



Theses and Dissertations

2022-04-12

Critical Incidents in Sustaining a Behavior Management Level System With Special Education Students in a Self-Contained School

Stephanie Johnson
Brigham Young University

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Counseling Commons](#)

BYU ScholarsArchive Citation

Johnson, Stephanie, "Critical Incidents in Sustaining a Behavior Management Level System With Special Education Students in a Self-Contained School" (2022). *Theses and Dissertations*. 9389.
<https://scholarsarchive.byu.edu/etd/9389>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.

Critical Incidents in Sustaining a Behavior Management Level System With
Special Education Students in a Self-Contained School

Stephanie Johnson

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Master of Science

Cade T. Charlton, Chair
Christian V. Sabey
Ellie L. Young

Department of Counseling Psychology and Special Education
Brigham Young University

Copyright © 2022 Stephanie Johnson

All Rights Reserved

ABSTRACT

Critical Incidents in Sustaining a Behavior Management Level System With Special Education Students in a Self-Contained School

Stephanie Johnson

Department of Counseling Psychology and Special Education, BYU

Master of Science

Level systems provide systematic support for teachers and students who need additional behavior support. The purpose of this study was to evaluate the perspectives of faculty members from a self-contained school, for special education students with severe behavior needs, which has sustained the use of a school-wide, behavior management level system for nearly 30 years. A total of 11 faculty members were interviewed using the critical incident technique, a qualitative methodology used as an exploratory tool in early stages of research. Interviews detailed the observable behaviors and specific events, critical incidents, that helped or hindered the sustainability of their school's level system. Faculty identified eight helping categories, six hindering categories, and six "wish list" categories, items participants wished had happened or could happen in the future. Helping categories included building rapport and teaming; collecting, recording and using data; communicating and collaborating; schoolwide consistency; relationships and student feedback; student engagement and buy-in; faculty buy-in and adaptability. Hindering categories included scoring and accountability inconsistencies; philosophical differences; insufficient staff training and experience; multifaceted system requiring nuanced decision making; staff mental health concerns; and community factors and external pressures. The findings of this study illustrate the importance of varied communication portals to support sustainability.

Keywords: level systems, behavior management systems, school wide behavior support, critical incident technique, sustainability

ACKNOWLEDGMENTS

I wish to express my sincerest gratitude for my research supervisor Dr. Cade Charlton. Without his motivational support, dedicated involvement, and patient mentoring over the past three years, this thesis would not have been accomplished. I also express thanks for my committee members Drs. Ellie Young and Christian Sabey. Their thoughtful feedback and inspired direction in my professional development was essential. I acknowledge my team members Jesse Rhodes and Maddy McGee for their support and efforts in this work. Lastly, I'd like to express my deepest appreciation for my husband and family. Their love and encouragement, sacrifice and patience, and laughter and love have kept me going. I would not be here without each of you.

TABLE OF CONTENTS

TITLE PAGE	i
ABSTRACT.....	ii
ACKNOWLEDGMENTS	iii
TABLE OF CONTENTS.....	iv
LIST OF TABLES.....	viii
DESCRIPTION OF THESIS STRUCTURE AND CONTENT	ix
Introduction.....	1
Statement of the Problem.....	6
George Washington School: An Exemplary Schoolwide Level System.....	6
Statement of the Purpose	7
Research Questions.....	7
Method.....	7
Participants.....	8
Setting	8
Measures	9
Research Design.....	10
Procedures.....	11
Data Analysis	12
Credibility Checks	13
Interview Fidelity.....	13
Extraction Check.....	13
Category Check.....	14

Participant Check	15
Expert Check.....	15
Participation Rate.....	16
Theoretical Validity	16
Results.....	16
Helping Categories.....	16
Building Rapport and Teaming.....	17
Relationships and Student Feedback	18
Collecting, Recording, and Using Data	19
Student Engagement and Buy-in	20
Communicating and Collaborating.....	21
Faculty Buy-In	23
Adaptability.....	24
Schoolwide Consistency.....	24
Miscellaneous	25
Hindering Categories	25
Philosophical Differences	26
Staff Scoring and Accountability Inconsistencies	27
Multifaceted System Requiring Nuanced Decision Making	28
Community Factors and External Pressures	29
Insufficient Staff Training and Experience.....	30
Staff Mental Health Concerns.....	31
Wish List Items.....	32

Evidence of Effectiveness.....	34
Discussion.....	34
Explanation of Key Findings	35
Limitations	38
Recommendations for Research	39
Recommendations for Practice	39
Conclusion	40
References.....	42
Tables.....	46
APPENDIX A: Review of the Literature.....	51
Teaching Students With EBD.....	54
Alternative Placements for Students With EBD.....	54
Evidence-Based Practice in Alternative Schools Placements.....	55
Systemic Supports in Alternative School Settings.....	57
Level Systems	58
Foundations of Level Systems	61
Level Systems in Schools	62
The Engineered Classroom.....	62
The Madison School Plan.....	63
Achievement Place.....	63
Boys Town	64
Level Systems in Special Education.....	65
Special Schools.....	65

Resource Room.....	67
Self-Contained EBD Students in Mainstream Classrooms.....	69
Implementation Challenges	70
Special Schools.....	71
Resource Classroom.....	72
Mainstream Classroom	72
Sustainability.....	73
Qualitative Research in Sustainability	74
Statement of the Problem.....	75
References.....	77
APPENDIX B: Institutional Review Board Approval Letter and Consent Form	84
APPENDIX C: Instruments	87

LIST OF TABLES

Table 1	<i>Participant Demographic Information</i>	46
Table 2	<i>Category Agreement Rate</i>	47
Table 3	<i>Incident Count and Percent Reporting for Helping Incident Categories</i>	48
Table 4	<i>Incident Count and Percent Reporting for Hindering Incident Categories</i>	49
Table 5	<i>Incident Count and Percent Reporting for Wish List Categories</i>	50

DESCRIPTION OF THESIS STRUCTURE AND CONTENT

This thesis, *Critical Incidents in Sustaining a Behavior Management Level System With Special Education Students in a Self-Contained School*, is written with journal publication formats and traditional thesis requirements. The preliminary pages of this thesis conform to the university requirements for submission. It is presented as a journal article and meets the length and style requirements for submission to *Behavioral Disorders*.

The literature review is included in Appendix A. Appendix B contains the Institutional Review Board consent form. Appendix C includes the interview template, the instrument used for data collection. This thesis format contains two reference lists. One reference list contains the references included in the journal-ready article and a second list includes all citations used in the literature review entitled “Review of the Literature,” which follows Appendix A.

Introduction

Children and adolescents with emotional and behavioral disturbances (EBD) are some of the most challenging students to support in school settings. Students with EBD demonstrate persistent behavior problems with both internalizing and externalizing behaviors that inhibit learning and can negatively impact the classroom environment (Klopfer et al., 2019; Oliver & Reschly, 2010; Reid et al., 2004; Wagner, 1995). Internalizing behaviors, while not as noticeable or disruptive to the classroom, can also negatively impact learning and in some cases immobilize students in social situations (McIntosh et al., 2014). Externalizing behaviors are far more disruptive to the classroom; they impede the instructional process and require deliberate attention because they can be harmful to both the student and others (Reddy et al., 2009). As a result, students with EBD often struggle to develop academic skills, earn the lowest grades, fail more classes, and have the highest drop-out rate of any other category of special education classifications (Oliver & Reschly, 2010; Reddy et al., 2009; Kostewicz et al., 2008; Reid et al., 2004; Sutherland & Singh, 2004). Low academic proficiency is compounded by high rates of absenteeism, disengagement from school, and long-term increases in failure rates in school for students with EBD (Wagner, 1995).

Students identified with EBD may or may not be identified under special education law. Within the Individuals with Disabilities Education Act (IDEA) special education students in need of behavior support could be identified under any of the 13 classifications, but often emotional disturbance (ED) is considered the behavior classification for students who need intensive behavior support. Within the literature, terms such as ED and EBD have been used interchangeably, even though ED typically refers to a classification under the IDEA and EBD refers to a number of behavioral disorders. Complicating this issue further is the term behavioral

disorder (BD) which is used to describe students who need additional behavior support, often in the form of a behavior plan, and is used to describe both special education and general education students. A more contemporary term used to describe students, in much the same way as BD, is social, emotional, or behavioral (SEB) needs. The term SEB also refers to students who may or may not have special education services. For the purposes of this thesis, EBD will be used as a broad term which describes students with intensive social, emotional and behavioral needs, who are classified under IDEA, but not necessarily as ED.

Challenging classroom behaviors may also lead to poor social relationships with peers and teachers. Externalizing behaviors are often irritating and can create negative feelings in teachers and peers, alienating students with EBD from healthy academic and social experiences (Reid et al., 2004). Veneziano and Veneziano (1988) found that a student's knowledge of social skills was highly related to his/her demonstration of pro-social behaviors. However, students with EBD struggle to interpret and respond appropriately to social cues (Lane et al., 2005). Students with limited social skills awareness (who struggle to recognize and understand social cues) will likely have more behavior problems, thus perpetuating a poor performance in both academics and interpersonal relationships.

Often, because of the complexities of educating students with EBD, these students are placed in alternative settings (Mattison, 2011). Alternative settings include self-contained classrooms and self-contained schools. These settings are more restrictive in nature but allow for a smaller class-size with more adult supervision. However, teachers in these settings struggle to create environments that are conducive to productive learning and improved behavior. According to Wehby et al. (2003) teachers in self-contained classrooms with EBD students spend only 30% of the day in academic instruction. To avoid aggressive behaviors, less academic demand is

given. During instruction wait time is high and structure and clarity are low (Kostewicz et al., 2008). Disruptive behaviors occur with more frequency in settings where classroom procedures are unclear and expectations are inconsistent, as well as in settings where teachers respond in untimely and ineffective ways (Eldar et al., 2018). Effective teachers promote student accountability by providing clear and explicit expectations, with steady positive feedback and consistently delivered consequences (Kostewicz et al., 2008; Swinson & Cording, 2002). Group contingencies can also reduce problem behavior and provide opportunities to teach appropriate classroom behaviors (Weeden et al., 2016). Purposeful praise and increased opportunities to respond improve the classroom environment by supporting an increase in appropriate behavior and a decrease of inappropriate behavior (Partin et al., 2009). A highly structured, organized, and positive environment is essential to the success of a self-contained classroom. Teachers need a comprehensive strategy that will help them to provide needed classroom structures and management procedures with clarity and consistency. Students need increased motivation to improve their circumstances. To meet the needs of teachers and students, a systematic approach is necessary.

Level systems can provide systemic support for teachers and EBD students in alternative placements. Fundamentally, level systems are an independent group contingency where performance criteria are given to the entire group, but reinforcement is delivered individually based on each student's unique performance (Cooper et al., 2007). Students move through different levels of criteria and earn privileges based on consistent demonstration of acceptable behaviors. When increased level expectations are met, more desirable privileges are available to students, motivating students to increase their appropriate behaviors. The ultimate goal of a level system is self-management through increased student responsibility and generalization of skills

in the least restrictive environment possible. However, level systems also help support teachers in creating a structured classroom with explicit rules and procedures. When used consistently, level systems provide order in the self-contained classroom, which promotes desired social, emotional, and academic behaviors (Cancio & Johnson, 2013).

Originally designed for use in institutional systems, Smith and Farrell (1993) suggested the contemporary development of level systems stems from the behavioral technology of the early 1960's in the field of ABA. Level systems rose in popularity during the 1980s in response to federal government demands to support students with EBD in the least restrictive environment. Level systems at that time were handed down as an effective technology for behavior management with little research to support the efficacy of their use in schools (Smith & Farrell, 1993). In 1998, to determine the extent of level system use, Farrell et al. surveyed 172 out of 200 randomly selected teachers of students with EBD, across grade levels and settings, in a southeastern state. At that time, 122 teachers of students with EBD were using level systems (71%). Teachers in special schools used level systems more than any other group within a service delivery model (93.3%) followed by teachers in self-contained classrooms (75.2%). The authors concluded, "Because so many teachers are using level systems, it is important to continue with thorough and systematic investigations of level systems so that their efficiency and effectiveness can be optimized for the benefit of all students" (p. 98). Yet research on level systems over the past few decades is limited, with much of the research published in grey literature (Farrell et al., 1998).

Level systems have several components to design and implement which can make them complicated to utilize and difficult to research. Teachers, students, and parents need to learn about the behavioral expectations and privileges associated with each level, as well as the

system for tracking student behaviors and the criteria for moving between levels. When combined with a token economy or a response cost system, additional criteria are implemented for earning or losing points and privileges. As a result, clear communication is critical when implementing a system; and doing so with fidelity adds another measure to consider. Consistency across time and people is essential. However, turnover in schools is constant, making sustained efforts in level system implementation and use challenging.

Sustainable practices help offset the challenges of a continually changing environment within schools. With student enrollment changing each year, as well as the constant turnover with personnel, schools need effective practices that can be taught to new hires and new students as an integral part of the environment. When schools stay with effective interventions over time, rather than trying something new every few years, they are better able to support training for school personnel, help students improve and use limited school resources more effectively. Student outcomes are improved not just through implemented effective practices, but also by sustained ones (McIntosh et al., 2009).

Qualitative research allows the more subjective aspects of sustainability to be identified through careful observation of interventions and firsthand communication with participants about their personal experiences. This method is appropriate when seeking to learn more about the less understood factors that contribute to or detract from effective performance of such experiences and activities (Butterfield et al., 2009). This methodology has never been used with level systems. Additionally, there is no empirical evidence to indicate which factors help in sustaining level system use.

Statement of the Problem

There is a current deficit in level systems literature and much of what is available was published in grey literature, a source only recently spurned. While a qualitative methodology has been used for other educational interventions, it has never been used as a way to measure level system sustainability. This study intends to use a qualitative approach to determine factors of sustainability for level systems that could be applied to other behavior management interventions and contribute more qualitative research to level system literature.

Understanding what helps or hinders sustainability with level systems will help teachers and school interventionists know what factors to consider when implementing level systems. This knowledge can be applied not just to special schools, but also to self-contained classrooms and other settings where students with EBD are taught. Furthermore, the findings from this study may contribute to sustainability research in general, providing guidance on supports and barriers to implementation of other school behavior interventions.

George Washington School: An Exemplary Schoolwide Level System

George Washington School (pseudonym), a self-contained school in the northwestern United States, provides intensive support for special education students with severe social, emotional, and behavioral needs. This school was the most restrictive behavior setting in the school district. Each student was referred into the school through a process of determining the least restrictive environment, a process which ensured each student had an Individual Educational Plan (IEP) and Behavior Intervention Plan (BIP). George Washington School was selected because of the sustained use of a schoolwide level system for more than 30 years across multiple administrative changes. faculty consistently voted to retain the level system.

