level higher than baseline. Roya's use of PE remained variable with PF + G. She had an immediate increase in PSI with PF + G, which stabilized at a level higher than the previous condition (i.e., email PF alone). Roya's use of PSI, PE, and LE did not maintain

when PF + G ceased but ended with an increasing trend. Rachel resigned from her position after the 26th observation and maintenance data were not collected.

Covert and Generalization Observations

Target teacher behaviors during covert sessions were low and stable during all baseline conditions. However, teachers' use of the target behaviors did not increase with email PF (nor with graphical feedback for Roya or Rachel) during covert observations with one exception. Courtney increased her use of RCE and PE to levels similar to classroom observations, which were maintained. However, the limited number of sessions restricts interpretations of the results.

Levels of target teacher behaviors during the generalization sessions were low and stable across all teachers with the introduction of email PF and maintained when email PF was removed. All teachers demonstrated notable increases in at least one target behavior during generalization sessions when feedback was provided at levels similar to classroom observations. However, Carmen's use of PE remained low, Tea had a slight decrease in RCE and LE, and Rachel's use of PE and LE returned to baseline levels during generalization sessions.

Child Challenging Behavior

Children exhibited low levels of challenging behavior, with no conditional variation, across classrooms and throughout the study (Table 3; range = 0-3). No functional relation was identified between challenging behavior and email PF. Graphed data are available via email.

Discussion

In this study, we replicated previous research by examining the use of email PF on target verbal behaviors with early childhood teachers in inclusive classrooms. We used an email PF format from previous research, teachers with similar characteristics as those from previous research (i.e., both preservice and in-service teachers), and examined discrete verbal behaviors included in previous studies. Our replication systematically varied the following essential components: (a) we focused on behaviors selected by teaching teams, (b) we examined novel and understudied target behaviors (i.e., PE, PSI, EL, and RCE), (c) we examined outcomes in covert observations when teachers were not aware they were being observed, and (d) we included brief, self-guided training (i.e., five PowerPoint™ slides with voice-over) to describe the target behaviors.

Our findings extend current knowledge in several ways. First, we hypothesized that providing the same feedback to members of a classroom teaching team and introductory PowerPoint™ slides would increase responsiveness to feedback. However, the data patterns show similar latency to change and similar or lower levels of the target behaviors with more variability when compared with previous studies (Barton, Ledford, et al., 2016; Barton, Pribble, & Chen, 2013). This might be due to the specific target behaviors included in our study, which might be more complex and difficult to consistently implement when compared with target behaviors in previous studies. For example, play expansions require the teacher attend to the child's play behavior, imitate it, *and* model a new behavior. This might be more difficult than simply repeating what the child says and adding to it (i.e., language expansions; Barton, Pribble, & Chen, 2013) or providing descriptive praise (Hemmeter et al., 2011). Furthermore, play expansions also require the child to engage with toys, which might not consistently occur.

Second, we examined several target behaviors that have been infrequently studied in previous PF research. These include PSI, PE, EL, and RCE. Although all teachers increased their use of these behaviors after receiving the introductory PowerPoint™ slides and email PF, we had higher rates of variability in the current study than noted in previous research (Barton, Fuller, & Schnitz,

2016; Barton, Rigor, et al., 2018). For several teachers, levels of target behaviors decreased when email PF on a new target was added. For example, Courtney's and Carmen's use of RCE decreased slightly when email PF on PE commenced. Similarly, Tessa's and Tea's use of PE notably decreased when email PF on EL commenced. Although levels of target behaviors eventually recovered for each teacher and target behavior, the decreases suggest it might be important to focus on one target at a time rather than simultaneously. For example, Hemmeter and colleagues (2011) demonstrated increased levels and sustained change in teachers' use of descriptive praise with email PF, but email PF focused on only one behavior—descriptive praise. This might be a behavior-specific phenomenon and particularly true for more complex behaviors, which should be examined in future replications.

Third, we included covert observations of teacher behaviors. Rates of target behaviors did not consistently increase during the covert observations or during generalization observations without feedback. This might be because the target behaviors were generally too complex to use consistently or simultaneously across settings. Conversely, the observer could have functioned as a discriminant stimulus, that is, the teachers used the target behaviors only when they knew it would result in email PF. Although this might indicate that teachers did not acquire a robust repertoire of the target behaviors, it also might indicate that teachers did not have the same expectation—that they should use the target behaviors consistently—across settings. The lack of generalization replicates findings from previous research (Barton, Ledford, et al., 2016; Barton, Pribble, & Chen, 2013; Barton, Rigor, et al., 2018) and suggests the impact of email PF might not generalize without intentional supports in those settings. Future research should continue to examine teacher use of target behaviors across settings, including those in which the observer is concealed, when expectations have been clearly set for teacher's use of target behaviors across settings. Additional replications might examine the use of more resource-intensive PF such as might be provided through live or BIE coaching. For example, Coogle et al. (2015) used BIE coaching to support preservice teachers' use of communication strategies within an activity-based intervention approach. In this case, preservice teachers were taught to use new, complex strategies. However, these more intensive technologies should be reserved for teachers or target behaviors that require more support. Future replications should continue to examine hypotheses related to PF such that evidence-based practices can be matched to the individual, outcomes, and available resources.

