A Single-Case Examination of Token Reinforcement and Self-Evaluation in a Self-Contained Classroom

Shiri Ayvazo[1], Avner Fraidlin[2], Ronit Kankazil - Maimon[3]

[1] School of Social Sciences and Humanities, Kinneret Academic College, Ztemach, Emek Hayarden, Israel
[2] Western Michigan University
[3] Applied Behavior Analysis Program, David Yellin College, Jerusalem, Israel

ABSTRACT:

Students with disruptive, impulse-control, and conduct disorders exhibit persistent problem behaviors that interfere with learning and instruction. Support interventions for these students should match their specific needs, such as learning to direct their behavior in the school setting. In this pilot study, we compared the efficacy of two universal teacher practices on a self-contained fifth-grade classroom student's Active Student Responding (ASR) and generalized on-task behavior. Token reinforcement, which was teacher-managed and commonly used in the classroom, and self-evaluation (SE), which was student-managed and novel for the teacher and the student. Interventions were evaluated during language arts lessons using a single-case Alternating Treatments Design. Findings demonstrated that the student's ASR comparatively improved under both interventions and on-task behavior mildly increased. The moderately positive impact of the two practices on ASR is discussed with emphasis on the appropriateness of SE for participants with emotional-behavioral disorders.

INTRODUCTION

Students lacking in emotional or behavioral regulation are often described in the research literature as students with or at risk for Emotional Behavioral Disorder (EBD). According to the Diagnostic Statistical Manual of Mental Disorders (American Psychiatric Association, 2013), specific diagnoses of disorders associated with emotional and behavioral regulation include Oppositional Defiant Behavior (ODD), Intermittent Explosive Disorder, Conduct Disorder, Antisocial Personality Disorder, Pyromania, and Kleptomania. Attention Deficit Hyper-Activity disorder (ADHD) is a common co-occurring condition together with the above disruptive, impulse-control and conduct disorders (American Psychiatric Association, 2013). This paper uses the prevalent acronym EBD to refer to the cluster of the aforementioned disorders, particularly ODD and ADHD. ODD and ADHD are associated with deficits in emotional and/or behavioral regulation that occasion conflict with social norms and authority figures as well as behaviors that violate the rights of others (American Psychiatric Association, 2013). Students with EBD, who are diagnosed with ODD and ADHD, exhibit persistent problem behaviors such as disruptive behavior, non-compliance, distractibility, off-task behavior, and physical aggression, that interfere with learning and instruction (Hansen et al., 2014). They struggle with managing their academic and social behaviors (Mooney et al., 2005). These persistent behavior problems and low academic engagement require teachers to provide evidence-based support interventions for students with EBD learning in general education schools. The type of support should match the students' specific needs (Zaheer et al., 2019). One imperative need is for students to be able to manage their lives independently and adapt well to their environment. School settings require self-regulation and self-direction in following the school norms, appropriately interacting with adults and peers, and actively participating in the learning process (Bruhn et al., 2015).

Engaging students with EBD in the learning process requires increasing the number of opportunities students receive to actively respond to ongoing instruction. Active Student Response (ASR) occurs when the student emits a detectable response such as raising his or her hand, using a sign or writing an answer following a teacher-posed question (Ayvazo et al., 2020; Haydon et al., 2010; Lambert et al., 2006). The research on ASR with students with EBD has long shown that ASR is associated with improved students' engagement measures such as on-task behavior, and reduced levels of inappropriate behaviors such as inattention and disruptive behavior during instruction (Ayvazo et al., 2020; Haydon et al., 2010; Lambert et al., 2006; Rila et al., 2019).

An approach that aims to cultivate optimal conditions for learning is the Response to Intervention (RTI) framework (Greenberg & Abenavoli, 2017). The Positive Behavior Support (PBS) approach is an RTI-based model designed to improve behavior across the school via a three-tiered approach (Greenberg & Abenavoli, 2017; Mccurdy et al., 2016). The first tier entails universal practices that are applied to all students. The second tier focuses on targeted interventions for students who do not respond to the universal practice and need more specialized support. The third tier involves intensive function-based individualized interventions designed specifically for individual students who do not respond to the previous intervention supports.

Universal practices are interventions that support teachers in structuring the classroom, and in relating to students in a positive manner. These are also interventions that promote active response, self-control, positive peer relations, and reduce disruptive behaviors. A growing body of literature documents the positive impact of universal teacher practices on classroom conduct, student behavior, and the use of disciplinary and aversive procedures (Flannery et al., 2014; Mccurdy et al., 2016; Miller et al., 2005; Trussell et al., 2016). Universal practices are also relatively low in cost per person in the classroom compared to the other tier levels' supports (Greenberg & Abenavoli, 2017). A more intensive intervention such as a function-based individualized intervention, albeit effective in reducing problem behaviors, is more expensive, more demanding to implement in terms of teachers' time, and reduces academic instruction (Greenberg & Abenavoli, 2017; Mccurdy et al., 2016; Trussell et al., 2016).