Their school's level system was originally designed and implemented as a schoolwide behavior management system for student motivation and behavior support, with EBD students as the target. All aspects of behavior management were taught, tracked and supported through the utilization of their schoolwide level system. The level system was used throughout the school in every classroom, all school common areas, and even on the bus. Daily point cards were used to track individual and classroom target behaviors for each student and a number of inappropriate student behaviors, including those requiring in-school-suspension. Points were earned or lost based on current behavior and then attached to a token economy. Students could purchase reinforcements of their choice from the school store, including activities. However, some activities and options were accessible only at higher levels in the system.

Statement of the Purpose

The purpose of this study is to use qualitative measures to examine what factors support or hinder the sustained use of a schoolwide level system in a specific self-contained school that serves students with severe behavior problems, including those identified with EBD.

Research Questions

This study will address the following research questions:

1. What incidents help the sustainability of this level system?
2. What incidents hinder the sustainability of this level system?
3. What could improve sustainability of this level system in the future?
4. What is the perceived effectiveness of this level system for students?

Method

This study used a qualitative method called the Critical Incident Technique (CIT). Through structured interviews, with staff members from George Washington School, researchers

identified helping and hindering factors which impacted the sustained use of a school wide level system. All procedures within this study were approved by the university and district Institutional Review Board (IRB). Risks and benefits to participants were outlined in the consent form. All participants signed and returned the consent form.

Participants

Faculty members from George Washington School were selected as participants for this study and included instructional staff members who had involvement in utilizing the level system on a daily basis, for a minimum of one year. Utilizing the level system was defined as tracking behavior, awarding points and delivering classroom feedback to students. Instructional staff included certified teachers, paraeducators, and teacher specialists. Staff members who were not asked to participate included administration, front office staff and related servers (speech language pathologist, school psychologist). Five instructional staff members, one teacher and four support personnel, were not included because they had worked with the level system for less than one year. Eleven faculty members fit the inclusion criteria and 100% consented to participate. The group included eight females and three males; and consisted of six teachers, four support staff and a teacher specialist. Experience with the level system ranged from three to 25 years. The majority of participants received initial training on level systems through hands-on experience. Two participants had additional training outside of regular faculty reviews of the level system. See Table 1 for a complete list of participant background information.

Setting

Research for this study was conducted with faculty from George Washington School. During the 2020-2021 school year, George Washington School enrolled 10 students in grades 4-7, and employed six full-time teachers. The student population consisted of eight males and two

females, with 30% eligible for free and reduced lunch. The racial make-up of the school was eight white students, one African American student, and one Latino student per the National Center for Education Statistics (NCES). However, because of the nature of the school, the student population is highly variable within and across years. Students ages 8-22 years old have attended the school and enrollment has been as high as 45 students.

Measures

This study used a structured interview adapted from Butterfield et al. (2009). The interview consisted of four demographic questions, focused primarily on years of experience and training with level systems, and four project-focused questions on level systems. The purpose of the demographic questions was to build rapport with participants so they felt comfortable expressing their experiences in the interview (Butterfield et al., 2009). The purpose of the remaining questions was to allow participants to describe specific events and observable behaviors, particularly those that supported or inhibited the longevity of their school's level system. Four questions were asked:

1. What are the specific events that have helped you sustain the level system in your school?
2. What are the specific events that have hindered your efforts to sustain the level system in your school?
3. Looking back, are there other things that you wished had happened or could happen to improve the level system in your school?
4. In your experience, what have you noticed makes your school's level system effective for students?

Follow-up questions and probing were also used for clarification and additional details as participants answered. See Appendix C for complete interview guide.

Research Design

This study used the Critical Incident Technique (CIT) developed as a research method by Dr. Flanagan in 1954 (Butterfield et al., 2005). It is a qualitative research method, used to identify helping and hindering factors in specific experiences or activities. Grounded in industrial and organizational psychology, CIT was created as an exploratory method helpful in the early stages of research. Because the focus of CIT was on real-life experiences, researchers gained insight into real-world problems through reports of specific, observable behavior. This method was appropriate for seeking to learn more about the less understood factors that contributed to or detract from effective performance of such experiences and activities (Butterfield et al., 2009). Within the CIT methodology, observable and measurable events perceived as helping or hindering the activity or experience of focus, are called critical incidents (CIs).

More substantive research is needed on level systems since much of the literature is descriptive in nature, explaining the process for designing and implementing a level system. Smith and Farrell (1993) state, “Research about level system technology will demand varied and intricate methodology, complex analysis, and applicable results. Extensive ongoing research, both quantitative and qualitative, needs to be implemented to uncover the nature of level system use” (p. 262). CIT was an effective research method for this study because it contributes meaningfully to the current research on level systems. CIT utilized research data that is both qualitatively obtained, through interviews and observation, and quantitatively analyzed, through coding and credibility checks.

Procedures

The first author met with the second author, the research team expert on CIT, to review key components of fidelity within the CIT interview, and to practice using the script. Key ideas reviewed included: ensuring the participant agreed to the recording of the interview, clearly stating that specific events shared should be observable behaviors of professionals, and the importance of summarizing participant responses to allow for clarification or further detail. The techniques of probing for further detail and using follow-up questions were reviewed and practiced. Using general probes, for example, “Tell me more about that,” to avoid being leading was recommended. Lastly, the script was reviewed for verbiage that could be made more user friendly if needed. For example, the word sustainability could be explained as “a long time.”

Next, the first and second authors met with a member of the research team. The primary researcher trained the additional team member, as described above, and the expert supervised this training. The primary researcher and team member practiced the script together and received feedback from the expert throughout the process. Finally, the lead researcher met with a final team member and coached her on the interview process and script, as previously explained. During this final training meeting, a non-participant familiar with level systems was interviewed by the team member and received feedback throughout the practice session from the primary researcher.

All interviews were conducted and recorded through Zoom. They were completed over a three-month period by three trained members of the research team, including the research team expert, who had conducted or participated in several studies using the CIT methodology, as well as the primary researcher and a member of the research team, who were both new to this type of research. Interviews were conducted with an interview guide (see Appendix C) to ensure the

same questions were asked in the same order, that all information was covered, and to keep the interviewer focused on the participants' answers (Butterfield et al., 2009). At the onset of each interview the project was reviewed with participants and demographic information collected. Interviewers used personal interest, curiosity, and/or active listening, to facilitate the interview. As content questions were being answered, researchers noted the participant's responses on the interview guide. Before moving on to the next question, researchers summarized and repeated back participant responses, and allowed participants an opportunity to further clarify their response. Anything added, revised, or deleted from their response was notated on the interview guide. The use of follow up questions allowed the same amount of detail to be given by all participants as recommended in the literature (Butterfield et al., 2009).

Data Analysis

Transcriptions of all interviews were completed through Zoom. A member of the research team refined the transcripts by taking time stamps out, putting complete sentences together and checking against video footage. When workable copies of the transcripts were available, the primary researcher divided each transcript by question onto google slides, sectioning out first the helping incidents, followed by the hindering incidents, wish list (WL) items and effectiveness evidence for each participant's transcript. Each slide was notated with the participant number and a critical incident or wish list code (e.g., P3 Hinder 2), then the main point of each slide was highlighted and also used as a working title. Helping CI slides from nine transcripts were then copied into a new slide presentation where slides were moved into initial groups based on similarities and differences between the main ideas and working titles. This process was repeated for hindering CIs, WL items and effectiveness evidence. Next, categories were named and all slides were examined to see if they fit the groupings. Some groups that had

overlap were merged together and other groups, where different and distinct ideas began to form, were broken into smaller groups. An operational definition was written for each category and all previously categorized items were checked against the operational definitions. One more participant's CIs and WL items was then added into the slides and adjustments to groupings and definitions were made. Two new categories were created with hindering CIs. The last participant's CIs and WL items were added into the slides and no adjustments to groupings and definitions were made, indicating that saturation had been reached.

Credibility Checks

Credibility checks provided a systematic way to analyze the data. The following credibility checks were implemented to ensure the research was completed with fidelity and the analysis and interpretation of data were consistent.

Interview Fidelity

Interview fidelity was checked by the second author because of his previous experience and expertise with the CIT research method. To ensure the interview script was followed and all major questions within the protocol were appropriately addressed, a member of the research team reviewed one interview transcripts from each of the interviewers, approximately 30% (Butterfield et al., 2009). Specifically, the reviewer verified that all essential questions were asked and that critical incidents were reviewed with the participant prior to moving on to the next section. The reviewer noted 100% agreement between the transcripts and the fidelity check indicating no further review was needed.

Extraction Check

The member of the research team who reviewed the Zoom transcriptions, also helped with the extraction check by creating an additional list of helping and hindering CIs, and WL

items. Before extraction, the research team met and discussed the key features of extraction, namely looking for the CIs in participants' answers, which could be supported by examples or descriptions of impact (Butterfield et al., 2009). As the researcher reviewed transcripts, critical incidents were bolded. A chart for each transcript was created showing the number of each type of CI and WL items for each participant. Transcripts were randomly selected from each interviewer pool, 36% in total, and a rate of agreement was calculated between the primary researcher and team member to determine the rate of agreement. The number of agreements on CIs and WL items (24) was divided by the total number of CIs and WL items (25), then multiplied by 100 to get a percentage of 96% agreement. The cited incidents were deemed critical to level system sustainability because of the high agreement rate (Butterfield et al., 2005).

Category Check

A copy of each of the original slide presentations (helping, hindering, and wish list) was sent to a member of the research team. All title and definition slides were moved to the top of each presentation and the remaining CI slides were scrambled. The primary researcher explained the categorization process as identifying the CI and then using titles and definitions to determine correct placement. For CIs without an obvious placement decision, operational definitions were used for clarification. A sampling of slides was reviewed for formatting. The team member then independently categorized the critical incidents. An agreement was counted for each CI that was placed by the independent researcher into the same category as in the original coding. Butterfield et al. (2009) recommend following the 80% or higher match rate guideline for this credibility check, which was originally suggested by Andersson and Nilsson in 1964. The overall percent agreement for all categories was 75%. However, because of the limited number of CIs (81), which was below Flanagan's (1954) recommended minimum of 100, reaching a categorical

agreement rate of 80% was difficult for categories with only a few CIs. With only one disagreement, the percent agreement could drop as low as 75% in categories with only four CIs. Primary areas of disagreement were between helping categories associated with teaming, communication and collaboration. Detailed information can be found in Table 2.

Participant Check

After the categories were created and extraction and categorization checks were completed, all participants were contacted a second time via email. The follow up email included the critical incidents identified in their individual report, and the associated category each incident was placed in. Participants were asked to review the information and indicate if the items looked correct, there was anything out-of-place, and if there was anything additional, they wanted to share. Six out of 11 participants responded to the email and all indicated that everything “looked good” to them.

Expert Check

After the categories were created and checked, both independently and by each participant, they were sent to experts for further analysis. Two experts were selected, based on their published research on level systems, to review the categories. They were asked the following questions.

1. Did you find the categories to be useful?
2. Are you surprised by any of the categories?
3. Do you think there is anything missing, based on your experience?

Their responses indicated that these categories and findings are consistent with their expectations.

Participation Rate

Each category was checked for a participation rate. Participation rate was calculated by counting the number of participants who referenced a specific event as a critical incident, dividing that number by the total number of participants, and multiplying by 100 to get a percentage. This calculation was repeated for each category. Consistent with other CIT studies, categories with a participation rate of 25% or higher were considered to be valid (Butterfield et al., 2005). Three hindering categories and five wish list categories did not meet the threshold of 25% participation rate. Tables 3, 4 and 5 contain a summary of the participation rate data for all categories.

Theoretical Validity

Theoretical agreement was obtained through comparing the underlying assumptions drawn from participant responses on level system sustainability to sustainability research on positive behavior supports. The helping and hindering categories created in this study were compared to the positive behavior support sustainability categories listed in the literature. McIntosh et al. (2013) identified ten different factors critical to sustainability. These include staff commitment, administrator support, integration into existing and new efforts, ongoing resources, perceived effectiveness, implementer skill and knowledge, teaming, efficiency, collection and use of data, and capacity building.

Results

Helping Categories

From the 11 interviews conducted by the research team, a total of 38 helping incidents were identified. All eight helping categories, the helping incidents within each category, and the percent of participants represented in each category are summarized in Table 2.

Building Rapport and Teaming

Participants identified building rapport and teaming as critical to sustaining the school-wide level system. They listed the following ways to build team rapport: spending time together talking, listening, and supporting each other after school; participating in team building activities; and getting to know each other throughout the school year. Participants indicated that teaming included weekly meetings in which teams discuss individual students, their behaviors and progress, through evaluating each student's data and determining their needs, including possible program modifications. They noted that teaming also included clarifying expectations and confusion around specific terms and procedures associated with the level system. While teaming does include data discussions, the collection, recording and use of data will be addressed in an additional category.

Building rapport and teaming included six critical incidents, as recorded by 45% of the participants. One participant stated that having a good rapport with fellow faculty members helped with feeling supported after a difficult day of work. "There's such hard days that we all just kind of get together and have to let it all go at the end of the day and think about other things." Several participants indicated that weekly team meetings were helpful in determining student needs. One replied, "we go through every single kid individually and talk about them and what has gone on for the week, and if changes need to be made." Additionally, participants said that team meetings were helpful in clarifying any confusion around the level system. "...we're talking about okay, has there been any drift? You know from center basically of the things that are like the core pieces of the system." The participant continued to explain helpful incidents as "those events where it's kind of a stop, recheck, make sure everybody's on the same page, [and] move forward."

Relationships and Student Feedback

Relationships between students and staff were noted as a critical incident in the sustained longevity of the level system by participants. They indicated that relationships were built and strengthened through student feedback provided by faculty members, which included holding students accountable for their actions as well as providing encouragement and reminders for improvement. Respondents noted that building rapport with students, through good communication with students and their families, allowed staff to get to know more about a student's life and created a closer bond of trust.

This category included six critical incidents, reported by 45% of the participants. One participant indicated that the foundation and strength of the level system was, in their opinion, the relationships between students and faculty. Another noted that the design of the program incorporated a teacher advisory (TA) period for the purpose of building a stronger rapport between the students and staff in each advisory. Since the teacher advisory for each student is responsible for communicating with each student's family, they are more familiar with their students and able to connect with them in a more personal way. Noted one participant,

“My understanding is it was designed, because you want like a stronger rapport with your TA. They're like your extra support people that you know who encourage. Good communication and parental involvement comes from the TA. That model works, I mean, I think it works good, and you know, because you just get a little bit closer bond, you're familiar with things going on in the family.”

Student feedback is provided both in the morning and afternoon when students check in and out of the day through their teacher advisory period. Respondents indicated that this time provided opportunities for more specific and direct encouragement and reminders for students

about their behaviors. For example, one participant stated, “we're able to say you know, like ‘you got some, you need to be careful with you know talking out. It looks like you got some inappropriate verbals. So what's going on?’” Another participant indicated this was a good time to set students up for success each day by having them set a goal each morning with their TA. Commented another participant,

“the relationships are part of this, that helps our system ...I mean it's not just an empty robot say[ing], ‘Okay, you did not do this,’ ...we talked to them about it. We gravitate toward the kid[s], they gravitate toward us, ...”