Fourth, we demonstrated that email PF was related to increased use of five different recommended practices across four teachers in two different inclusive preschool classrooms. These replications are important for establishing PF as an evidence-based practice for early childhood teachers. We also advanced current knowledge by showing email PF alone did not work for two teachers, Rachel and Roya. However, when graphical feedback and goal setting were added to the email PF (PF + G), target behavior(s) increased. These results demonstrate that minor, low-effort adaptations to the email PF can be effective for nonresponders. Rachel's data provide strong support for these adaptations given her levels of target behaviors maintained when PF + G ceased. Furthermore, we maintained experimental control by staggering the introduction of PF + G across tiers and showing immediate behavior change after PF + G was introduced.

Finally, at least two of the three target behaviors across teaching teams required specific child behaviors. The observed variability might be due, at least in part, to the children the teachers interacted with during each observation which was not controlled. For example, Classrooms 1 and 2 comprised infants and toddlers. In Classroom 1, the children had heterogeneous play repertoires; the teacher's ability to use PE might have depended on which children she was interacting with during the observation. Similarly, in Classroom 2 the children had heterogeneous communication skills, and the teacher's ability to use LE might have depended on which children she was talking to during the observation. Also, PSI required the children were *not* already socially interacting. Thus, if the teachers were in a center with children who were socially

interacting, opportunities for PSI were limited. Although this heterogeneity is typical of inclusive early childhood classrooms, a more precise measure of behavior change might be the ratio of emitted skills to the total opportunities to use skills.

Limitations

There are several limitations. First, coaches made some PF errors with Rachel and Roya (i.e., the coach did not send the email on the same day, include a clear request for response, or include examples of target behaviors). This likely did not affect data as the errors were minimal and the teachers consistently received email PF. The presence of PF errors and the subsequent increase in target behaviors despite the errors increase the ecological validity of the study. Supervisors and classroom coaches might not always be able to send emails on the same day of the observation, resulting in inconsistent PF. Also, the participants did not always respond to emails. Future replications should examine strategies for ensuring teachers regularly check email when receiving email PF and other feasible forms of PF delivery (e.g., text messaging, phone calls). Second, we also had several sessions with low IOA. This was likely due to observing three target behaviors—many of which also required monitoring child behaviors—simultaneously during live observations. Our IOA averages meet design standards and we are confident that our results are reliable and accurate. We graphed and reviewed results across both observers and ensured systematic bias was unlikely. Third, we did not directly address the social validity of the outcomes, goals, or procedures. Each of these has been shown to have strong social validity in previous studies (Barton, Rigor, et al., 2018), but should continue to be examined in future research. Finally, the coaches had existing professional relationships with some of the teachers; however, the consenting process was carefully designed and conducted to reduce the likelihood that they felt pressure to participate in the study. For example, the school director carefully explained that their participation was optional and did not affect their employment or their status in the graduate program.

Future Replications

The current findings should be replicated in the following ways. First, the use of graphical feedback and goal setting increased levels of target behaviors across two teachers in one teaching team. These relatively low-effort strategies might represent important ways to supplement or enhance email PF. Graphical feedback, although not studied extensively, is a promising practice in early childhood settings (Casey & McWilliam, 2008). Goal setting, however, has been shown to be effective in a variety of settings and populations (Epton, Currie, & Armitage, 2017). Research should continue to examine these strategies for increasing teacher's use of recommended practices in early childhood settings. Second, research should systematically examine the use of PF on more complex target behaviors. For example, perhaps supporting teachers in using PE independent of other behaviors might result in more effective and efficient learning. Similarly, PSI and PE might initially require more intensive in vivo coaching, and PF can be used to support teachers' maintained and generalized use of these or other recommended practices. Finally, we did not discern notable changes in children's challenging behaviors. This might have been an artifact of the measurement system (i.e., it was not adequately precise) or the duration of the observations (i.e., they were not long enough to detect change). However, we anecdotally noted changes in child behaviors directly associated with changes in target teacher behaviors. For example, we noted increases in children's social interactions in Classroom 1 as Carmen and Courtney increased their use of PSI. We also noted children talking about their emotions more often when Tessa and Tea used EL. Although the target behaviors were selected based on their known relation to positive child outcomes, future replications should examine changes in child behaviors directly related to the target teacher behaviors (e.g., social interactions, PSI).