This pilot study examined two universal practices, Token reinforcement (TR) and Self-Management (SM), that entail different contingency locus of control. TR, also known as token economy, is an entirely adult-mediated intervention (Briesch & Chafouleas, 2009). TR is a contingency reinforcement system in which students can earn immediate reinforcement (i.e., token) contingent on exhibiting target behaviors and later exchange it for a reward (Zaheer et al., 2019). The reinforcement contingencies in TR are contrived to occur more immediately and more frequently than how they naturally occur in the classroom and therefore may be more appropriate for students who exhibit challenging behaviors (DuPaul et al.,

2011). TR interventions are well-known in school settings and have long been used by teachers to reduce students' disruptive and off-task behavior, regulate behavior, increase academic engagement and encourage prosocial behavior (Maggin et al., 2011; Zaheer et al., 2019).

"Self-Management" (SM) is an umbrella of interventions where concontingency management placed with the student. SM interventions have been documented to improve students' on-task behaviors, increase academic engagement, decrease disruptive behaviors, and improve social skills (Briesch & Chafouleas, 2009; Bruhn et al., 2015; Busacca et al., 2015; Denune et al., 2015; Maggin et al., 2013; Mooney et al., 2005). Self-evaluation (SE) is a variant of SM, where students attend to their own behavior and evaluate it relative to an objective standard or an external evaluator (Ardoin & Martens, 2004; Hansen et al., 2014; Mooney et al., 2005). Students also receive reinforcement for the accuracy of their self-rating, in terms of how close their rating resembled the teacher's rating (Ardoin & Martens, 2004). It has been two decades since Mooney et al. (2005), in their literature review on SM applications for students with EBD, noted the lack of research in general education inclusive settings. Briesch et al. (2019) recently conducted another literature review examining the configurations of SM in school-based research. Their review indicates only a few studies in elementary school settings that have incorporated SE. There is still more to learn about SE applications in elementary school settings and with students with EBD.

An early example of SE can be found at Ardoin and Martens (2004), who studied how skilled elementary students (age 9-11) exhibiting behavioral and/or academic problems associated with ADHD can accurately evaluate their social behavior, disruptions, and out-of-seat behavior. Students were required to evaluate the amount of time they thought they were engaged in the target behavior (e.g., being out of their seat) using a four-point Likert scale. A symbol representing time (i.e., open and partially closed circles) and a written description (e.g., little amount of the time) appeared next to each point to aid understanding of the evaluation procedure. Students' disruptive behaviors decreased, particularly after they received training in SE accuracy. The researchers concluded that elementary students with problem behaviors can be trained to accurately evaluate their behavior (Ardoin & Martens, 2004).

Both SE and TR are considered universal teacher practices aimed to support students' behavior and minimize behavioral escalation. Mccurdy et al. (2016) pinpointed

the limited research of universal evidence-based practices in educational settings serving students with EBD. More research is needed in these settings.

Briesch & Chafouleas (2009) argued that although TR interventions are commonly used as a universal practice to support students with special needs, they might be limited in promoting students' self-control, independence and generalization, due to their inherent teacher-control locus. SE interventions could alleviate these limitations. As a student-managed intervention, SE could also be cost-effective by making more instructional time available for the teacher.

Considering the importance of providing behavioral support to students where their actual placement is, and the great need of students with EBD for behavioral support interventions, the purpose of this pilot investigation was to examine the efficacy of two universal teacher practices: TR and SE intervention. The main research question was, "What are the differential effects of TR and SE interventions on ASR of a student in an EBD self-contained classroom." The research also examined the collateral effects of the interventions on on-task responding. Finally, the study examined the student's level and accuracy of SE. Social validity data on the interventions' social acceptability were also collected and examined. The TR intervention has been commonly used in this classroom, while SE was novel to the teacher and the student. The study contributes to the existing literature by (a) evaluating the effectiveness of the universal practices (TR and SE) in a comparative single-case design, (b) expanding the research on SE interventions in inclusive elementary school settings and (c) documenting research on universal teacher practices in a middle-eastern country.

METHOD

This research was reviewed and approved by the chief science officer of the Israeli Ministry of Education. The participant, his parents, and the teacher gave their informed permission and assent to participate in the study. Methods of this case study were based on previous validated procedures (e.g., Ardoin & Martens, 2004; Moore et al., 2013).

THE PARTICIPANT

The participant in this study was a fifth-grade male student named Yaron (pseudonym). Yaron was 10.6 years old and this was his first year attending a self-contained classroom for students with EBD. Yaron was diagnosed

with ADHD, ODD and a learning disability. He had been treated with Methylphenidate for 2 years prior to the onset of the study. His medical treatment was gradually faded, as requested by his parents. He ceased medication entirely a month prior to the onset of the study and was not taking any medication during the study. The weening period was abundant with disruptive behaviors.

The Functional Analysis Screen Tool (FAST; Iwata et al., 2013) was used for the completion of Yaron's Functional Behavior Assessment (FBA). The FBA outcomes depicted Yaron was displaying severe disruptive behaviors for over 12 months, such as screaming at staff and students, interrupting class, loitering outside the classroom, physically assaulting students, and destroying school property. Yaron's homeroom teacher described him as short-tempered, impulsive and very talkative. He continuously imposed his thoughts and ideas on others, to the extent of disrupting the lesson. Most of his remarks were irrelevant to the ongoing learning in class and therefore pervasively interrupted the lesson. When asked to delay or restrain his irrelevant engagement (e.g., you can share this idea with us during recess time), Yaron typically responded by screaming and refusing to comply with instructions. Scatter plot data indicated that on average, Yaron was off-task and disruptive in 54% of the weekly lessons. These behaviors occurred daily in every lesson aside from physical education and science. The teacher assessed his problem behaviors as highly disruptive but with little risk to property or health. He was the most behaviorally concerning student in the classroom. The FAST scoring summary resulted in social attention as the primary potential source of reinforcement for Yaron's problem behaviors, and escape from tasks or activities as a secondary source.