On the other hand, respondents also indicated that holding students accountable for their actions was just as important when giving feedback. A phrase one participant ascribed to was, “Holding students accountable when necessary and encouraging otherwise.” Having a program to follow allowed one participant to remain neutral when providing corrective feedback, which, in their opinion, preserved the teacher-student relationship.

“I tell that to the kids all the time. ‘This is your behavior. I have to mark it, that's my job.’ Um, so, it keeps it less personal with the kids so that they're not feeling as attacked when we mark these things.”

Collecting, Recording, and Using Data

Collecting, recording and use of data was identified by respondents as critical to sustaining the level system. Collecting data was noted as consistently marking interval data throughout each day on each student’s point card. Respondents noted that recording data was completed at the end of each school day, as teachers and paras together discussed and reflected on individual student data, checked it for accuracy, and entered it into a database in preparation for the next day. They indicated that using data consisted of reviewing data weekly, during

faculty meetings, to determine student progress and to make needed individual modifications; analyzing data for patterns when there are behavior issues; and utilizing data for student programming. Additionally, making data more visual, through point cards for students and by using graphs for faculty and parents, was listed as a helping factor in the effective use of data. With a total of five critical incidents, the data category included 36% of participants. Some respondents indicated that having data allows for better programming with individual students because specific behaviors can be tracked. For example, if a student continuously missed points for inappropriate language “then we know that that's a real issue when we look at all the data. Then we can focus on that, and put that up on their target and work on that specifically.” Several participants noted that having data was helpful in seeing the big picture around student behavior as well as convenient when trying to pinpoint specific issues with students. One participant, after reviewing a student’s data, recognized that the student was “not just having bad days, he's having bad fifth periods.” Discovering this pattern allowed the faculty “to take that data and figure out, okay now where can we go from here to figure out what's going on during the period.” Participants indicated that utilizing data in visual ways was not only helpful in making decisions about student needs but was also helpful in motivating students to improve their behavior. Daily student data, as noted by several participants, was collected on individual point cards during five-minute intervals. One respondent determined that the visual nature of this data would “help some of them [students] to guide themselves back to kind of regroup.”

Student Engagement and Buy-in

Student buy-in of the level system was determined to be crucial to the sustained longevity of the level system. Five critical incidents from 36% of respondents are represented in this

category. Respondents indicated that student buy-in was obtained through providing students with opportunities to get something they wanted, including rewards, privileges, and activities.

“To sustain it [the level system] the kids have to have buy-in. So it has to be something that they want to achieve...”

“...buy-in with the students. That is the biggest most important part, because if we don't get them to buy in there is no program. The rewards at the end are very significant. It's not just the store oriented stuff. So freedoms within the school, ... you can listen to music, while you work on the computer. ... you can use the staff stairs, ... you can wear a hat in school. So it's little things.”

“We have free time Fridays, where everybody comes [to] hang out in the gym. We play games ...and they get to play against the staff ... I don't know that many kids that don't want to come and defeat their teachers at something.”

Respondents noted that rewards were more desirable when students could see them on display or when faculty members reminded students of their available options. Additionally, they noted that continued student engagement in the level system was gained through offering novel rewards and privileges within the higher levels of the system. Random reinforcement from administration was also identified as a means for increasing student buy-in.

“Well, recently, the admin has come around for any student that doesn't have any ISSs [In-School Suspensions] on particular days. There's no rhyme or reason, when they come, they just come. They'll get a full candy bar. And the kids really like that.”

Communicating and Collaborating

Daily communication and collaboration were indicated as helping factors in sustaining a school-wide level system. Participants noted that communication among faculty members was

necessary in two important ways. In times of crisis, immediate communication informed other faculty members of a need for additional support, initiating an immediate collaborative response to a student's behavior. Throughout the day, communication among the faculty addressed general student concerns, for example, clarification on a missed point. Additionally, participants indicated daily discussions and collaboration were helpful in determining what was and was not effective for individual students, during times of crisis or success.

This category included four critical incidents and 36% of participants. Several participants noted that communication was critical in a time of crisis and being in constant communication with each other over walkie talkies improved the immediacy of a collaborative response and intervention. One participant stated, "even if it's not my student and they're having problems, whoever is available you step out and do what you need to do." Another related, "I think at this school it's good the way we're in immediate contact, it's not something you find out two days later...which is way after the fact. An intervention that late isn't going to be too useful. Ours are immediate."

Some respondents indicated that collaborating in the moment of decision-making was helpful.

"If one kid upstairs is having a behavior and they're not seeing that anyone up there is helping, like getting them turned around, then sometimes they call for one of us in the elementary to come up and try and redirect and help do that. Just a different face giving the same direction they were given originally. But then we're collaborating back and forth [on] what worked/what didn't work."

This same collaborative approach of “going back and forth” to make decisions was also mentioned as helpful when determining missed points in situations that weren't as straightforward.

Faculty Buy-In

With a total of four critical incidents, faculty buy-in was indicated by 27% of participants as a critical incident in sustaining the level system. Participants indicated that when staff members were able to recognize the effectiveness of the level system through improved student outcomes and changes in student behavior, this facilitated buy-in from the faculty. “A student that previously had walked out of class consistently. Because of our level system ... that behavior had been shaped and changed and ... walking out of class has reduced or eliminated completely.”

“you see weeks, weeks of you know them getting low percentages, ... and then the progression of that kid getting better and better and better and then the rise, the phase system, you ... see their excitement, and then you know when they get to that independence ... I've seen it just be successful with a majority of all our kids.”

Respondents noted that consistently talking about the program, how it worked and previous success stories, helped new staff to get onboard with the level system.

“At least every year there's a specific amount of time set aside where it's like we're sitting down and talking about, you know, where we're at with it, how it works, do we need to get anybody on board or you know some reeducation on it.”

It was also noted that faculty members who looked for and recognized the success of the functioning level system were more likely to buy-in to the system.

Adaptability

Having a system that was individual and personal for students was listed by 27% of participants as a critical incident in sustaining the level system. They noted that individualization occurred through the point card in two different ways. First, by including personal student goals and/or plans on the point card. “Everyone's on a point card...[and they are] each are on their own individual plan as well, so there are certain things on the point card that are individualized for each of the kids.” Second, by adjusting points and/or time intervals as needed so students can be more successful on the level system. “Specifically, if we have a student that is struggling staying in class, we adjust his missed points or we adjust his time in ISS to reflect what is best for him ... to be more successful in the classroom.” Keeping the point card personal, which allowed students to track their own performance, was also listed as a critical incident.

“...with the card, with the kid it seems a little more, it's it's not like put out on display in public, you know. It seems a little more private for the kid that they could sit and look and track at their own behaviors.”

Schoolwide Consistency

Respondents noted that consistency was critical to the sustainability of the level system. Specifically, they noted two areas where consistency was essential. First, consistency across classrooms with the structure of the level system, including using the same vocabulary and point card. Second, consistency in staff responses to student performance including delivering feedback and enforcing consequences. Nonexamples: This category does not include data collection, which was addressed in a previous category.

This category has three incidents including 27% of participants. Respondents indicated that having the same structure across the school helped students know what to expect. One

participant indicated that consistency helped new students settle into their new environment, stating,

“when they get here, sometimes it's rough for them. But we just keep the same structure of no, this is what we do here, and they get into it and then it's good for them to be in that structure.”

Another participant noted that using the same vocabulary supported schoolwide consistency stating,

“the vocabulary is the same, no matter what room they go into. The vocabulary is the same on the point cards. The time used within the point card is the same...so the students are very consistently using the same vocabulary. Yeah, so I think that is what works so well, is because no matter where they go it's consistent.”

Ensuring consistency across staff members when addressing student performance was also noted. One respondent indicated that “consistency across the board with everyone being on the same page of what is a behavior that we're that we're watching for and calling students on,” was critical when delivering feedback to students.

Miscellaneous

One additional critical incident was mentioned but, because the incident was about a specific situation with a specific student, it was disregarded. Single incidents may or may not be reflective of helping factors which impact the sustainability of the schoolwide level system.

Hindering Categories

A total of 27 hindering incidents were identified from the 11 interviews conducted by the research team. All hindering categories, the hindering incidents included within each category, and the percentage of participants within each category are summarized in Table 3.

Philosophical Differences

Participants indicated that differences in staff opinions, for example the “right” way to implement the program, as well as philosophical differences, for instance the purpose of the level system, were hindering factors in the sustainability of the level system. Participants noted that some of these differences appeared in the following ways: gender, tolerance or sensitivity levels, and relationship style with students. Eight (55%) hindering critical incidents were associated with this category. Disagreements among staff members about “how things should be done” was a hindering factor. One participant reported, “The only real hindrance I’ve ever found with it is just when somebody isn't doing it right.” Differences in opinion created discord among faculty and inconsistencies with students.

“...we [the staff] have our biases, we all strive to do what we think is right and not everybody agrees with everybody. I’m more patient and I won't, like, let things bother me. There's other staff that are like ‘no you can't do that,’.... There’s that kind of mentality. It's not, like, bad, it's just like some people are really on it, some are more relaxed.”

Others expressed that different perspectives on, and tolerance for, inappropriate student behavior hindered sustainability.

“...it's not personal, it's usually never about that behavior you're talking about or dealing with. It's, it's a bigger issue, sometimes, and I think a lot of teachers misread situations... You just cannot ever take anything personal because it's not about you it's about the kid.”

Differing philosophies among staff members about the purpose of the level system was also mentioned as a hindering factor. Some staff members used the level system as a way to show favoritism to certain students, while others used it as a form of punishment. “Sometimes

you get people in that, you know, might get a little too cozy, or tell too much, [to students] and that doesn't work well here.” Another participant stated, “I think the system is used to just get the kid away, instead [of] building a relationship to make some level of understanding between that teacher and that student.”

Staff Scoring and Accountability Inconsistencies

Participants noted that scoring inconsistencies were the result of human error and included both differences in how the data was recorded as well as inaccuracies in what data was recorded. Teacher perception of student intent resulted in varying degrees of student accountability among staff members. Six critical incidents from 45% of participants were associated with this category. Participants indicated that students received mixed messages through inconsistent scoring and accountability and that created confusion as well as opportunities for students to exploit the system. “If somebody is doing it wrong, it messes with the kids. [They] figure it out too, a little bit, and then they might exploit it.” They also noted that inconsistent collection of interval data and awarding of inflated bonuses created imbalance in the monetary system associated with the level system. “It definitely threw things off and ... it really hurt some of the other students.” One of the most frequent inconsistencies reported was not holding students accountable for their behavior in the same way.

“The biggest problem we have is people forgetting to, um, mark down some of these behaviors. ... because we're just warning, warning, warning and then not marking it down, so we have some kids who move up the level system even though their behavior really hasn't shown that they have earned that, ... so the data is not showing what we are seeing...”

Discretion in accountability was also the result of teacher perception of a student's intent, when demonstrating problem behavior. "...some students, they will push buttons and they do need that constant 'okay well that's a missed point.' Other students that are just messing around ... they're just too goofy." Differences in classroom management also created discrepancies in accountability. One teacher reported that in one classroom "it's like too by the book... you know cross your t's dot your i's, do everything. Whereas [in] the next room it's like they're just playing..."

Multifaceted System Requiring Nuanced Decision Making

Participants (36%) noted that making nuanced decisions within a complicated and multifaceted level system hindered the ability to sustain the schoolwide level system. They indicated that learning the structure of the level system was challenging because of the various components of the system and remembering how to implement them in a fair and consistent way, while providing individualization to particular students, added to the difficulty. Participants reported that it took experience and time to learn the nuances of the system and how to make appropriate decisions with different age groups and populations within the school. They noted that some elements of the level system were also difficult for students to remember, "some of our kids really have a hard time keeping track of things, but on the other hand, that's also something they need to work on 'cause that's life." While some participants indicated confusion in learning the components of the level system was hindering, others expressed that knowing how to apply the level system with individual students was hindering.

"It's not something you can even teach because a lot of it is discretionary. ... It just takes time, it's like a ballet dance almost, you can't just teach anybody that [and it] doesn't matter how much education you have, because we all have different ideas. And every

single one of these kids is different, some of them, you just have to learn like, well this child has you know this issue so we gotta do it a little bit different than this child and, so you have to constantly gauge who you're dealing with. But stay within the fairness too, because that would be a problem, too, if one of them's getting marked up for sitting on the floor and another one's not, that would also blow it, you know, so it's just a constant thing, and you just have to learn it, you know.”

When new faculty join the staff, this problem is compounded because the adults and students don't know each other. “Then you have a whole new piece of all these new people coming in, that the students don't know, the adults don't know and then it's that whole learning curve...”

Community Factors and External Pressures

Participants (18%) indicated that community factors and external pressures were a hindrance to the sustainability of the schoolwide level system. They noted that changes within the community impacted the student population at their school and also lessened support from outside agencies.

“...our population was, more of a lot of conduct disorder, we had a heavy gang population. ... so our system works differently ... it's the same system that we're running, but you run it a little differently than with the conduct disorder gang member, than you do with an autism and some emotional disorders.”

One participant emphasized that outside support, like truancy court, was once helpful in helping students get what they need beyond the level system. “...we have these consequences and things so streamlined, but when it gets to that point to where we need help outside of this place

it's hard getting it.” External pressures, such as familial influences, were also mentioned as hindering factors in the sustainability of the level system.

“The apple does not fall far from the tree. And you can see that there's a lot of things that go on at home so then if we're the only ones holding this kid accountable, it feels like it we're at a loss. ...if you don't have the parents that are on board, or at least attempt to try to...work on the behavior to make it better, then you're at a loss, like you're you're basically going up the hill with the boulder every day.”

Insufficient Staff Training and Experience

Participants indicated that new staff members were a hindering factor to the sustainability of the level system. They reported that new staff were mostly trained on-the-job by watching others implement the level system. This inadequate training created inconsistencies and misunderstandings among faculty members about the correct way to implement the various components of an intricate level system. Included in this category are three critical incidents from 18% of participants. It was reported that there was difficulty in making sure that all new employees were trained in a standardized way since each new employee was initially assigned a classroom position and their training took place by watching and learning from that particular group of faculty members. One participant noted,

“...you had then of course training from multiple different areas [in the school]. You know, multiple different TA partners. Right, you know a little thing here and a little thing that got slightly off course, maybe with each one. Right? Then pretty soon, people are like ‘Well wait a second I learned this way. Wait a second I learned that way.’ So there were a few little things that slipped through the cracks and definitely needed a little bit of a readjustment.”

It was also reported that in order to get new staff “trained on a fairly intricate system like adequately, ... word of mouth wasn't good enough, to get them trained.”