Furthermore, future replications might measure child behaviors during *and after* teacher observations to measure immediate *and delayed* changes in child behaviors associated with improvements in teacher behaviors.

Conclusion

The current study was a systematic replication of previous research examining the efficacy of email PF. We extended previous research to demonstrate the efficacy of PF on teaching teams' use of verbal behaviors. We also examined several target behaviors that have been infrequently studied, which is important for understanding ideal contexts and targets for using email PF. Although PF has been identified as an evidence-based practice, continued replications are needed to understand efficient delivery methods, behaviors amenable to change using PF, and the impact of PF on child learning.

Acknowledgments

The authors would like to acknowledge the children, families, teachers, and staff of the Susan Gray School and the Special Education graduate students at Peabody College whose participation and support made this study possible.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Preparation of this article was supported in part by the U.S. Department of Education, Office of Special Education and Rehabilitative Services, Grant H325K140110. However, the opinions expressed do not necessarily reflect the policy of the U.S. Department of Education and no official endorsement should be inferred.

References

Artman-Meeker, K., Fettig, A., Barton, E. E., Penney, A., & Songtian, Z. (2015). Applying an evidence-based coaching framework to the Early Childhood Professional Development literature. *Topics in Early Childhood Special Education*, *35*, 183-196.

Artman-Meeker, K. M., & Hemmeter, M. L. (2013). Effects of training and feedback on teachers' use of classroom preventive practices. *Topics in Early Childhood Special Education*, *33*, 112-123.

Banerjee, R., Movahedazarhouligh, S., Millen, K., & Luckner, J. L. (2018). Building evidence in early childhood special education: A systematic review of replication intervention studies. *Topics in Early Childhood Special Education*, *37*, 246-256.

Barton, E., Chen, C., Pribble, L., Pomes, M., & Kim, Y. (2013). Coaching preservice teachers to teach play skills to children with disabilities. *Teacher Education and Special Education*, *36*, 330-349.

Barton, E., Fuller, E., & Schnitz, A. (2016). The use of email to coach preservice early childhood teachers. *Topics in Early Childhood Special Education*, *36*, 78-90.

Barton, E., & Wolery, M. (2007). Evaluation of email feedback on the verbal behaviors of pre-service teachers. *Journal of Early Intervention*, *30*, 55-72.

Barton, E. E., Ledford, J. R., Lane, J. D., Decker, J., Germansky, S. E., Hemmeter, M. L., & Kaiser, A. (2016). The iterative use of single case research to advance the science of EI/ECSE. *Topics in Early Childhood Special Education*, *36*, 4-14.

Barton, E. E., Lloyd, B. P., Spriggs, A. D., & Gast, D. (2018). Visual analysis of graphic data. In J. R. Ledford & D. L. Gast (Eds.), *Single case research in behavioral sciences* (3rd ed., pp. 179–214). New York, NY: Routledge.

Barton, E. E., Meadan-Kaplansky, H., & Ledford, J. R. (2018). Independent variables, fidelity, and social validity. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (pp. 133-156). New York, NY: Routledge.

Barton, E. E., Pribble, L., & Chen, C. I. (2013). The use of e-mail to deliver performance-based feedback to early childhood practitioners. *Journal of Early Intervention*, *35*, 270-297.

Barton, E. E., Rigor, M., Pokorski, E. A., Velez, M. S., & Domingo, M. (2018). Using text messaging to deliver performance feedback to preservice early childhood teachers. *Topics in Early Childhood Special Education*, *39*, 88-102.

Casey, A. M., & McWilliam, R. A. (2008). Graphical feedback to increase teachers' use of incidental teaching. *Journal of Early Intervention*, *30*, 251-268.

Casey, A. M., & McWilliam, R. A. (2011). The characteristics and effectiveness of feedback interventions applied in early childhood settings. *Topics in Early Childhood Special Education*, *31*, 68-77.

Coogle, C. G., Ottley, J. R., Storie, S., Rahn, N. L., & Burt, A. K. (2017). ECoaching to enhance special educator practice and child outcomes. *Infants & Young Children*, *30*, 58-75.

Coogle, C. G., Rahn, N. L., & Ottley, J. R. (2015). Pre-service teacher use of communication strategies upon receiving immediate feedback. *Early Childhood Research Quarterly*, *32*, 105-115.

Coogle, C. G., Rahn, N. L., Ottley, J. R., & Storie, S. (2016). ECoaching across routines to enhance teachers' use of modeling. *Teacher Education and Special Education*, *39*, 227-245.