Academically, Yaron was in the 70th percentile of his class. His Individualized Education Plan (IEP) consisted of language arts, math, and emotional-behavioral objectives. The educational team was challenged by his impulsive behavior and was ill-equipped to design appropriate objectives and activities to address the impulsivity. Yaron's language arts proficiency varied by topic. His reading comprehension and grammar skills were at the fourth-grade level and his oral expression was at the third-grade level. Writing skills were Yaron's main area of difficulty, as he performed at the second-grade level. Yaron's handwriting was sloppy and hard to decipher. Due to these ongoing difficulties, the homeroom special education teacher invited the research team, who were behavior analysts, to assist with an intervention aimed at improving Yaron's behavior regulation.

SETTING AND MATERIALS

The study was conducted at an elementary school in a central city in Israel serving approximately 600 students. Yaron was attending a fifth-grade classroom for students with EBD (*N*=10). Similar to the description in Steiner et al. (2013), classrooms for students with EBD in Israel are primary placements for students whose main diagnosis is ADHD, ODD, and conduct disorder. Students placed in an EBD classroom might have comorbid diagnoses such as anxiety disorder or a learning disability. These classrooms are characterized by a high frequency of externalizing problem behaviors and low rates of desired behaviors.

The study was conducted during language arts lessons, three times a week (i.e., two consecutive lessons each time, totaling six weekly lessons). Lessons included grammar, reading comprehension, and written and oral expression. Language arts sequencing, the pace of content, and level of difficulty were tailored for each student based on their performance levels as detailed in their IEP.

The homeroom teacher had been a special educator for 10 years and this was her second-year teaching that group of students. She had been implementing a TR system regularly for more than 4 years, targeting classroom rule-following behaviors (e.g., be ready with your learning materials on the desk). The teacher was unfamiliar with the SE intervention but was enthusiastic about the opportunity to learn a new behavioral skill and address students' low engagement in class. She requested that the two interventions be applied to all students in her class. The research team collected data only on Yaron's behavior.

For the SE intervention, the researchers used a wooden clip and a laminated 10X30 cm rectangle cardboard with a four-square column printed on it, which represented a medal scale of gold, silver, and bronze medals and an empty (i.e., white) square at the bottom. For the TR intervention, the researchers used a 5X8 cm laminated ticket called "Like" showing a thumbs-up picture, and a 5X8 cm "Golden Like" ticket. Discriminative stimuli in the form of visual symbols were alternately hung on the classroom whiteboard to signal which intervention was in effect. An enlarged medal scale (a 20X40 cm laminated cardboard with a green background) signaled SE, an enlarged golden "Like" ticket (40X40 cm laminated cardboard with orange background) signaled TE.

Response Definitions and Measurement
Dependent variables were ASR and on-task behavior. Active student response was defined based on previ-

ous research (i.e., Lambert et al., 2006) as a hand raise above head height within 10 sec following the teacher's question and before another student's response. Observational data were collected live using an event recording method during the 20 min of acquisition time. Data were calculated and presented as the mean percentage of ASR out of total OTR.

On-task behavior was also defined, based on previous research (Moore et al., 2013), as any observable response made by the student that complied with instructions given by the teacher. Observers collected on-task data during a 20-min independent practice segment of the lesson, which was not supplemented with the intervention. Data were collected using a 10-sec whole-interval recording method. The teacher validated the definitions of ASR and on-task behavior.

The study also measured data on permission to respond and the accuracy of SE. Permission to respond was defined as responding to the teacher's question after being permitted either by a head nod or by verbal permission (e.g., "Yes, Yaron"). Permission data were collected using an event recording method. Data were calculated and presented as the mean percentage of permissions to respond. Accuracy of the evaluation was defined as the match between the participant's SE of ASR and the actual raw data collected by the observers on ASR. Accuracy was recorded using the permanent product of Yaron's evaluation and the ASR percentage data collected by the observers.

All aforementioned dependent variables were collected live in the classroom by three graduate-level behavior analysts with experience in observation and data collection methods. All observers were trained to collect live data on the dependent variables across all research conditions. The live observation was advantageous as the observers positioned themselves in vantage locations that allowed the detection of the student's and the teacher's verbal and non-verbal responses. Their presence in the classroom allowed great familiarity with the lessons' contextual variables that could be undetected in video recordings. Participants' reactivity to the observers in the classroom is a known drawback of live coding. One observer attended the class a few times before the beginning of the study to reduce possible students' reactivity and to collect preliminary data on participation levels and types of questions asked during language arts lessons. A second observer has been working as a behavior analyst in this classroom and the participant was acquainted with her. Finally, the observers stood in a permanent location at the back of the classroom away from students' sight.

The study also examined social validity like Lambert et al. (2006) social validity assessment. A short interview was conducted with Yaron and a questionnaire was given to the teacher at the end of the study. The interview included seven questions regarding the execution of the interventions, level of difficulty, and level of satisfaction, for example: "Which of the interventions did you prefer most?" The teacher's questionnaire comprised 16 questions about the implementation of the interventions and their feasibility and suitability to the teacher, to the class, and to Yaron. For instance, one question was, "Which of the two interventions was more suitable for Yaron and why?"