Staff Mental Health Concerns

While only mentioned by one participant, this category represents a hindering factor worth reporting and reflecting upon. Staff mental health was indicated as a hindering factor in the sustained use of the schoolwide level system. The participant indicated that teacher burnout alters the intended use of the level system. “Sometimes people get burned out on some kids, and really will use the system to remove the kid from the situation because they themselves can't deal with the kid adequately.” The participant further indicated that past experiences of victimization or trauma could also impact a teacher’s ability to use the level system in the best interest of the student.

“You know we're not dealing with neurotypical kids most of these kids have been traumatized before age three you know, and I think there has to be a lot of that understanding, you know that these kids are victims they were victimized in a in previous settings in their life, you know, and they are hurt, their full of pain and they don't know how to adequately deal with that pain and react to that that severity of pain. ... I think a lot of teachers react differently towards kids' behaviors. And I always just got to tell them, you know you got to really know who this kid is and where they're coming from to see how they react because it's the way teachers react to some of the behaviors that become more of an escalation than there needs to be ...it's difficult because we all come from various backgrounds. So you got to remove your background from the situation for the best interest of the kid.”

Wish List Items

Participants were asked to list items or factors they wished they would have had in the past or wanted in the future to continue the sustained use of the schoolwide level system at their school. Sixteen wish list items were identified from the 11 interviews conducted by the research team. Five categories were created from the wish list items and the majority of those categories corresponded with previous helping or hindering categories. For example, participants indicated a desire for improved training, expressing a desire for a standardized procedure which could be improved with something like a training manual. More participants, than in other categories, expressed a desire to have stronger reinforcements and a greater variety of rewards associated with the level system. Primarily, they indicated the desire for students to have more social experiences with typical peers. Making the rewards more immediate through adjusting the monetary system was also mentioned. Other participants expressed a wish for a simplified version of the level system, indicating a desire for a more straightforward system which would be easier for students to understand, and uniform point cards throughout the school. To keep engagement high, as well as fit the level system to individual needs, some participants indicated a desire to modify the movement criteria through the levels. All wish list categories, the items included within each category, and the percentage of participants within each category are summarized in Table 4.

One novel category was created from wish list items: Technology. Participants indicated a wish for additional technology, so all data could be tracked electronically, which included digital daily point cards and digital bankbooks for students as well as a means for adults to electronically record all quarterly reports and other data currently marked in binders on paper spreadsheets.

“We did make the improvement of making it more electronic by putting all the data online, which was a big change. But I think it's good that we're moving in that direction. But if there was a way to do it, I know, like on an iPad or something if ... the special ed kids had an iPad and they could just have a digital version of it [the point card] that you can mark.”

With technology, like a tablet for each student, daily points could be tracked electronically; staff members would mark digital point cards and students could continue to track and see their daily progress. Participants indicated, in other helpful incident categories, that using digital data made decision making easier, particularly when collaborating with other team members and family. “...our weekly meetings when we sit down and talk about all the students in the school, we have all the data displayed, of course, like you know through that software and it really is convenient.”

“We do use that [digital data] in our meetings on Fridays and look at it ...that's really helpful and we're able to look up and see if there's an incline or a decline, um, on behaviors, on how they're doing ... It's a nice visual.”

More access to digital data could help in other aspects of the level system and improve sustainability.

Another notable point of interest was that the majority of participants, when asked if there were things they wished had happened in the past or could happen in the future to improve the level system and sustainability, immediately expressed their fondness of the level system in its current state. Most participants initially made a comment indicating they didn't want to make any changes to it, for example, “I don't know, I really like the level system,” or “I don't know if I'd make any changes,” before then mentioning any wish list items.

Evidence of Effectiveness

A final question, not part of the CIT, was included in the interview protocol to determine what participants believed was evidence of the level system's success. Most participants highlighted evidence of effectiveness throughout the interviews, before the question was asked. "[I see] the progress that they make. ... that kid getting better and better and better." "We see those baby steps, but baby steps in our world are huge." Participants indicated that students were motivated on the level system because they were able to recognize their own growth. Immediate feedback was given to students which established clear expectations for them and resulted in less anxiety.

"A lot of these kids just need the consistent feedback, the immediate feedback and ... that reduces anxiety, I think it reduces the need to act out. When ... you're in a comfort zone, knowing what you need to do, I feel like behaviors decrease."

Students stayed motivated because they learned that each day was a new day, and all was not lost if they had an "off" day. "Every day's a new day. They start out with % every day, it doesn't hang over and they can have a bad day ... and the next day they're % again." Evidence of success was also indicated by participants as an increase in student responsibility and prosocial behaviors, as well as a decrease in aggressive behaviors. One participant highlighted, "it's helping them focus on the one thing they need to work on, as well as the other behaviors, to teach them, um, appropriate behavior that's gonna make them successful out in the world."

Discussion

Current research is lacking in information for effective implementation of level systems and is void of evidence indicating the sustaining factors of level systems. This study used the CIT to examine the perspectives of faculty members from a special school which has

implemented a schoolwide level system, for students with behavior disabilities, for over 30 years. Interviews focused on identifying the helping and hindering factors which contributed or deterred from the sustainability of their level system. Participants also shared wish list items and evidence showing that the level system has been effective for students. Categories were created, from reported helping, hindering and wish list items, which aligned with previous research on implementation and sustainability. Previous discussions from the research indicated that critical factors for effective level systems implementation were communication and consistency across time and people. Challenges with implementation also included the multifaceted nature of level systems, making them difficult to learn for both students and facilitators. Key sustainability indicators which aligned with the CIT categories were utilizing a team approach, using data to make team-based decisions, teacher buy-in and acceptability, and contextual fit.

Explanation of Key Findings

This research confirmed the need for utilizing a team approach in two particular ways: high rates of communication and building team rapport. Level system implementation requires constant communication, to ensure that teachers, students and parents have the same understanding of each part of the system, how it works, and the criteria associated with it. For this team, communication happened in a number of ways throughout each day/week. Weekly team meetings provided a forum for discussions about student needs, clarification on expectations, and a space for collaborative decision-making. During team meetings student data were used for decision-making about student programming. Communication was important at the end of each day when faculty members discussed each student and checked data for accuracy. Most critical was daily, walkie talkie, team communication when managing crisis situations where immediate input was needed. Building team rapport allowed faculty members to find a

shared sense of purpose and to work more effectively as a team by supporting each other in their work. As discussed in the lit review, teams who work together with a common purpose, sharing the workload, maintain morale (McIntosh et al., 2013). High morale was helpful in this setting because of the intense nature of working with students with EBD and the resulting emotional factors that come from managing students with persistent and intense behaviors. Philosophical differences among staff members hindered their ability to work together effectively as a team, particularly when personnel became entrenched in their beliefs about the “right” way to implement certain components of the levels system and the fundamental purpose of utilizing a level system in their school. Once teams were able to come together and communicate their differences of opinion or belief, they were able to find common ground and problem solve their disagreements.

Building student-teacher relationships was identified as a helpful category. Although building student relationships was not discussed in the previous literature as an implementation or sustainability factor, participants reported that building relationships with students, through meaningful feedback and encouragement, helped them feel like they were making a difference and that they were involved in something worthwhile. Giving feedback in meaningful ways helped personnel feel less like a robot and more like a coach. When teachers give feedback more frequently, research shows they tend to increase praise given to students (Partin et al., 2009). Increased teacher praise is associated with increased student engagement. A benefit of level systems is that teachers evaluate student behavior more frequently, which provides students with more consistent and specific feedback, and faculty members more opportunities to connect with their students. When faculty members developed relationships with students, they felt invested in

the student's daily outcomes and were motivated by the long-term changes in students' behaviors.

Staff scoring and accountability inconsistencies was a highly ranked hindering category. As discussed in the literature, consistency is essential for effective level system implementation. Staff inconsistencies in data collection and accountability undermine the entire system because of how it changes the value of reinforcers, which can then modify a student's motivation and interest. Furthermore, without good data, ineffective decisions are made, and students' needs are unclear. Using data for decision-making is a critical predictor of sustainability. But, if the data are not accurate or representative of a true picture of student behavior and schoolwide concerns, it cannot be as effectively used. As a result of regular data discussions, staff were more aware of scoring inconsistencies and how that impacted the decisions that were made. Technology was indicated as a wish list item that would help with improved data collection and efficiency in access to data.

Of the ten different factors identified by McIntosh et al. (2013) as critical to sustainability, the one most noticeably missing was administrative support, which was not identified as a CI by any participants. The sustainability of this school's level system happened in spite of administrative support or with lack of support from the new administration. Other aspects of sustainability which included staff commitment, integration into existing and new efforts, ongoing resources, perceived effectiveness, implementer skill and knowledge, teaming, efficiency, collection and use of data, and capacity building, were identified as helping or hindering factors within the categories of this research.

As reviewed in the literature, sustainability is affected by both the perception of effectiveness by implementers and actual effectiveness quantified with data (McIntosh et al., 2013). This research supports that idea. Participants indicated that level system effectiveness was shown through improved student outcomes including increased independence, motivation and responsibility. These behaviors are difficult to measure, yet participants indicated their endorsement of the level system based on these perceived ideas. Because teachers could see visible behavior changes, they sustained the level system believing it to be what contributed to their students' improved behavior.

Limitations

Limitations of this research result primarily from the limitations of the CIT. Since data are collected from participants through their recall of specific events, time and memory can impact the accuracy of that information. Additionally, there is potential error from interviewers who could lead or mislead participants and influence participant responses. This study is limited in its scope because of a limited number of participants. Data were only collected from one small setting, at one particular point in time. Interviews were conducted over a three-month period and researchers did not control for potential communication among participants. Faculty members could have talked with each other and shared their responses before all interviews were completed, thus potentially influencing another participant's response. Additionally, participants may be biased towards their school's level system because of the all-encompassing nature of it and having no other form of behavior management system to compare it with. Some participants have only known this one schoolwide behavior management system. Another limitation is the small number of critical incidents for all categories (81) which is below Flanagan's (1954) recommended minimum of 100 CIs.

Recommendations for Research

Future research should focus on replicating this study on a larger scale, with more participants over more settings. This could include not just self-contained schools, but self-contained EBD classrooms as well. Research could focus on how critical incident categories and participation rates hold true or change with a greater participant group from differing EBD settings. In addition, future research could focus on the quantitative aspects of level system implementation. Level systems operate with multiple components, making variable isolation difficult. Researchers could manipulate different components within the level system, through a single-subject research design and evaluate how eliminating different facets of a level system impacts student outcomes as well as faculty opinion about it.

Recommendations for Practice

This study has highlighted the comprehensive nature of level system implementation and sustainability. While challenging, research previously discussed indicated that the obstacles associated with implementation could be overcome with a clear grasp of effective behavioral principles and a good understanding of how to develop and implement effective systems. This requires a specific knowledge base and skill set, which would require training and resources to ensure staff members develop this foundation. Training of this nature was not mentioned as part of the onboarding for participants in this study. However, weekly meetings and a number of communication portals helped their team establish a forum for continued training and correction of drift. Practitioners can learn from this example that the amount of communication needed for sustainability is both varied and comprehensive.

Establishing a firm foundation for new employees could be more standardized to ensure uniformity in training. When new members join a team, particularly in this type of high stakes setting, it is helpful for them to see the success of the system and catch the vision of it before jumping right into the work. One way to accomplish this is through pre-training materials, which would include training modules on behavior basics coupled with hands-on experience. New faculty members would spend the first 30 minutes of their workday, over the course of their first four weeks on the job, participating in the modules. Modules would include video modeling examples and other visuals, along with a paired assignment to shadow a specific experienced employee while watching for a specific behavior. This would allow new employees to become more proficient as they learned and practiced one aspect of the system or element of implementation at a time. Role plays could also be incorporated to build confidence and proficiency. Modules could be incorporated into a digital training manual so employees can refer back to it with questions and it could also be used during regular faculty meetings to correct misunderstandings and drift.

Mental health concerns are real for those adults working in education, but particularly in settings with intense student needs. Incorporated into new staff member training should be a number of coping strategies, including those to use during crisis situations as well as those to implement as needed for individual mental health. Employees should be educated on the nature of their work but also be given the tools they need to manage their mental health in helpful ways.

Conclusion

This research adds to the limited research on level systems, illustrating their complicated nature and implementation challenges. This study also contributes to findings on implementation and sustainability by providing evidence for key indicators in implementation and sustainability

research including teaming, communication, and consistency in implementation and data collection. Although limited in scope, the findings can inform other schools about the helping and hindering factors to sustaining school wide behavior systems.

References

- Andersson, D. L., & Nilsson, S. (1964). Studies in the reliability and validity of the critical incident technique. *Journal of Applied Psychology, 48*(6), 398-403.
<https://www.doi.org/10.1037/h0042025>
- Butterfield, L. D., Borgen, W. A., Amundson, N. E., & Maglio, A. T. (2005). Fifty years of the critical incident technique: 1954-2004 and beyond. *Qualitative Research, 5*(4), 475-497.
<https://www.doi.org/10.1177/1468794105056924>
- Butterfield, L. D., Borgen, W. A., Maglio, A. T., & Amundson, N. E. (2009). Using the enhanced critical incident technique in counseling psychology research. *Canadian Journal of Counselling and Psychotherapy, 43*(4), 265-282.
- Cancio, E. J., & Johnson, J. W. (2013). Designing effective classwide motivation systems for students with emotional behavioral disorders. *Preventing School Failure, 57*(1), 49-57.
<https://www.doi.org/10.1080/1045988X.2012.664582>
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Pearson Merrill Prentice Hall.
- Eldar, E., Ayvazo, S., & Hirschmann, M. (2018). Descriptive analysis of the instructional control of teachers in a classroom of students with behavioral disorders. *Journal of International Special Needs Education, 21*(1), 14-20.
<https://www.doi.org/10.9782/2159-4341-20.2.111>
- Farrell, D. T., Smith, S. W., & Brownell, M. T. (1998). Teacher perceptions of level system effectiveness on the behavior of students with emotional or behavioral disorders. *The Journal of Special Education, 32*(2), 89-98.
<https://www.doi.org/10.1177/002246699803200203>

- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51(4), 327-358.
<https://www.doi.org/10.1037/h0061470>
- Kerr, M. M., & Valenti, M. W. (2009). Controls from within the classroom: Crises or conversations? *Reclaiming Children and Youth*, 17(4), 30.
- Klopfer, K. M., Scott, K., Jenkins, J., & Ducharme, J. (2019). Effect of preservice classroom management training on attitudes and skills for teaching children with emotional and behavioral problems: A randomized control trial. *Teacher Education and Special Education*, 42(1), 49-66. <https://www.doi.org/10.1177/0888406417735877>
- Kostewicz, D. E., Ruhl, K. L., & Kubina, R. M. (2008). Creating classroom rules for students with emotional and behavioral disorders: A decision-making guide. *Beyond Behavior*, 17(3), 14-21. <https://www.jstor.org/stable/24011934>
- Lane, K. L., Wehby, J. H., Little, M. A., & Cooley, C. (2005). Academic, social, and behavioral profiles of students with emotional and behavioral disorders educated in self-contained classrooms and self-contained schools: Part I—Are they more alike than different? *Behavioral Disorders*, 30(4), 349-361.
<https://www.doi.org/10.1177/019874290503000407>
- Mattison, R. E. (2011). Comparison of students classified ED in self-contained classrooms and a self-contained school. *Education and Treatment of Children*, 34(1), 15-33.
<https://www.doi.org/10.1353/etc.2011.0003>
- McIntosh, K., Horner, R. H., & Sugai, G. (2009). Sustainability of systems-level evidence-based practices in schools: Current knowledge and future directions. In Sailor, W., Dunlap, G., Sugai, G., & Horner, R (Eds.), *Handbook of positive behavior support* (pp. 327-352). Springer.