Cook, B. G., & Therrien, W. J. (2017). Null effects and publication bias in special education research. *Behavioral Disorders*, *42*, 149-158.

Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Coyne, M. D., Cook, B. G., & Therrien, W. J. (2016). Recommendations for replication research in special education: A framework of systematic, conceptual replications. *Remedial and Special Education*, *37*, 244-253.

Division for Exceptional Children. (2014). DEC recommended practices in early intervention/early childhood special education 2014. Retrieved from http://www.dec-sped.org/recommendedpractices

Epton, T., Currie, S., & Armitage, C. J. (2017). Unique effects of setting goals on behavior change: Systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*, *85*, 1182-1198.

Fallon, L. M., Collier-Meek, M. A., Maggin, D. M., Sanetti, L. M., & Johnson, A. H. (2015). Is performance feedback for educators an evidence-based practice? A systematic review and evaluation based on single-case research. *Exceptional Children*, *81*, 227-246.

Fullerton, E. K., Conroy, M. A., & Correa, V. I. (2009). Early childhood teachers' use of specific praise statements with young children at risk for behavioral disorders. *Behavioral Disorders*, *34*, 118-135.

Gast, D. L., Lloyd, B. P., & Ledford, J. R. (2014). Multiple baseline and multiple probe designs. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (3rd ed., pp. 239-281). New York, NY: Routledge.

Hebbeler, K., Spiker, D., & Kahn, L. (2012). Individuals with Disabilities Education Act's early childhood programs: Powerful vision and pesky details. *Topics in Early Childhood Special Education*, *31*, 199-207.

Hemmeter, M. L., Snyder, P., Kinder, K., & Artman, K. (2011). Impact of performance feedback delivered via electronic mail on preschool teachers' use of descriptive praise. *Early Childhood Research Quarterly*, *26*, 96-109.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*, 165-179.

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development*. Alexandria, VA: Association for Supervision & Curriculum Development.

Lemons, C. J., King, S. A., Davidson, K. A., Berryessa, T. L., Gajjar, S. A., & Sacks, L. H. (2016). An inadvertent concurrent replication: Same roadmap, different journey. *Remedial and Special Education*, *37*, 213-222.

Makel, M. C., Plucker, J. A., Freeman, J., Lombardi, A., Simonsen, B., & Coyne, M. (2016). Replication of special education research: Necessary but far too rare. *Remedial and Special Education*, *37*, 205-212.

Metz, A., & Bartley, L. (2012). Active implementation frameworks for program success: How to use implementation science to improve outcomes for children. *Zero to Three*, *32*, 11-18.

Miltenberger, R. G. (2012). Behavior skills training procedures. In R. G. Miltenberger (Ed.), *Behavior modification: Principles and procedures* (5th ed., pp. 217-235). Belmont, CA: Wadsworth.

Odom, S. L. (2009). The tie that binds: Evidence-based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education*, *29*, 53-61.

Rathel, J. M., Drasgow, E., Brown, W. H., & Marshall, K. J. (2014). Increasing induction-level teachers' positive-to-negative communication ratio and use of behavior-specific praise through emailed performance feedback and its effects on students' task engagement. *Journal of Positive Behavior Interventions*, *16*, 219-233.

Sheridan, S. M., Edwards, C. P., Marvin, C. A., & Knoche, L. L. (2009). Professional development in early childhood programs: Process issues and research needs. *Early Education and Development*, *20*, 377-401.

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. New York, NY: Basic Books.

Snyder, P. A., Hemmeter, M. L., & Fox, L. (2015). Supporting implementation of evidence-based practices through practice-based coaching. *Topics in Early Childhood Special Education*, *35*, 133-143.

Snyder, P. A., Hemmeter, M. L., & McLaughlin, T. (2011). Professional development in early childhood intervention: Where we stand on the silver anniversary of PL 99-457. *Journal of Early Intervention*, *33*, 357-370.

Solomon, B. G., Klein, S. A., & Politylo, B. C. (2012). The effect of performance feedback on teachers' treatment integrity: A meta-analysis of the single-case literature. *School Psychology Review*, *41*, 160-175.

Stormont, M. A., Reinke, W. M., Newcomer, L., Marchese, D., & Lewis, C. (2015). Coaching teachers' use of social behavior interventions to improve children's outcomes: A review of the literature. *Journal of Positive Behavior Interventions*, *172*, 69-82.

Stormont, M. A., Smith, S. C., & Lewis, T. J. (2007). Teacher implementation of precorrection and praise statements in Head Start classrooms as a component of a program-wide system of positive behavior support. *Journal of Behavioral Education*, *16*, 280-290.

What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education. (2017, October). *Standards handbook* (Version 4.0). Available from http://whatworks.ed.gov