Experimental Design and Procedures
This study entailed a single-case design (SCD) approach, that is commonly used to define and examine educational practices at the level of the individual learner (Busacca et al., 2015; Horner et al., 2005). SCD use individuals as their own control and collect repeated measures of one or more dependent variables (DV) in the presence and absence of a practice manipulated by the researcher (i.e., the independent variable, IV) over time (Horner et al., 2005). The important elements to consider when evaluating SCD are methodological features associated with internal validity such as adequate interobserver agreement (IOA), treatment integrity data and repeated measurement over time; and experimental control evaluated by visually inspecting the graphed data to assess functional relations between the IV and the DV (Busacca et al., 2015).

One of the appropriate and acceptable SCD is an Alternating Treatments Design (ATD). In ATD the researcher conducts rapid and frequent alternations of conditions in a manner that allows for comparisons between them. Manolov et al. (2022) argue that ATD is particularly useful in applied research due to its ability to compare different interventions' efficacy in applied settings, as opposed to the comparison that baseline interventions offer. The rapid and frequent alternation and comparison of treatments also allow for several demonstrations of the interventions' impact in a brief period. The current study used ATD with an initial baseline phase (Barlow & Hayes, 1979) to compare the effects of TR and SE on Yaron's classroom performance. The baseline condition consisted of 10 data points, and the two interventions (i.e., TR and SE) were delivered randomly and interchangeably across 24 days.

Baseline Phase
Typical instruction during the baseline phase did not involve any special contrived contingencies for impacting ASR beyond occasional verbal praise and prompts to

participate (e.g., "I want to see more participants"). Each lesson began with a brief review of previous objectives (e.g., prefixes, suffixes, and roots) followed by a whole-class 20-min segment for the acquisition of new content. Acquisition entailed questions and students were expected to respond with a hand raise. The interventions were applied throughout this segment. Next, students continued with personal-level independent practice until the lesson's end. The teacher guided as needed during this segment. The interventions were non-active during the independent practice time. These three segments (i.e., review, acquisition, and practice) were maintained throughout the study.

TR and SE Training

Two training sessions (one for each intervention) were delivered for all students in the class. Training for TR included the practice of appropriate ASR using examples and nonexamples. The teacher introduced the token that would be awarded for appropriate and frequent ASR. Students then practiced ASR during language arts lessons. At the conclusion of the practice, the teacher praised students who demonstrated high levels of ASR and awarded them with the "like" token. She also encouraged those who did not meet the criteria to try harder next time.

Training for SE began with the teacher's explanation of SE, the medal scale, and contingencies pertaining to each of the medals. She explained that gold to bronze medal levels varied from responding to a majority of the questions to responding to only a few questions (accordingly) and that no-medal (i.e., the empty square on the scale) represented no participation at all. Each student received their own medal scale and a wooden clip and practiced placing the different squares. Next, students practiced attending to their participation level and evaluating it during short mock teaching segments. The teacher provided feedback to students on their SE and participation. These short simulations were repeated several times during the training session to promote students' acquisition of SE skills related to ASR. On day 22 all students in the class received a booster practice to enhance their SE accuracy.

Token Reinforcement Intervention

Each lesson began with the teacher reminding students of her expectation for ASR and the token reinforcer (i.e., the "like" ticket) for appropriate active participation. She placed the TR corresponding visual discriminative stimulus on the board to signal the intervention in effect. At the end of the acquisition segment, the teacher praised students who actively responded to more than 50% of

the questions posed and gave them a token that could be exchanged for a social reinforcer (e.g., a picture with the teacher) or a tangible item (e.g., erasers). The teacher briefly consulted the researchers, who observed and collected data, to determine students' participation levels. After 16 intervention sessions, and following a class-wide preference assessment, a backup reinforcer was introduced, and students could redeem their earned and accumulated tokens toward the classroom token economy system that had been used regularly during other classes (excluding language arts).

Self-Evaluation

Each lesson began with the teacher reminding students of her expectation for ASR and the SE (i.e., medal scale) to be used to evaluate active participation. She placed the SE corresponding visual discriminative stimulus on the board to signal the intervention in effect and reminded the students of the four levels of participation performance. At the end of the acquisition segment, the teacher asked the students to evaluate their participation using their medal scale. After visually scanning students' reports, she praised those who participated at the gold level, provided corrective feedback to students whom she thought evaluated their performance inaccurately, and praised a selected student for an honest evaluation. Finally, the teacher encouraged students to strive for higher-level participation in the next lesson.

Interobserver Agreement (IOA) and Treatment Integrity

The percentage of IOA for ASR, permission to respond, and on-task behavior were calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. IOA for ASR and permission to respond was obtained for 35% of the sessions; IOA for ASR averaged 93% (range: 82%–100%) and IOA for permission to respond averaged 97% (range: 87%–100%). Interobserver agreement for on-task behavior was obtained for 41% of the sessions and averaged at 89% (range: 60%–100%).

Two observers assessed treatment fidelity at the conclusion of every lesson via systematic direct observation of the occurrence or absence of the main components of the interventions: visual discriminative stimulus on the board, reminder of the expected behaviors, token and verbal praise given for responding to most questions during acquisition, abstention from exceedingly powerful verbal praise or prompts for participation and SE using the medal scale and contingent praise. Both interventions were

executed with 100% accuracy on all items, excluding abstention which was achieved with 92.5% accuracy.