- McIntosh, K., Mercer, S., Hume, A., Frank, J., Turri, M., & Mathews, S. (2013). Factors related to sustained implementation of schoolwide positive behavior support. *Exceptional Children, 79*(3), 293-311.
- McIntosh, K., Ty, S. V., & Miller, L. D. (2014). Effects of school-wide positive behavioral interventions and supports on internalizing problems: Current evidence and future directions. *Journal of Positive Behavior Interventions, 16*(4), 209-218.
<https://www.doi.org/10.1177/1098300713491980>
- National Center for Education Statistics. (n.d.). *Search for public schools*.
https://nces.ed.gov/ccd/schoolsearch/school_detail.asp?ID=490042000308
- Oliver, R. M., & Reschly, D. J. (2010). Special education teacher preparation in classroom management: Implications for students with emotional and behavioral disorders. *Behavioral Disorders, 35*(3), 188-199.
<https://www.doi.org/10.1177/019874291003500301>
- Partin, T. C. Moore, Robertson, R. E., Maggin, D. M., Oliver, R. M., & Wehby, J. H. (2009). Using teacher praise and opportunities to respond to promote appropriate student behavior. *Preventing School Failure, 54*(3), 172-178.
<https://www.doi.org/10.1080/10459880903493179>
- Reddy, L. A., De Thomas, C. A., Newman, E., & Chun, V. (2009). School-based prevention and intervention programs for children with emotional disturbance: A review of treatment components and methodology. *Psychology in the Schools, 46*(2), 132-153.
<https://www.doi.org/10.1002/pits.20359>
- Reid, R., Gonzalez, J. E., Nordness, P. D., Trout, A., & Epstein, M. H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of*

Special Education, 38(3), 130-143.

<https://www.doi.org/10.1177/00224669040380030101>

Smith, S. W., & Farrell, D. T. (1993) Level system use in special education: Classroom intervention with prima facie appeal. *Behavioral Disorders*, 18(4), 251-264.

<https://www.doi.org/10.1177/019874299301800408>

Sutherland, K. S., & Singh, N. N. (2004). Learned helplessness and students with emotional or behavioral disorders: Deprivation in the classroom. *Behavioral Disorders*, 29(2), 169-

181. <https://www.doi.org/10.1177/019874290402900208>

Swinson, J., & Cording, M. (2002). Assertive discipline in a school for pupils with emotional and behavioural difficulties. *British Journal of Special Education*, 29(2), 72-75.

<https://www.doi.org/10.1111/1467-8527.00243>

Veneziano C., & Veneziano L. (1988). Knowledge of social skills among institutionalized Juvenile delinquents: An assessment. *Criminal Justice and Behavior*, 15(2), 152-171.

<https://www.doi.org/10.1177/0093854888015002003>

Wagner, M. M. (1995). Outcomes for youths with serious emotional disturbance in secondary school and early adulthood. *The Future of Children*, 5(2), 90-112.

<https://www.doi.org/10.2307/1602359>

Weeden, M., Wills, H. P., Kottwitz, E., & Kamps, D. (2016). The effects of a class-wide behavior intervention for students with emotional and behavioral disorders. *Behavioral Disorders*, 42(1), 285-293. <https://www.doi.org/10.17988/BD-14-12.1>

<https://www.doi.org/10.17988/BD-14-12.1>

Wehby, J. H., Lane, K. L., & Falk, K. B. (2003). Academic instruction for students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders*,

11(4), 194-197. <https://www.doi.org/10.1177/10634266030110040101>

Tables

Table 1

Participant Demographic Information

Participant	Gender	Current Role	Years Experience	Experience w/ LS	Training on LS	Additional Training
1	F	Support	25	25	On-the-job	Ongoing review
2	F	Teacher	19	4	On-the-job	Ongoing review
3	F	Support	12	6	On-the-job	Ongoing review
4	F	Teacher	24	19	On-the-job	College & PD
5	F	Teacher	10	5	On-the-job	Ongoing review
6	F	Support	5	3	On-the-job	Ongoing review
7	F	Teacher	10	5	On-the-job	College
8	M	Teacher	19	15	Job shadow	Ongoing review
9	F	Support	20	20	On-the-job	Ongoing review
10	M	Teacher	6	6	On-the-job	Ongoing review
11	M	Specialist	15	15	On-the-job	Ongoing review

Note. Experience implementing level systems (LS) is reported in years. PD = Professional

development.

Table 2*Category Agreement Rate*

Category	Incident Count		Percent Agreement
	Original Coding	Credibility Check	
Building Rapport and Teaming	6	4	66%
Relationships and Student Feedback	6	7	86%
Collecting, Recording, and Using Data	5	5	100%
Student Engagement and Buy-in	5	4	80%
Communicating and Collaborating	4	8	50%
Faculty Buy-in	4	1	25%
Adaptability	4	3	75%
Schoolwide Consistency	3	4	75%
Philosophical Differences	8	7	88%
Scoring & Accountability Inconsistencies	6	8	75%
Multifaceted System Requiring	4	3	75%
Nuanced Decision Making	4	2	50%
Community Factors and External Pressures	4	5	80%
Insufficient Staff Training and Experience	3	3	100%
Staff Mental Health Concerns	2	2	100%

Table 3*Incident Count and Percent Reporting for Helping Incident Categories*

Category	Count of Incidents	Percent of Participants
Building Rapport and Teaming	6	45%
Relationships and Student Feedback	6	45%
Collecting, Recording, and Using Data	5	36%
Student Engagement and Buy-in	5	36%
Communicating and Collaborating	4	36%
Faculty Buy-in	4	27%
Adaptability	4	27%
Schoolwide Consistency	3	27%

Table 4*Incident Count and Percent Reporting for Hindering Incident Categories*

Category	Count of Incidents	Percent of Participants
Philosophical Differences	8	55%
Scoring & Accountability Inconsistencies	6	45%
Multifaceted System Requiring Nuanced Decision Making	4	36%
Community Factors and External Pressures	4	18%*
Insufficient Staff Training and Experience	3	18%*
Staff Mental Health Concerns	2	9%*

Note. Percentages with an asterisk (*) did not meet the minimum 25% participation rate.

Table 5*Incident Count and Percent Reporting for Wish List Categories*

Category	Count of Incidents	Percent of Participants
Reinforcement	5	36%
Unique	3	18%*
Movement	2	18%*
Training	2	18%*
Technology	2	18%*
Simplification	2	9%*

Note. Percentages with an asterisk (*) did not meet the minimum 25% participation rate.

APPENDIX A

Review of the Literature

Children and adolescents with emotional and behavioral disturbances (EBD) are some of the most challenging students to support in the school settings. Students with EBD are impacted by emotional, behavioral or psychiatric disorders, yet for a number of reasons, do not always receive the support they need through special education services. Furthermore, when students with EBD do receive eligibility for special education services, they are not always classified under the category of Emotionally Disturbed (ED), the main designation for students with EBD (Forness et al., 2012). Under the Individuals with Disabilities Education Act (IDEA), students classified as having ED must demonstrate one or more of the following characteristics: (a) an inability to learn that cannot be explained by intellectual, sensory, or health factors, (b) an inability to build or maintain satisfactory interpersonal relationships with peers and teachers, (c) inappropriate types of behavior or feelings under normal circumstances, (d) a general pervasive mood of unhappiness or depression, or (e) a tendency to develop physical symptoms or fears associated with personal or school problems. These characteristics must persist over time and to such a degree that it impacts educational performance. However, because there is no further definition to the term *educational performance*, there is confusion about what qualifies as poor performance. Is poor performance solely poor academic achievement, limited functioning in forming and maintaining personal relationships with peers and teachers, or both? As a result, students with EBD are often under-identified or misidentified under IDEA (Forness et al., 2012). According to the National Center for Educational Statistics (NCES, 2021), during the 2019-2020 school year 5% of the 7.3 million students receiving special education services under IDEA were classified as ED, compared to 33% of students classified with learning disabilities and 19% with

speech language impairments. Considering that behavior is highly affected by the environment, and also the variability of behavior with time and setting, it is likely that some students with EBD are being served in other categories. For example, students with a diagnosed disability such as attention deficit hyperactivity disorder (ADHD), a disability characterized by impulsivity, inattentiveness, and hyperactivity, can legally be served under the category of Other Health Impairment if this is the primary disability which is impacting the student's educational performance. IDEA also requires students with disabilities to be educated in the least restrictive environment possible, alongside their typically developing peers, yet Farrell et al. (1998) report that students classified with EBD are placed in more restrictive settings and segregated from their peers more often than students in other disability categories. Additionally, because of the stigma attached with a classification of ED, which labels students as having mental health disorders, some teachers and parents are hesitant to choose this category for placement of students with EBD, choosing instead to seek classification in a more socially acceptable category, such as OHI (Forness et al., 2012).

Students with EBD demonstrate persistent behavior problems through both internalizing and externalizing behaviors that inhibit learning and negatively impact the classroom environment (Klopfer et al., 2019; Oliver & Reschly, 2010; Reid et al., 2004; Wagner, 1995). Internalizing behaviors, which are self-directed, include anxiety, nervousness, depression, somatization, and obsessive-compulsive behaviors. These behaviors, while not as noticeable or disturbing to the classroom, can immobilize some students and prohibit their learning (McIntosh et al., 2014). Far more disruptive to the classroom are externalizing behaviors, which are by definition directed towards the external environment. Examples of these behaviors include defiance (including ignoring and/or resisting a teacher's directives or requests), verbal aggression

(including criticizing and antagonizing peers), physical aggression, and/or destruction of property. Externalizing behaviors require deliberate attention at school because they can be harmful to both the student and others, and they impede the instructional process in the classroom (Reddy et al., 2009).

As a result, students with EBD struggle to develop academic skills, including attending to academic assignments by accurately completing classwork and homework (Oliver & Reschly, 2010; Reddy et al., 2009). Students with EBD typically perform academically 1-2 years below grade level, earn the lowest grades, fail more classes, and have the highest drop-out rate of any other category of special education classifications (Kostewicz et al., 2008; Reid et al., 2004; Sutherland & Singh, 2004). Low academic proficiency combined with high rates of absenteeism for students with EBD result in disengagement from school, thus perpetuating a low success rate in school (Wagner, 1995).

Challenging classroom behaviors may also lead to poor social relationships with peers and teachers. Reid et al. (2004) explained in their research that because externalizing behaviors are often irritating and create negative feelings in others, students with EBD can be alienated from healthy learning and socializing experiences. Additionally, students with EBD struggle to interpret and respond appropriately to social cues (Lane et al., 2005). Veneziano and Veneziano (1988) found that lack of social skills was highly related to behavior problems. Students with limited social skills awareness, who struggle to recognize and understand social cues, will likely have more behavior problems, thus perpetuating a poor educational performance in both academics and interpersonal relationships.

Teaching Students With EBD

Meeting the diverse needs of students with EBD is challenging, and teachers struggle to create an environment conducive to productive learning and improved behavior. Research shows that teachers of students with EBD too often fail to implement the empirically supported practices that would create this type of environment in both general education and self-contained classrooms (Kostewicz et al., 2008). Carr et al. (1991) indicated that teachers provide less rigorous educational instruction and fewer academic interactions to students who display problem behavior, compared to students who display fewer behavior problems. To avoid potential triggers from task demands and the resulting escalating behaviors, teachers avoid interacting with students who demonstrate problematic behaviors and choose instead to give their instructional time to students with more appropriate behaviors (Kostewicz et al., 2008). Students with EBD receive more teacher attention following inappropriate behaviors and limited teacher attention for choosing appropriate behavior. As a result, research shows that teacher behaviors are not predictive of praise or positive interactions to the EBD student and are therefore not reinforcing (Partin et al., 2009). This is in part due to the low rates of correct responding and the opportunities to respond provided to students with disruptive behavior. However, a lack of positive reinforcement is shown to produce a lack of confidence in students (Sutherland & Singh, 2004).

Alternative Placements for Students With EBD

Often, because of the complexities of educating students with EBD, these students are placed in alternative settings; 40% of students with EBD are educated in self-contained classrooms or self-contained schools (Mattison, 2011). Steinberg and Knitzer (1992), in their efforts to identify best practices impacting students with EBD, visited 26 programs across 13

different states. Although considered to be exemplary settings, they repeatedly witnessed what they considered “pervasive boredom and apathy,” including students sleeping at their desks and an excessive amount of time each day spent waiting. They also observed limited academic content and few academic challenges, a lack of effective teaching strategies used, and the absence of student differentiation (Steinberg & Knitzer, 1992). Wehby et al. (2003) indicates that teachers in self-contained classrooms of EBD students spend only 30% of the day in academic instruction, in an environment that is lacking in praise and high in reprimands. In an effort to avoid aggressive behaviors, less academic demand is given, resulting in less chance for academic growth. Additional research in EBD classrooms indicates that more time is spent in negative teacher-student interactions than in positive ones (Sutherland & Wehby, 2001), wait time is high, which can be a trigger for disruptive behavior (Kerr & Valenti, 2009), and there is a lack of clear instruction and structure (Kostewicz et al., 2008). Research shows that more disruptive behaviors occur in settings where classroom procedures and expectations are unclear and inconsistent, and also occur in classrooms where teachers respond in untimely and ineffective ways (Eldar et al., 2018). Students need proactive measures to help them identify and manage their problems. Highly skilled teachers are needed to create a safe and orderly classroom environment and to efficiently manage students with such complex needs. Yet, teachers of students with EBD report feelings of inadequate preparation and a lack of classroom management skills, as well as increased exhaustion and depersonalization (Cancio & Johnson, 2007; Ginns & Begeny, 2019; Klopfer et al., 2019; Oliver & Reschly, 2010).