RESULTS

Figure 1 depicts the average percentage of Yaron's ASR and of permission given to respond across all research conditions. Yaron's ASR during the baseline phase was 32% (range: 9%–53%), and he increased his ASR to 72% (range: 43%–96%) under the TR intervention and 74% (range: 43%–95%) under the SE intervention. Yaron received permission to respond on 23% (range: 0%–50%) of ASR occurrences during the baseline phase. He was given permission to respond on 21% (range: 4%–52%) of the ASR occurrences during TR and on 16% (range: 5%–33%) of occurrences during SE.
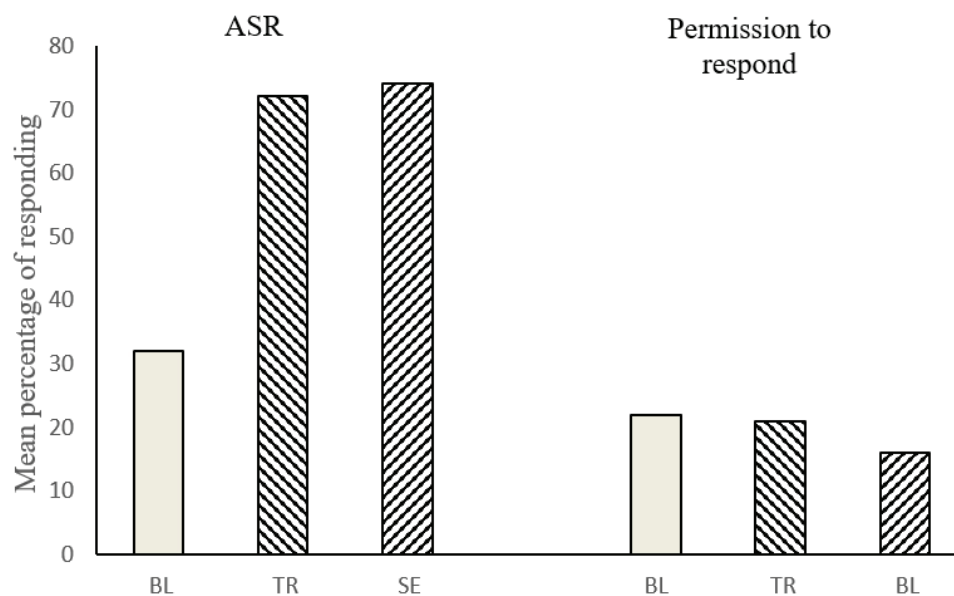
Fig. 1. Average percentage of Yaron's Active Student Responding (ASR, left half) and of permission given by the teacher to respond (right half) across Baseline (BL), Token Reinforcement (TR) and Self-evaluation (SE).
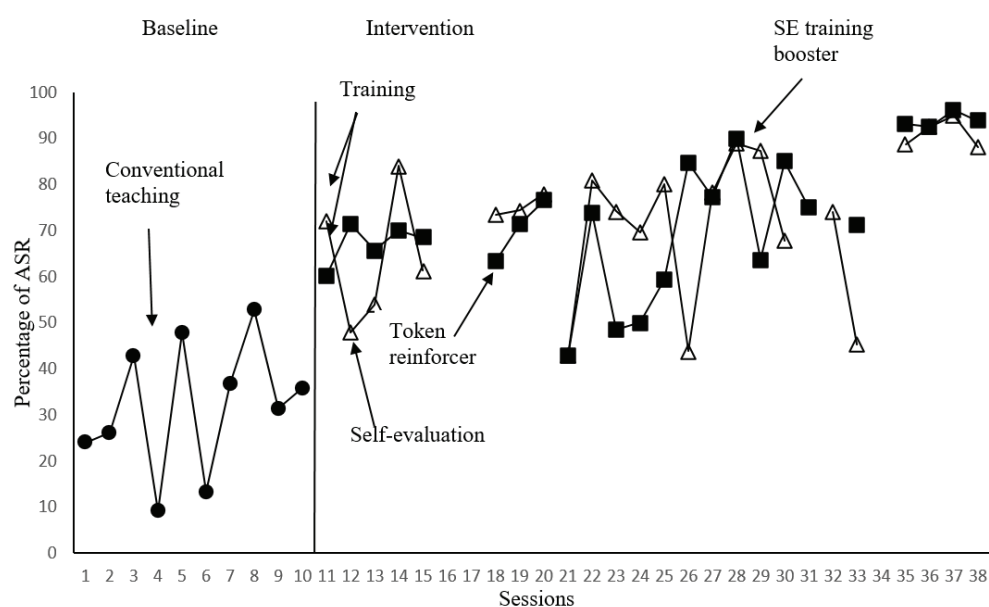
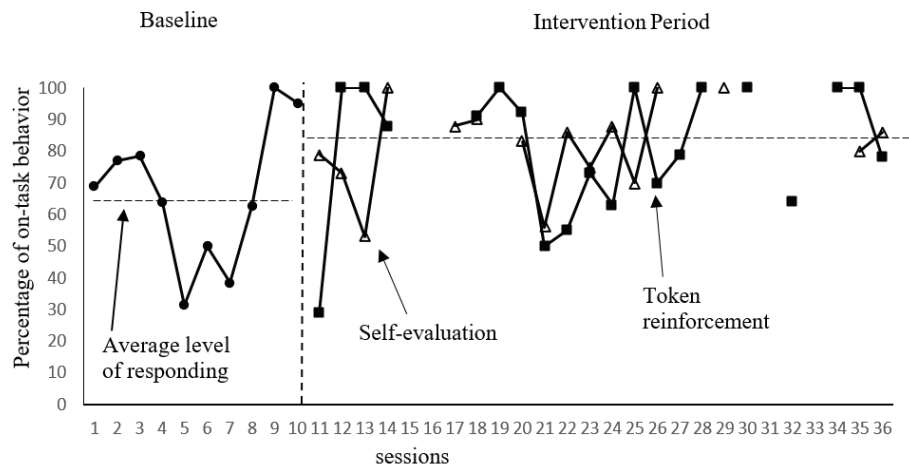Fig. 2. Yaron's percentage of ASR throughout the study.

Fig. 3. Yaron's on-task behavior during the practice segment of each lesson. Dotted line represents average level of on-task behavior.

Table 1. **Yaron's self-evaluation and its accuracy**

Day	Participant SE	Accuracy		Day	Participant SE	Accuracy		Day	Participant SE	Accuracy
11	G	A		21	G	I		31	-	-
12	G	A		22	G	A		32	G	A
13	G	I		23	G	A		33	G	A
14	G	A		24	G	A		34	-	-
15	S	A		25	G	A		35	G	A
16	_	-		26	G	I		36	G	A
17	-	-		27	G	A		37	G	A
18	G	I		28	G	A		38	G	A
19	G	A		29	G	A				
20	G	A		30	G	A				
Total correct SE		83%=20/24								

Note. SE=Self-evaluation. I=Inaccurate; A=Accurate; G=Gold level; S=Silver level.

Figure 2 depicts Yaron's ASR point-by-point performance throughout the study. His ASR performance during the baseline condition averaged 32% with variability and a slightly increasing trend. His ASR improved once interventions were introduced to 72% under the TR intervention and 74% under the SE intervention. An increasing trend was inspected under the two interventions with variability. Yaron's ASR performance stabilized above 90% in the last four sessions of the study, for both interventions. The percentage of non-overlapping data (PND) for each intervention was calculated by dividing the number of intervention data points that exceeded the highest baseline data point by the total number of intervention data points. The PND for TR was 87% and for SE it was 83%, yielding a moderately effective effect size (Kratochwill et al., 2010).

Figure 3 depicts Yaron's on-task behavior during the practice segment of each lesson. Data path appears as a dotted line since on-task measures were collected subsequent to the intervention. Average on-task behavior during the baseline phase was 67% (range: 31%–100%) with high variability and a decreasing trend followed by a decreasing trend. During the TR intervention, Yaron's on-task behavior increased to an average of 81% (range: 29%–100%) and of 81% during SE intervention (range: 53%–100%). On-task data during the interventions showed a variable trend, with frequent overlap with baseline data.

We measured the accuracy of Yaron's SE by tracking his SE selection and then dividing the number of accurate evaluations (compared to the observer's evaluations) by the total number of evaluations made. Data are presented in Table 1. Yaron evaluated his participation at the gold level 23 times (i.e., 92%) and once at the silver level. Twenty of the evaluations were accurate, reaching 83% accuracy.

Social validity was assessed at the end of the study. Yaron reported in the interview that he liked both interventions, but preferred SE over TR since it was easier for him to SE his behavior and achieve an award using the medal scale than earning a token from the teacher under TR. Yaron thought he participated equally under each intervention. He also wanted to continue using both strategies in lessons other than language arts. The teacher reported she found both interventions useful, feasible, and suitable for her students. She thought TR was more suitable for Yaron because, in her opinion, he needed external reinforcement from an adult. She also thought his engagement in class improved. He seemed more regulated and less impulsive and made efforts to attend class discussions. She also noticed his on-task behavior during independent practice markedly improved upon success in the preceding acquisition segment. Lastly, his behavior conduct under intervention was better than in other lessons she taught without those interventions.

DISCUSSION

This pilot investigation compared the differential effects of two universal practices, TR and SE, on EBD self-contained classroom students' ASR. The study also examined the collateral effects of the interventions on on-task responding. Finally, the study examined the student's level and accuracy of SE and examined the social validity of the interventions. The TR intervention has been in use by the classroom teacher, while SE was a novel intervention for the teacher and the student. The two evaluated interventions resulted in a moderate positive impact on the student's ASR. The improvement in ASR performance was comparable in both interventions, although SE, was novel for the teacher and student while TR was conventionally used. These findings possibly demonstrate the potential value of SE on the behavior management of students with EBD. The current data show the participant, who was severely lacking self-control and experiencing disturbing behavior and conduct problems, evaluated his participation at the gold level 92% of the opportunities. He was accurate 83% of the

time, based on the actual data collected by the observers. One hypothesis could be that Yaron rated himself "gold" all of the time, suggesting he was responding to the gold stimulus in the evaluations, rather than to his actual performance. However, since Yaron's SE data co-occurred simultaneously with an increase in his actual appropriate engagement in the classroom, we cautiously hypothesize that the participant's repeated assessment of his engagement at the gold level indicates increased effort to self-manage and meet the teacher's expectations regarding the target behavior. Previous findings by Ardoin and Martens (2004) and DuPaul et al. (2011) offer similar observations. Elementary students with problem behaviors can successfully evaluate their behavior.