Evidence-Based Practice in Alternative Schools Placements

As a result of the number of issues associated with educating students with EBD, there is a need for more effective and empirically supported practices implemented with fidelity in the

classroom. Empirically supported practices have considerable research supporting the effectiveness of particular interventions and the use of them in the classroom setting. In evaluating evidence-based practices in classroom management, Simonsen et al. (2008) organized their identified practices into five empirically-supported, critical features of classroom management, including: (a) maximize structure; (b) post, teach, review, monitor, and reinforce expectations; (c) actively engage students in observable ways; (d) use a continuum of strategies for responding to appropriate behaviors; and (e) use a continuum of strategies for responding to inappropriate behaviors. Structure promotes appropriate academic and social behavior. Students exhibit more on task and engaged behaviors as well as less aggression in classrooms that have explicitly defined rules, minimized crowding and distractions, and have a balance of teacher-directed and independent learning. When teachers post, teach, and review expectations, and provide feedback through active supervision, it decreases disruptive behavior and increases conflict resolution (Simonsen et al., 2008). Empirically supported practices that focus on positive behavior in the classroom include specific and contingent praise, group reinforcement contingencies, and token economies. Token economies improve desirable behavior and decrease undesirable behavior by motivating students with the use of tokens. When students demonstrate appropriate target behaviors, they immediately earn a token which can later be exchanged for a more attractive object or activity. Two empirically supported interventions that focus on responding to inappropriate behavior in the classroom include differential reinforcement and response cost. More a function of effect, response cost only occurs when a student's behavior decreases as the result of removing a previously determined reinforcer for the given behavior (Cooper et al., 2007).

A highly structured, organized, and positive environment is essential to the success of a self-contained classroom. Research shows that effective teachers monitor classroom behaviors through providing clear and explicit expectations, with steady positive feedback and consistently delivered consequences, to promote student accountability (Kostewicz et al., 2008; Swinson & Cording, 2002). This allows teachers to manage the classroom in a more preventative way (Eldar et al., 2018). To promote appropriate student behavior, Partin et al. (2009) recommends a purposeful design to the classroom, which promotes good behavior and sets students up for success. Teachers, as “agents of prevention,” use preventative tactics like purposeful praise and increasing opportunities to respond (questions, tasks, demands) to influence the classroom environment by consistently increasing appropriate behavior and decreasing inappropriate behavior (Partin et al., 2009). More research indicates that group contingencies can reduce problem behavior, as well as provide opportunities to teach appropriate classroom behaviors for students with EBD (Weeden et al., 2016). Increased teacher praise increases student engagement for at-risk students. Increased engagement is shown to mitigate academic failure and improve social and behavioral outcomes (Downs et al., 2019).

Systemic Supports in Alternative School Settings

Research indicates that students with emotional and behavioral challenges perform better academically and have fewer behavioral challenges when empirically supported interventions are implemented in the classroom. Yet, research also suggests that teachers of students with EBD are not utilizing these practices to a degree that is improving student outcomes, providing instead fragmented programs and isolated efforts. The reasons for ineffective implementation vary but could include one or more of the following: inadequate teacher preparation and training; lack of teacher support through performance feedback, mentoring, or coaching; high levels of stress; or

increased exhaustion and depersonalization. Teachers need a comprehensive strategy that will help them to provide needed classroom structures and management procedures with clarity and consistency. When students are placed in special schools, the most restrictive educational environment, it is assumed the placement is warranted as a result of severe problem behaviors demonstrated over time. Students in this setting require more intensive support (Hendrickson et al., 1998). Often this placement is viewed by school personnel as the “last resort” within the school system and students may intuit this as well. Consequently, students in this placement need an increase of motivation to improve their circumstances. To meet the needs of both teachers and students, a systematic approach is needed. Approaching the management of students with EBD in special schools in a systematic way allows teachers and administrators to coordinate all efforts and aspects of intervention towards a specific and common objective.

Level Systems

An example of a systemic support system used with EBD students in alternative placements is a level system. With the ultimate goal of self-management, level systems are designed to shape students’ desired social, emotional, and academic behaviors, through the application of the principles of applied behavior analysis (ABA). Level systems allow students to earn privileges for exhibiting prosocial behaviors during specific scoring periods. Students can move through different levels of criteria based on consistent demonstration of learned behaviors. As expectations increase, so does access to more desirable privileges. According to Cancio and Johnson (2013) level systems should also be designed to promote generalization, increase self-sufficiency, and provide an organizational framework for the classroom of EBD students. Level systems provide a clear structure and consistent order in the self-contained classroom to promote prosocial behaviors in students with EBD. Fundamentally, level systems are an independent

group contingency where performance criteria are given to the entire group, but reinforcement is delivered individually based on each student's unique performance (Cooper et al., 2007).

Teachers can use level systems for supporting both social and academic behaviors and the system can easily be paired with social skills and/or self-management curriculums. Creating a level system helps support teachers in creating a structured classroom with explicit rules and procedures. There is some flexibility in using level systems as they can be paired with other management technologies, which can be removed for more simplicity. They can also be used with an individual, a group of students, or an entire school, depending on student need. Though the purpose of level systems may vary from setting to setting, the ultimate goal of a level system is self-management through increased student responsibility and generalization of skills in the least restrictive environment.

Critical to the success of well-designed level systems, are clearly defined behavioral expectations. These expectations are stated as the desired student behaviors, or target behaviors. Target behaviors must be observable and measurable and typically fall within general categories of social and academic behaviors (e.g. use appropriate language or stay on task; Cancio & Johnson, 2013). Once target behaviors are clearly defined and performance criteria is set, reinforcement contingencies are established. Typically, point values are assigned to each expected behavior. Students may earn points for exhibiting desired behaviors or possibly lose points for demonstrating adverse behaviors. All those who participate in a level system, including students, teachers, and administration, need to clearly understand the defined target behavior, as well as the expectations and consequences associated with each desirable behavior.

Each level of the system could be considered a different contingency arrangement on a hierarchy of behavioral expectations. Movement through each level of the intervention is contingent upon a very precise criteria of behavior expectations as well as a time frame of demonstrated sustained performance for each level. Additionally, for adherence to each level's strict criteria, student's earn specific privileges associated with each level. Often a token economy is paired with a level system. In this case, for demonstration of established criteria, students earn points or tokens to purchase the rewards and privileges available at each level. The token economy allows students to earn points which they can exchange for privileges, activities or tangibles. Lower levels require mastery of fewer behaviors and fewer privileges are available; while higher levels in the system require mastery of more behaviors and greater privileges associated with these levels. Privileges act as positive reinforcers, motivating students to increase their appropriate behaviors and decrease their negative behaviors as they move through the levels of a system. Level systems have also been paired with response cost. In this situation students lose points for inappropriate behaviors, which acts as a punishment and makes contact with positive reinforcement less likely. Using level systems allows behavior to be shaped and reinforcement to be faded, as motivation changes from external to more intrinsic along the way. At the highest level, performance of recently acquired behavior skills is expected to be demonstrated in natural settings, with natural reinforcers sustaining the behavior (Cancio & Johnson, 2007, 2013; Cruz & Cullinan, 2001; Scheuermann et al., 1994; Smith & Farrell, 1993; Walker & Shea, 1995).

Level systems are potentially beneficial to students with EBD in many ways. First, students with EBD start to see a connection, a cause-and-effect relationship between their behaviors and the consequences that follow. No longer is the adult seen as the cause of the

negative experiences in the student's life because they realize their earned privileges are the direct result of their personal improvements. This increases social maturity as students learn to take responsibility for their actions. Second, the design of level systems creates structure and routine which provides students with EBD a sense of order, fairness and security. Rules are explicitly stated, and privileges are contingent on a well-defined performance criterion. Third, level systems support generalization as the skills necessary for success in an inclusive setting are learned and practiced with increasingly less reinforcement. Daily point trackers help students learn self-monitoring skills and inventory their own behavior in preparation for transitioning to an inclusive environment (Cancio & Johnson, 2013; Smith & Farrell, 1993). Fourth, teachers evaluate student behavior more frequently, which provides students with more consistent and specific feedback. When teachers give feedback more frequently, they tend to increase praise given to students.

Foundations of Level Systems

Determining the precise origin of level systems is complicated. Some would argue that level systems have been around for quite some time. For example, Bremner (as cited in Smith & Farrell, 1993, p.252) refers to the writings of Charles Dickens in the mid-1880s where Dickens writes of the Boston House of Reformation for juvenile offenders. At the Boston House, offenders were assigned the lowest class of four upon entry and worked their way to higher levels through good behavior. Walker and Shea (1995) trace the origins of level systems to residential treatment centers in the US. They detail the system used in the 1970s at the Holy Cross Program in New York City. This adolescent substance abuse program used a four-level system, including a token economy, to provide feedback and positive reinforcement to improve tolerance for delayed gratification. Kazdin (1977) points out that level systems have frequently

been used with a token economy in psychiatric hospitals as a way to move patients towards the ultimate goal of discharge. Examples include level programs described by Atthowe and Krasner (1968) for psychiatric patients; the level system and token economy of Achievement Place in 1971 where pre-delinquent youth moved from token level to merit level too homeward bound level; and the hospital for drug addicts (Melin & Gotestam, 1973) who advanced through three levels including detoxification, treatment, and rehabilitation. While there is no exact date to the inception of level systems, Smith and Farrell (1993) suggest their contemporary development stems from the behavioral technology of the early 1960's in the field of ABA. While the true origin may be unclear, the use of level systems with the specific population of emotionally and behaviorally disturbed is consistent throughout history.

Although level systems appear to have evolved from institutions, there is limited research on the use of them for educational settings. Kazdin (1977) explains that level systems are not new to education because public school is essentially a level system in and of itself, moving from grade level to grade level based on performance. Furthermore, the movement from elementary to secondary to college and to post graduate school can't be viewed as a level system as each level of completion earns higher paying jobs and greater benefits.

Level Systems in Schools

The Engineered Classroom

Hewett's work with children exhibiting emotional and behavioral disorders led to his development of the Engineered Classroom in 1968. Based on the behavior modification model for education, Hewett saw the teacher as the behavioral engineer tasked with the responsibility of engineering a classroom to maximize student success. The design included both the learning triangle and a hierarchy of educational skills including seven levels of developmentally

appropriate skills essential for educational success (Hewett, 1967). Both of these elements influenced current level systems. The learning triangle gave the teacher a structure for engineering the classroom which included (a) educational goals directed towards helping students achieve competence at one or more of the levels of learning in the hierarchy (b) positive consequences for manifested efforts (c) highly structured conditions maintained for optimal learning. The hierarchy of levels, or educational skills, promoted moving from level to level through developmentally appropriate steps.

The Madison School Plan

The engineered classroom strategies were applied to a resource classroom in 1972, a classroom designed to help students with disabilities transition into a regular classroom. It was called the Madison School Plan. The plan utilized four levels. Mastery of specific behaviors at each developmentally appropriate level increased time in the regular classroom, with level four being placement in the regular classroom and included all of the expectations, behavioral and academic, of that setting. This was the first documented use of a treatment package aimed at the generalization of skills and mainstreaming of students (Cancio & Johnson, 2007; Smith & Farrell, 1993).

Achievement Place

During the 1970s research independent of Hewett's Engineered Classroom was being done at Achievement Place, a home for pre-delinquent boys. Researchers used a token economy as a behavior management system and successfully changed the boys' behavior. Several modifications to the plan were made over the research period, adding various techniques to the system, resulting in a comprehensive program (Hewett, 1967).

Boys Town

Long before the 1960s technology of radical behaviorism was being developed, Father Flanagan used a \$90 loan, in 1917, to open a boys' shelter for troubled teens called Boys Town (Jendryka, 1994). Father Flanagan believed, "There are no bad boys; only bad environments, bad example, bad thinking," and this became the guiding philosophy of his work (Jendryka, 1994, No Such Thing as a Bad Boy section, para. 1). This passion led him to take in numerous renegades from off the streets of Omaha, Nebraska and rehabilitate them. He believed that with the right environment and positive reinforcement all boys could be reformed. The purpose of Boy's Town was to create an environment that could encourage and preserve good character (Jendryka, 1994). Originally founded upon principles of hard work, structure, and religious values, boys from off the streets lived together in a family environment. Father Flanagan kept the boys busy farming, performing, and serving in the community, which significantly reduced problem behaviors. Now, more than 100 years later, Boys Town has grown into an education model deeply grounded in principles of ABA and social learning theory.

Consistent with this foundation, one of the primary components of the Boys Town Education Model is a graduated level system with three levels of support. The purpose of the levels is to promote generalization and maintenance of appropriate behavior through fading dependence on reinforcers. These two components together form the Boys Town Motivation System. Students' progress through three levels. Initially, students receive frequent and immediate feedback through points earned or lost. Use of feedback and points are faded as the student is able to demonstrate appropriate social skills behavior in a variety of social settings and their successful interactions become the predominant motivating factor. The Boys Town Education Model of today is designed to support students who have been labeled as having

discipline problems, as well as those identified as being emotionally disturbed, through explicitly teaching social skills. By 1990 the Boys Town Education Model was being used in 20 states. Although there is only non-empirical evidence from the participating school districts, reports indicate satisfaction and success in their implementation of the system, particularly with behavior disordered students (Wells, 1990).

Level Systems in Special Education

In response to federal government demands to support students with EBD in the least restrictive environment, use of level systems in classroom settings rose in popularity during the 1980s. Smith and Farrell (1993) suggest that level systems at that time were handed down as an effective technology for behavior management with little research to support the efficacy of its use in schools. Their review of the literature described only three available studies that could be used as evidence in support of the use of level systems in the classroom. These were the only three studies that collected and analyzed data in any way, the other studies being descriptive in nature.

Special Schools

In a descriptive case study Klotz (1987) explained the implementation of school-wide interventions for emotionally disturbed adolescents. The program entitled the Behavior Management Level System (BMLS) combined a token economy with a hierarchical level of expected behavior mastery. The study took place during the 1983-1984 school year at a bridge school, a school for adolescents with emotional disturbance in Montgomery County, Maryland. The 36 students enrolled at the school participated in the BMLS. Students were given daily point sheets to track their progress. Up to two points could be earned during each class forward to hearing to class rules, which were operationalized and consistent across all classrooms and the

school. These rules included: on time, in place, on task, respect for property, appropriate language, required materials, homework, ignoring inappropriate behavior, refraining from inappropriate gestures, and refraining from set-ups. Students also earned points for demonstrating appropriate behaviors in areas that were particularly challenging for them. Performance criteria for behaviors was determined by the current level of the student. There were six levels. Students were expected to take more responsibility for their behavior as they progressed through the levels. They demonstrated this by earning more points each day. At the highest level, behavior was monitored through a weekly behavior contract; and daily point sheets were discontinued. Greater privileges were available as students moved up through the levels which included activities that allowed more student choice unless an adult supervision (Klotz, 1987).