This pilot investigation contributes to our clinical insights. Both TR and SE made an equally-comparable impact at the student's individual level. We carefully suggest that if a student with EBD can be taught to self-evaluate, then SE can be more valuable over a teacher-managed intervention in several ways. First, by promoting the acquisition of self-management among students who may need it most. Second, by extending the adaptive skillset and independence of students who demonstrate the greatest deficits in these areas. Third, by teaching a strategy that could assist students to generalize conduct and achievements across different conditions. Fourth, by reducing the behavior-management load from the teacher and possibly making more time available for instruction (Ardoin & Martens, 2004; Briesch & Chafouleas, 2009). The social validity data also support these insights. The teacher acknowledged that SE allowed for more time to be dedicated to instruction and feedback as the student took responsibility for regulating his behavior. The participant also reported preferring SE over TR.

On-task behavior showed only mild improvement. We suspect improvement in on-task behavior is better explained as a collateral positive change. Increases in on-task behavior may have been a function of increased ASR. The participant may have been more likely to engage in behaviors related to the assignment (i.e., staying on task) during the independent practice due to the preceding activity where he was responding to discriminative stimuli (i.e., OTR) and his response was differentially reinforced (e.g., praise or tokens for correct responses). The on-task finding does not shed light on a generalization of classroom engagement to un-intervened times and should be further examined under more rigorous research conditions.

This pilot investigation might also contribute to ecological validity by examining the validity of two universal

practices in a real-world self-contained classroom setting that accurately reflect students' placement as suggested by Mooney et al. (2005). The behavior analysts providing the service in this setting pursued a research plan and empirically monitored the classroom interventions within a case-study capacity. These conditions also explain the limitations of this pilot study. The first limitation is the measurement of performance of a single participant in a SCD, which critically limits the generalization of findings. Future applications should specifically examine the effects of SE on a larger sample of participants with EBD in their natural educational settings. Second, the variable and comparable nature of the findings of the alternating treatments posed a threat to internal validity. A return to baseline consequently of the alternating treatments phase and reintroduction of each condition discretely could have substantiated the superiority of one intervention over the other. Third, this study is limited by the fact that it did not measure academic outcomes associated with increased ASR.

All in all, the preliminary results of this study contribute to translating research into practice in two ways. First, this investigation adds information on SE as an IV in an elongated ATD. Second, this study generates empirical findings from two universal teacher practices executed in a real-world setting that accurately reflect students' placement (i.e., a self-contained classroom within a public-school elementary setting) as encouraged by Mooney et al. (2005) and Briesch et al. (2019).

IMPLICATIONS FOR PRACTICE

Three implications for special educators and professionals (e.g., behavior analysts) who support students with EBD emerge from this study. First, teachers of students with EBD should carefully consider the suitability of interventions to students' needs. Self-management and SE techniques could be as good as TR in terms of lesson conduct, but better than TR in terms of the skillset it allows students to acquire. Second, although we collected data on a single participant, SE was implemented with the entire class classroom and served the teacher well as a universal intervention. SE could also be implemented at the individual level to support those students with EBD, who are placed in general education classrooms, or those who are based in a self-contained classroom and participate in a few weekly classes in the inclusive classroom. Third, when using SE, we recommend utilizing a criteria system of gold, silver, and bronze medals. Students in the current study were excited and motivated to use the medals' system, possibility due to their general acquaintance with the reinforcing qualities of medals in sports, and in other areas of performance.

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Ardoin, S. P. & Martens, B. K. (2004). Training children to make accurate self-evaluation: Effects on behavior and the quality of self-rating. *Journal of Behavioral Education, 13,* 1-23. https://doi.org/10.1023/B:JOBE.0000011257.63085.88

Ayvazo, S., Fraidlin, A., & Kankazil-Maimon, R. (2020). Promoting active responding of students with Emotional-Behavioral Disorders through behavioral strategies. *Issues in Special Education and Inclusion, 30,* 7-24. [Hebrew].

Barlow, D. H., & Hayes, S. C. (1979). Alternating treatments design: One strategy for comparing the effects of two treatments in a single subject. *Journal of Applied Behavior Analysis*, *12*(2), 199-210. https://doi.org/10.1901/jaba.1979.12-199

Briesch, A. M., & Chafouleas, S. M. (2009). Review and analysis of literature on self-management interventions to promote appropriate classroom behaviors (1988-2008). *School Psychology Quarterly, 24*(2)*,* 106-118. https://doi.org/10.1037/a0016159

Briesch, A. M., Daniels, B., & Beneville, M. (2019). Unpacking the term "Self-Management": Understanding intervention applications with the school-based literature. *Journal of Behavioral Education, 28,* 54-77. https://doi.org/10.1007/s10864-018-9303-1

Bruhn, A., McDaniel, S., & Kreigh, C. (2015). Self-monitoring interventions for students with behavior problems: A systematic of current research. *Behavioral Disorders, 40*(2), 102-121. https://doi.org/10.17988/BD-13-45.1

Busacca, M. L., Anderson, A., & Moore, D. W. (2015). Self-management for primary school students demonstrating problem behavior in regular classrooms: Evidence review of single-case design research. *Journal of Behavioral Education, 24*(2)*,* 373-401. https://doi.org/10.1007/s10864-015-9230-3

Denune, H., Hawkins, R., Donovan, L., McCoy, D., Hall, L., & Moeder, A. (2015). Combining self-monitoring and interdependent group contingency to improve the behavior of sixth graders with EBD. *Psychology in the Schools, 52*(6), 562-577. https://doi.org/10.1002/pits.21846

DuPaul, G. J., Weyandt, L. L., & Janusis, G. M. (2011). ADHD in the classroom: Effective intervention strategies. *Theory into Practice, 50*(1), 35-42. https://doi.org/10.1080/00405841.2011.534935

Flannery, K., B., Fenning, P., McGrath Kato, M., & McIntosh, K. (2014). Effects of school-wide positive behavioral interventions and supports and fidelity of implementation on problem behavior in high schools. *School Psychology Quarterly, 29,* 111-124. https://doi.org 10.1037/spq0000039

Greenberg, M. T. & Abenavoli, R. (2017). Universal interventions: Fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness, 10*(1), 40-67. https://doi.org/10.1080/19345747.2016.1246632

Hansen, B. D., Wills, H. P., Kamps, D. M., & Greenwood, C. R. (2014). The effects of function-based self-management interventions on student behavior. *Journal of Emotional and Behavioral Disorders*, *22*(3), 149-159. https://doi.org/10.1177/1063426613476345

Haydon, T., Conroy, M. A., Scott, T. M., Sindelar, P. T., Barber, B. R., Orlando, A. (2010). A comparison of three types of opportunities to respond on student academic and social behaviors. *Journal of Emotional and Behavioral Disorders, 18,* 27-40. https://doi.org/10.1177/1063426609333448

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*(2), 165-179. https://doi.org/10.1177/001440290507100203

Iwata, B. A., DeLeon, I. G., & Roscoe, E. M. (2013). Reliability and validity of the functional analysis screening tool. *Journal of Applied Behavior Analysis, 46,* 271-284. https://doi.org/10.1002/jaba.31

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation.* What Works Clearinghouse. http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Lambert, M. C., Cartledge, G., Heward, W. L., & Lo, Ya-Yu. (2006). Effects of response cards on disruptive behavior and academic responding during math lessons by fourth-grade urban students. *Journal of Positive Behavior Interventions, 8*(2), 88-99. https://doi.org/10.1177/10983007060080020701

Maggin, D. M., Briesch, A. M., & Chafouleas, M. (2013). An application of the What Works Clearinghouse standards for evaluating single-subject research: Synthesis of the self-management literature base. *Remedial and Special Education, 34*(1), 44-58. https://doi.org/10.1177/0741932511435176

Maggin, D. M., Chafouleas, S. M., Goddard, K. M., & Johnson, A. H. (2011). A systematic evaluation of token economies as a classroom management tool for students with challenging behavior. *Journal of School Psychology, 49*(5)*,* 529-554. https://doi.org/10.1016/j.jsp.2011.05.001

Manolov, R., Tanious, R., & Onghena, P. (2022). Quantitative techniques and graphical representations for interpreting results from alternative treatment design. *Perspectives on Behavior Science, 45,* 259-294. https://doi.org/10.1007/s40614-021-00289-9

Mccurdy, B. L., Thomas, L., Truckenmiller, A., Rich, S. H., Hillis-Clark, P., & Lopez, J. C. (2016). School-wide positive behvioral interventions and supports for students with emotional and behavioral disorders. *Psychology in the Schools*, *53*(4), 375-389. https://onlinelibrary.wiley.com/doi/10.1002/pits.21913

Miller, D. N., George, M. P., & Fogt, J. B. (2005). Establishing and sustaining research-based practices at centennial school: A descriptive case study of systematic change. *Psychology in the Schools, 42,* 553-567. https://doi.org/10.1002/pits.20091

Mooney, P., Ryan, J. B., Uhing, B. M., Reid, R., & Esptein, M. H. (2005). A review of self-management interventions targeting academic outcomes for students with emotional and behavioral disorders. *Journal of Behavioral Education, 14*(3), 203-221. https://doi.org/10.1007/s10864-005-6298-1

Moore, D., Anderson, A., Glassenbury, M., Lang, R., & Didden, R. (2013). Increasing on-task behavior in students in a regular classroom: Effectiveness of a self-management procedure using a tactile prompt. *Journal of Behavioral Education, 22*, 302-311. https://doi.org/10.1007/s10864-013-9180-6

Rila, A., Estrapala, S., & Bruhn, A. L. (2019). Using technology to increase opportunities to respond. *Beyond Behavior, 28*(1), 36–45. https://doi.org/10.1177/10742956198352

Steiner, N. J., Sidhu, T. K., Pop, P. G., & Frenette, E. C. (2013). Yoga in an urban school for children with emotional and behavioral disorders: A feasibility study. *Journal of Child and Family Studies, 22*(6)*,* 815-826. https://doi.org/10.1007/s10826-012-9636-7

Trussell, R, P., Lewis, T. J., & Raynor, C. (2016). The impact of universal teacher practices and function-based behavior interventions on the rate of problem behaviors among at-risk students. *Education and Treatment of Children, 39*(3), 261-282. http://www.jstor.org/stable/44684109

Zaheer, I., Maggin, D., McDaniel, S., McIntosh, K., Rodriguez, B. J., & Fogt, J. B. (2019). Implementation of promising practices that support students with emotional and behavioral disorders. *Behavioral Disorders, 44*(2), 117-128. https://doi.org/10.1177/0198742918821331