Implementation took place in two phases, after an initial staff and parent training. During fall semester the training phase was initiated and only the point system was put into effect. The levels were implemented during the spring semester. Baseline data was collected during the training phase and used to make placement decisions for the implementation phase. Evaluative information was collected through student progress data, student preference votes, and a staff questionnaire. Student progress data indicated that 24 of the 36 students (71%) progressed one or more levels or maintained a high level of performance (either level three or four) during the implementation phase. Results of the student vote were nearly evenly split with 49% preferring the point system only and 51% in favor of the BMLS. The BMLS Evaluation Questionnaire was completed by 100% of the staff at the end of the implementation phase. It was organized into six categories including: planning and placement, mechanics of levels, preference, effectiveness, staff perception of parent understanding, and usefulness of class rules and target behavior, each

answered on a six-point scale. Teacher ratings were as high as 84% in preference for the BMLS over the point system alone, and 80% in using the system to make educational goals for students. The majority of teachers (79%) agreed the BMLS was a highly effective program and 78% of teachers ranked the usefulness of class rules as highly effective. Areas of mixed results included the mechanics of the levels, 69% of staff ranked criteria for levels just right and 26% scored it as too difficult; and in staff perception of parent understanding of the program (Klotz, 1987).

Resource Room

With limited evidence of the effectiveness of level systems in more inclusive settings, Mastropieri et al. (1988) implemented a level system into high school English resource classes for students with learning disabilities and/or behavior problems, to provide data on specific student outcomes, both behavioral and academic. The intervention took place at a rural high school in the west, with Caucasian students in 9th - 12th grade, during a daily 70min English class. Students were both male and female with average socioeconomic status and varying cognitive abilities.

The purpose of the level system in the class was to decrease student talk outs and out-of-seat behavior for the 15 students in the class. Four colored levels were implemented, and students were required to wear a name tag in the color of their current level. In addition to the classroom rules and each level has specific rules and privileges associated with it. Students self-monitored their behavior through the levels and the interval length was increased as students moved to less restrictive levels. Students could also earn additional privileges as they moved through the levels. At the end of each week students could request level changes that were determined during a class meeting and student vote. Granted level changes were accompanied by a new contract signed by the student, teacher, and three peer witnesses chosen by the student.

Baseline data were collected for one week prior to the intervention for both the whole class and each individual student by taking a daily average frequency for each behavior. The same measurement was used during the three weeks of implementation. Data analysis indicated that both behaviors decreased after the implementation of the level systems and all 15 students decreased talking out and out-of-seat behaviors. Upon completion of the three-week intervention period, students were asked to complete a questionnaire rating their attitude towards the criteria associated with each level. Students rated each level on a 1 through 5 scale, with one being very negative and five being very positive. Not surprisingly, the students rated the levels more negatively with more restrictive levels.

The purpose of the intervention, in the second classroom, was to improve assignment completion and accuracy. It took place in the same high school in a similar English resource class with four Caucasian boys with the same disabilities as in room #1. In this intervention the levels were modified to combine the two highest levels together because of their similarities; but otherwise, the system had the same criteria and functioned the same as the previous. Data were collected on behavior during the first three weeks of baseline and four weeks of intervention. This was followed by one week of return to baseline conditions and then the levels were re-implemented for one week. Data analysis for this classroom shows that student work completion and accuracy rate improved during implementation, dropped during the reversal phase, and increased again during re-implementation. This indicates a functional relationship between student behavior and use of levels. This design, however, incorporated a certain amount of peer pressure and it is difficult to determine if the behavior change is due to the levels being used or the peer pressure felt when the levels were in effect.

Self-Contained EBD Students in Mainstream Classrooms

Brennock et al. (1989) describe the mainstreaming process at Oak Park and River Forest High School in Oak Park, Illinois. The EBD program used a level system with a continuum of services aimed at transitioning students back into a regular education classroom through the organized and structured support of the system. To address the mainstreaming concerns of isolation of the EBD student from positive peer models, and transportation between campuses, Oak Park and Forest High Schools moved the EBD students to the main campus and implemented a level system. The goal of the system was to have all EBD students fully mainstreamed into regular classes. Participants in the five-semester program were 97 students, the majority with diagnoses associated with the American Psychiatric Association DSM-III categories, including major depressive disorder, oppositional disorder, conduct disorder, attention hyperactivity deficit disorder. As students progressed through the levels, each successive level had less structure and support. Students were given more responsibility for demonstrating what they had learned in previous levels. Requirements for the program were strict, including no absences or unexcused tardiness for each six-week grading period, completion of all class assignments and homework, and maintaining a C or higher average, to promote success in mainstream classes. Students were provided a mainstream facilitator to help with communication between all teachers, the student and parents. Students were also provided with additional support staff including special educators, a school psychologist, and social workers to help manage the anxiety associated with the transition. Outcome data was collected by comparing credits earned and attendance records prior to and after placement in the program. Students in the program increased in both the number of classes passed and in attendance. It is

unclear if the increases in performance were due to the initial criteria, the amount of personal support staff, the level system, or simply maturation over five semesters.

Implementation Challenges

The use of level systems has often exceeded existing research. Farrell et al. (1998) surveyed 172 out of 200 randomly selected teachers of students with EBD, across grade levels and settings, in a southeastern state. At that time, 122 teachers of students with EBD were using level systems (71%). Teachers in special schools used level systems more than any other group within a service delivery model (93.3%) followed by teachers in self-contained classrooms (75.2%). The authors concluded that “because so many teachers are using level systems, it is important to continue with thorough and systematic investigations of level systems so that their efficiency and effectiveness can be optimized for the benefit of all students” (p. 98). Yet research on level systems over the past few decades is limited, with much of the research published in grey literature.

One possible reason for the limited research on level system effectiveness in education could be because of the challenges in implementing level systems. Level systems have several components to design and implement which can make them complicated to utilize and difficult to research. Cancio and Johnson (2013) explain that there are eight steps to effectively implementing a level system: (a) identify target behaviors and develop some type of daily tracking sheet, (b) determine a timeframe for student feedback, (c) assign the point value for identified behaviors, (d) establish criteria for each level of the system and the movement between each level, (e) determine the privileges associated with each level in the system, (f) identify how and when students will have access to the reinforcers, (g) determine how points earned and spent will be tracked, and (h) develop a system for monitoring student progress as well as overall

system effectiveness. Since level systems are this extensive, consistency across time and people is critical. However, turnover in schools is constant, making sustained efforts in level system implementation and use challenging. Therefore, communication between teachers, staff, parents, and students is critical. Teachers, students, and parents need to learn about the behavioral expectations and privileges associated with each level, as well as the system for tracking student behaviors and the criteria for moving between levels. When combined with a token economy or a response cost system, additional criteria are implemented for earning or losing points and privileges. Cancio and Johnson (2013) explain that many of the obstacles associated with implementation can be overcome with a clear grasp on effective behavioral principles and a good understanding of how to develop and implement effective systems. Yet, this can necessitate considerable time and money to be spent on personnel training.

Special Schools

To implement a sustainable level system in special schools, Klotz (1987) noted that training took place over the course of an entire semester. Initially only points were used to allow for thorough training of staff and students as well as parents. During phase two of implementation, the levels were added. This allowed the change to a more comprehensive management system to happen gradually. Additionally, this school applied behavior principles known to improve the effectiveness of token economies to their level system. For example, they allowed students to preselect backup reinforcers and faded the use of tokens, replacing them with praise and activities to promote generalization, as these events are more naturally occurring in the school environment.

Resource Classroom

Mastropieri et al. (1988) provided no information on implementation challenges in the resource room. However, the environment was very structured which allowed for more effective implementation. Only one classroom was used for this study, so there may have been limited faculty to train.

Mainstream Classroom

Although no implementation or training information was provided in the mainstream classroom study (Brennock et al., 1989), this level system intervention planned for facilitators to create and maintain clear lines of communication and support staff to assist in the flow of all communication which could affect a student, including: teachers, parents, students, support staff, and administrators. Not all districts, schools or classrooms may have this level of support for students.

For effective implementation, teachers must know the rules, keep consistent records, and faithfully monitor behaviors and deliver feedback (Cruz & Cullinan, 2001). Implementation requires buy-in and cooperation from staff members. Furthermore, without a clear understanding of behavioral principles and the needed training, level systems can become primarily punitive and may rely too much on punishment as a tool for behavior reduction. Hewett (1967) cautioned, “One major problem with what has arisen from presenting the engineered classroom design to teachers in the field bears mention. Some of them are so desperate for ideas and directions to increase their effectiveness with emotionally disturbed children that they react to superficial aspects of the design and somewhat randomly apply them in the classroom” (p. 466). All of these problems make sustainability very challenging.

Sustainability

Sustainability is the process of sustained implementation with high fidelity over time (McIntosh et al., 2013). Sustainable practices help offset the challenges of continual change in schools. For example, student enrollment and faculty naturally change each year. Schools need effective practices that can be taught to new hires and new students over time. When schools stay with effective interventions over time, rather than trying something new every few years, these interventions become institutionalized. Institutionalization is usually accompanied by enhanced personnel training and integration of the intervention into the school culture, thereby enhancing cost-efficiency. Student outcomes are more likely to be improved when effective practices are implemented and sustained (McIntosh et al., 2009). There are some key indicators of sustainability within education in general: administrator support, a team approach and use of data to make team-based decisions. Teachers buy-in, as well as acceptability, contextual fit and the degree of individual and agency collaboration are also key indicators of sustainability within education. However, there is no research to indicate if any of these factors contribute to level systems sustainability.

There is, however, sustainability research on school-based interventions. Schoolwide positive behavior interventions and support (SWPBIS) is a framework which utilizes a systems-level approach for supporting student behavior. Preventative by design, SWPBIS incorporates principles of applied behavior analysis, contextual fit, data-based decision making and stakeholder collaboration to improve student outcomes (Andreou et al., 2015). Level systems utilize these same principles in supporting student behavior. While research on the sustainability of SWPBIS is limited, the existing literature can be used as a template to inform this study on sustainability of level systems.

In the literature there is a model of sustainability for schoolwide practices, proposed by McIntosh et al. (2009). At the foundation of the model are three frequently repeated steps: identifying valued outcomes; identifying and modifying practices; and implementing practices. Schoolwide practices that produce desired outcomes are seen as viable and are continually implemented and modified based on these three steps. It is throughout this process that contextual factors may improve or impede sustainability. For this model contextual factors are grouped into four hypothesized categories assumed to work together for improved sustainability. Priority includes staff commitment, administrative support, integration into new and existing efforts, and ongoing resources. Heightened priority leads to improved implementation which contributes to effectiveness and efficiency. Effectiveness includes perceived effectiveness, implementer skill and knowledge, and teaming. Efficiency is a standalone category used to evaluate available resources. Continuous regeneration includes collection and use of data, and capacity building, which constantly act upon the three other categories (McIntosh et al., 2013).

Qualitative Research in Sustainability

Qualitative research allows the more subjective aspects of sustainability to be identified through careful observation of interventions and firsthand communication with participants about their personal experiences. Andreou et al. (2015) states, “Qualitative research asks questions regarding ‘how’ or ‘why’ sustainability occurs under real-world conditions, which can help the field more fully understand sustainability, as well as what steps can be taken to enhance it” (p. 158). Within the field of education there is an increased need for qualitative research because this method will provide a better conceptualization of interventions (Andreou et al., 2015). This can lead to more thorough evaluation of practices. To better understand the level system technology Smith and Farrell (1993) suggest that research “will demand varied and intricate methodology,

complex analysis, and applicable results. Extensive ongoing research, both quantitative and qualitative, needs to be implemented to uncover the nature of level system use” (p.262). More substantive research is needed on level systems since much of the literature is descriptive in nature, explaining the process for designing and implementing a level system. The critical incident technique is a potential qualitative method to provide more substantive research.

The Critical Incident Technique (CIT) was developed as a research method by Dr. Flanagan in 1954 (Butterfield et al., 2005). Grounded in industrial and organizational psychology, CIT was created as an exploratory method helpful in the early stages of research, used to identify helping and hindering factors in specific experiences or activities. Because the focus of CIT is on real-life experiences, researchers gain insight into real-world problems through the analysis of observable behavior of those in the field doing the work. This method is appropriate when seeking to learn more about the less understood factors that contribute to or detract from effective performance of such experiences and activities (Butterfield et al., 2005; Butterfield et al., 2009). Within the CIT methodology, observable and measurable events perceived as helping or hindering the activity or experience of focus, are called critical incidents. Kent McIntosh (2016) is known to have said that CIT is “the most quantitative of the qualitative approaches” for research (personal communication, 2016). CIT utilizes research data that is both qualitatively obtained, through interviews and observation, and quantitatively analyzed, through coding and credibility checks to construct theories about specific practices.

Statement of the Problem

Smith and Farrell (1993) further stated, “Many questions will come from future research that will contribute to productive professional discourse about the use, design, and implementation of level systems as a way to assist students in managing their own behavior” (p.

262). One of these questions is, “Given the number of implementation challenges, what factors contribute to the sustained use of a level system?” Understanding what helps or hinders sustainability within level systems will help teachers and school interventionists know what factors to consider when implementing level systems and potentially other behavior intervention plans as well. This knowledge can be applied to settings where students with or without EBD are being taught. While the CIT methodology has been used for other educational practices, it has not been used as a way to measure level system sustainability. Currently, there is no empirical evidence to indicate which factors contribute to the sustainability of level systems.

References

- Andreou, T. E., McIntosh, K., Ross, S. W., Khan, J. D. (2015). Critical incidents in sustaining school-wide positive behavioral interventions and supports. *The Journal of Special Education, 49*(3), 157-167. <https://www.doi.org/10.1177/0022466914554298>
- Atthowe, J. M., & Krasner, L. (1968). Preliminary report on the application of contingent reinforcement procedures (token economy) on a "chronic" psychiatric ward. *Journal of Abnormal Psychology, 73*(1), 37-43. <https://www.doi.org/10.1037/h0025439>
- Brennock, T. M., Zemitzsch, A., & Simon, D. J. (1989). A high school behavior disorder program focused on mainstreaming. *The Pointer, 33*(3), 27-31. <https://www.doi.org/10.1080/05544246.1989.9944807>
- Butterfield, L. D., Borgen, W. A., Amundson, N. E., & Maglio, A. T. (2005). Fifty years of the critical incident technique: 1954-2004 and beyond. *Qualitative Research, 5*(4), 475-497. <https://www.doi.org/10.1177/1468794105056924>
- Butterfield, L. D., Borgen, W. A., Maglio, A. T., & Amundson, N. E. (2009). Using the enhanced critical incident technique in counseling psychology research. *Canadian Journal of Counselling, 43*(4), 265-282.
- Cancio, E., & Johnson, J. W. (2007). Level systems revisited: An important tool for educating students with emotional and behavioral disorders. *International Journal of Behavioral Consultation and Therapy, 3*(4), 512-527. <https://www.doi.org/10.1037/h0100820>
- Cancio, E. J., & Johnson, J. W. (2013). Designing effective classwide motivation systems for students with emotional behavioral disorders. *Preventing School Failure, 57*(1), 49-57. <https://www.doi.org/10.1080/1045988X.2012.664582>

- Carr, E. G., Taylor, J. C., & Robinson, S. (1991). The effects of severe behavior problems in children on the teaching behavior of adults. *Journal of Applied Behavior Analysis, 24*(3), 523-535. <https://www.doi.org/10.1901/jaba.1991.24-523>
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied Behavior Analysis* (2nd ed.). Pearson Merrill Prentice Hall.
- Cruz, L., & Cullinan, D. (2001). Awarding points, using levels to help children improve behavior. *Teaching Exceptional Children, 33*(3), 16-23. <https://www.doi.org/10.1177/004005990103300303>
- Downs, K. R., Caldarella, P., Larsen, R. A. A., Charlton, C. T., Wills, H. P., Kamps, D. M., & Wehby, J. H. (2019). Teacher praise and reprimands: The differential response of students at risk of emotional and behavioral disorders. *Journal of Positive Behavior Interventions, 21*(3), 135-147. <https://www.doi.org/10.1177/1098300718800824>
- Eldar, E., Ayvazo, S., & Hirschmann, M. (2018). Descriptive analysis of the instructional control of teachers in a classroom of students with behavioral disorders. *Journal of International Special Needs Education, 21*(1), 14-20. <https://www.doi.org/10.9782/2159-4341-20.2.111>
- Farrell, D. T., Smith, S. W., & Brownell, M. T. (1998). Teacher perceptions of level system effectiveness on the behavior of students with emotional or behavioral disorders. *The Journal of Special Education, 32*(2), 89-98. <https://www.doi.org/10.1177/002246699803200203>
- Forness, S. R., Freeman, S. F. N., Paparella, T., Kauffman, J. M., & Walker, H. M. (2012). Special education implications of point and cumulative prevalence for children with

- emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders*, 20(1), 4-18. <https://www.doi.org/10.1177/1063426611401624>
- Ginns, D. S., & Begeny, J. C. (2019). Effects of performance feedback on treatment integrity of a class-wide level system for secondary students with emotional disturbance. *Behavioral Disorders*, 44(3), 175-189. <https://www.doi.org/10.1177/0198742918795884>
- Hendrickson, J. M., Smith, C. R., Frank A. R., & Merial C. (1998). Decision making factors associated with placement of students with emotional and behavioral disorders in restrictive educational settings. *Education and Treatment of Children*, 21(3), 275–302. <https://www.jstor.org/stable/42940508>
- Hewett, F. M. (1967). Educational engineering with emotionally disturbed children. *Exceptional Children*, 33(7), 459-467. <https://www.doi.org/10.1177/001440296703300703>
- Jendryka, B. (1994). Flanagan's island. *Policy Review*, 69, 44-51.
- Kazdin, A. E. (1977). *The token economy: A review and evaluation*. Plenum Press.
- Kerr, M. M., & Valenti. M. W. (2009). Controls from within the classroom: Crises or conversations? *Reclaiming Children and Youth*, 17(4), 30.
- Klopfer, K. M., Scott, K., Jenkins, J., & Ducharme, J. (2019). Effect of preservice classroom management training on attitudes and skills for teaching children with emotional and behavioral problems: A randomized control trial. *Teacher Education and Special Education*, 42(1), 49-66. <https://www.doi.org/10.1177/0888406417735877>
- Klotz, M. E. B. (1987). Development of a behavior management level system A comprehensive school-wide behavior management program for emotionally disturbed adolescents. *The Pointer*, 31(2), 5-11.

- Kostewicz, D. E., Ruhl, K. L., & Kubina, R. M. (2008). Creating classroom rules for students with emotional and behavioral disorders: A decision-making guide. *Beyond Behavior, 17*(3), 14-21. <https://www.jstor.org/stable/24011934>
- Lane, K. L., Wehby, J. H., Little, M. A., & Cooley, C. (2005). Academic, social, and behavioral profiles of students with emotional and behavioral disorders educated in self-contained classrooms and self-contained schools: Part I—Are they more alike than different? *Behavioral Disorders, 30*(4), 349-361. <https://www.doi.org/10.1177/019874290503000407>
- Mastropieri, M. A., Jenne, T., & Scruggs, T. E. (1988). A level system for managing problem behaviors in a high school resource program. *Behavioral Disorders, 13*(3), 202-208. <http://www.jstor.org/stable/23886113>
- Mattison, R. E. (2011). Comparison of students classified ED in self-contained classrooms and a self-contained school. *Education and Treatment of Children, 34*(1), 15-33. <https://www.doi.org/10.1353/etc.2011.0003>
- McIntosh, K., Horner, R. H., & Sugai, G. (2009). Sustainability of systems-level evidence-based practices in schools: Current knowledge and future directions. In Sailor, W., Dunlap, G., Sugai, G., & Horner, R (Eds.), *Handbook of Positive Behavior Support* (327-352). Springer.
- McIntosh, K., Mercer, S., Hume, A., Frank, J., Turri, M., & Mathews, S. (2013). Factors related to sustained implementation of schoolwide positive behavior support. *Exceptional Children, 79*(3), 293-311.
- McIntosh, K., Ty, S. V., & Miller, L. D. (2014). Effects of school-wide positive behavioral interventions and supports on internalizing problems: Current evidence and future

directions. *Journal of Positive Behavior Interventions*, 16(4), 209-218.

<https://www.doi.org/10.1177/1098300713491980>

Melin, G. L., & Götestam, K. G. (1973). A contingency management program on a drug-free unit for intravenous amphetamine addicts. *Journal of Behavior Therapy and Experimental Psychiatry*, 4(4), 331-337.

[https://www.doi.org/10.1016/0005-7916\(73\)90003-7](https://www.doi.org/10.1016/0005-7916(73)90003-7)

National Center for Educational Statistics. (2021, May). *Students with disabilities*.

<https://nces.ed.gov/programs/coe/indicator/cgg>

Oliver, R. M., & Reschly, D. J. (2010). Special education teacher preparation in classroom management: Implications for students with emotional and behavioral disorders. *Behavioral Disorders*, 35(3), 188-199.

<https://www.doi.org/10.1177/019874291003500301>

Partin, T. C. Moore, Robertson, R. E., Maggin, D. M., Oliver, R. M., & Wehby, J. H. (2009). Using teacher praise and opportunities to respond to promote appropriate student behavior. *Preventing School Failure*, 54(3), 172-178.

<https://www.doi.org/10.1080/10459880903493179>

Reddy, L. A., De Thomas, C. A., Newman, E., & Chun, V. (2009). School-based prevention and intervention programs for children with emotional disturbance: A review of treatment components and methodology. *Psychology in the Schools*, 46(2), 132-153.

<https://www.doi.org/10.1002/pits.20359>

Reid, R., Gonzalez, J. E., Nordness, P. D., Trout, A., & Epstein, M. H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of*

Special Education, 38(3), 130-143.

<https://www.doi.org/10.1177/00224669040380030101>

Scheuermann, B., Webber, J., Partin, M., & Knies, W. C. (1994). Level systems and the law: Are they compatible? *Behavioral Disorders*, 19(3), 205-220.

<https://www.doi.org/10.1177/019874299401900304>

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children*, 31(3), 351-380. <https://www.doi.org/10.1353/etc.0.0007>

Smith, S. W., & Farrell, D. T. (1993) Level system use in special education: Classroom intervention with prima facie appeal. *Behavioral Disorders*, 18(4), 251-264.

<https://www.doi.org/10.1177/019874299301800408>

Steinberg, Z., & Knitzer, J. (1992). Classrooms for emotionally and behaviorally disturbed students: Facing the challenge. *Behavioral Disorders*, 17(2), 145-156.

<https://www.doi.org/10.1177/019874299201700208>

Sutherland, K. S., & Singh, N. N. (2004). Learned helplessness and students with emotional or behavioral disorders: Deprivation in the classroom. *Behavioral Disorders*, 29(2), 169-181. <https://www.doi.org/10.1177/019874290402900208>

Sutherland, K. S., & Wehby, J. H. (2001). The effect of self-evaluation on teaching behavior in classrooms for students with emotional and behavioral disorders. *Journal of Special Education*, 35(3), 161-171. <https://www.doi.org/10.1177/002246690103500306>

Swinson, J., & Cording, M. (2002). Assertive discipline in a school for pupils with emotional and behavioural difficulties. *British Journal of Special Education*, 29(2), 72-75.

<https://www.doi.org/10.1111/1467-8527.00243>

- Wagner, M. M. (1995). Outcomes for youths with serious emotional disturbance in secondary school and early adulthood. *The Future of Children*, 5(2), 90-112.
<https://www.doi.org/10.2307/1602359>
- Walker, J. E., & Shea, T. M. (1995). *Behavior management: A practical approach for educators*. Merrill Pub Co.
- Weeden, M., Wills, H. P., Kottwitz, E., & Kamps, D. (2016). The effects of a class-wide behavior intervention for students with emotional and behavioral disorders. *Behavioral Disorders*, 42(1), 285-293. <https://www.doi.org/10.17988/BD-14-12.1>
- Wehby, J. H., Lane, K. L., & Falk, K. B. (2003). Academic instruction for students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders*, 11(4), 194-197. <https://www.doi.org/10.1177/10634266030110040101>
- Wells, P. L. (1990). *Boys Town Educational Model*. Boys Town National Training Center.

APPENDIX B

Institutional Review Board Approval Letter and Consent Form**Memorandum**

To: Cade Charlton

Department: BYU - EDUC - Counseling, Psychology, & Special Education

From: Sandee Aina, MPA, HRPP Associate Director

Wayne Larsen, MAcc, IRB Administrator

Date: November 02, 2021

IRB#: E2020-168

Title: Perspectives from the Classroom: A Qualitative Analysis of the Helping and Hindering Incidents Associated with Implementing Behavior Intervention Plans

Brigham Young University's IRB has reviewed the amendment submitted to change the study title, add study personnel, make a change in enrollment, modify the consent form to emphasize the broadened focus on classroom management (along with a modification to the consent form to reflect this change), and including the addition of a new survey that includes the level system (effective classroom management) along with the BIP survey. The IRB determined that the amendment does not increase risks to the research subject and the aims of the study remain as originally approved. The amendment has been approved. The revised consent statement and recruiting script have been approved and stamped for your files.

All conditions for a continued approval period remain in effect. Any modifications to the approved protocol must be submitted, reviewed, and approved by the IRB before modifications are incorporated in the study.

Consent to be a Research Subject

Title of the Research Study: Perspectives from the Classroom: A Qualitative Analysis of the Helping and Hindering Incidents Associated with Implementing Effective Classroom Management Strategies **Principal Investigator:** Dr. Cade Charlton
IRB ID#: E2020-168

Introduction

This research study is being conducted by Cade Charlton and Ellie Young at Brigham Young University to determine the helping and hindering events that influence the implementation of effective classroom management strategies including the implementation of Behavior Intervention Plans (BIPs). You were invited to participate because you have had recent success implementing an effective classroom management strategy.

Procedures

If you agree to participate in this research study, the following will occur:

- You will be asked to fill out a brief demographic survey (~ 5 minutes).
- You will be interviewed for approximately 30-75 minutes about implementing classroom management procedures.
 - The interview will be audio recorded to ensure accuracy in reporting your statements.
- The interview will take place at a time and place (e.g., your office, classroom, or on Zoom) that is convenient for you.
- The researcher will contact you later via email to clarify your interview answers. Your reply will likely take <15 minutes.

Risks/Discomforts

You may experience difficulties associated with allocating time to participate in this research, you may experience some emotional distress as you recall their experiences working with challenging students. You also may experience a loss of confidentiality by disclosing the details of a behavior intervention plan to the researchers.



IRB NUMBER: E2020-168

IRB APPROVAL DATE: 11/02/2021

1 | BYUHRP v09/2021

Benefits

There will be no direct benefit to you. You may gain insight into what will help you better implement classroom management strategies in the future.

Confidentiality

Your name will be replaced by a pseudonym in order to maintain confidentiality. Additionally, all data from this study will be saved on a secure, password protected, online file sharing website.

Compensation

You will be compensated with a \$25 amazon gift card or cash equivalent gift at the end of the interview.

Participation

Participation in this research study is voluntary. You have the right to withdraw at any time or refuse to participate entirely without jeopardy to your professional status or standing with the university.

Questions about the Research

If you have questions regarding this study, you may contact Cade Charlton at cade_charlton@byu.edu for further information.

Questions about Your Rights as Research Participants

If you have questions regarding your rights as a research participant contact Human Research Protections Program by phone at (801) 422-1461; or by email: BYU.HRPP@byu.edu.

Statement of Consent

I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study.

Name (Printed): Signature Date:



IRB NUMBER: E2020-168

IRB APPROVAL DATE: 11/02/2021

1 | BYUHRPv09/2021

APPENDIX C

Instruments**Interview Script***(Estimated duration = 30 minutes)*

Date _____
 Interviewer _____
 Participant # _____

“Thank you for taking the time to participate in this study. Your time is important to us so we will do our best to complete this interview in under 30 min. As you may know, we are conducting research on school-wide behavior management systems, in particular level systems, and how to sustain them. Your school has utilized a school-wide level system for over 20 years. Our purpose in conducting this study is to identify what has sustained your efforts in using the level system at your school. You were selected for this study because you are a part of the faculty and have contributed to the longevity of your school’s level system.

I’d like to record our conversation so your answers can be evaluated by our team.

First, I’d like to ask a few questions about you.”

1. “How long have you been teaching/years in your current position? and years in this field? role? (Structured follow-up)

2. “When you started in your current position, what training was provided on level systems?”

3. “Have you received any additional training on level systems?” If yes, “In what ways?” (e.g., district training, professional development, conference, university class)

4. At any time, have you contributed to creation or updates of any part of your school’s level system?

“Thank you. Next, I’d like to ask you several questions about the sustained longevity of the level system at your school. Please, reflect carefully on the specific events, observable behaviors, and examples that come to mind with each question. These events should be things that you, members of your team, or other professionals did.”

1. “What were the important *events* that **helped** you sustain the level system in your classroom? Please describe each incident in as much detail as possible.”

- Probe for further detail.

HELPFUL FACTORS

“Let me briefly summarize the helpful factors you mentioned. I noted [List Helping Factors].”

2. “What were the specific events that **hindered** your efforts to sustain the level system in your classroom?”

- Probe for further detail.

HINDERING FACTORS

“Let me briefly summarize the factors you mentioned that hindered sustainability. I noted [List Hindering Factors].”

3. “Looking back, are there other things that you wished had happened or could happen to improve the level system in your school?”

- Probe for further detail.

WISH LIST ITEMS

“Let me briefly summarize the items you mentioned that could have helped sustainability. I noted the following items [List Wish List Items]. “

“Thank you so much for your time. We will be conducting other interviews and then will work on analyzing these data. We will share our results with the district as soon as possible.”

Length of the interview (min:sec) __