BULLETIN OF EDUCATIONAL STUDIES
http://bulletineducation.com
https://prensip.gen.tr

RESEARCH ARTICLE

Beyond Standard Measures: Crafting an Inclusive Evaluation Tool for Special Education Teachers in Self-Contained Settings

Kareem Thompson[1] • Jennifer Lesh[2]✉

Lynn University, College of Education, Department of Exceptional Student Education, Boca Raton, FL/USA

ARTICLE INFO

Article History
Received: 05.12.2023
Accepted: 14.12.2023
First Published: 31.12.2023

Keywords
Educational transformation
Individualized instruction
Self-contained classroom
Special education evaluation
Special education teacher accountability

ABSTRACT

Self-contained special education teacher evaluations are investigated in this study. The perceptions of special education teachers and building administrators are explored and the suitability of standard teacher evaluation tools, such as the Marzano Teacher Evaluation System, for special education teachers in self-contained classrooms are discussed. This study highlights the significant gap in these tools' relevance and efficacy due to the unique challenges and individualized instruction required in special education settings. The study utilizes a quantitative survey method to gather perceptions from teachers and administrators, revealing widespread dissatisfaction with current evaluation processes and a lack of alignment with special education needs. Results indicate a pressing need for tailored evaluation frameworks that accurately reflect the pedagogical realities of special education. The article advocates for reforming evaluation systems to foster professional growth and improve educational outcomes for students with disabilities, thus transforming the approach to educator accountability and development in specialized teaching environments.

Please cite this paper as follows:

Thompson, K., & Lesh, J. (2023). Beyond standard measures: Crafting an inclusive evaluation tool for special education teachers in self-contained settings. *Bulletin of Educational Studies*, *2*(2), 65-75. https://doi.org/10.61326/bes.v2i2.125

1. Introduction

In the special education profession, where specially designed instruction is often scaffolded and individualized, the effectiveness of the special education teacher evaluation process becomes a vital function to the continual improvement of the special education teacher's pedagogy. This study focused on the usefulness of standardized teacher evaluation tools such as *the Marzano Teacher Evaluation System* as an evaluation instrument for implementation with special education teachers (SETs) working in self-contained classrooms. There have been concerns about the relevance of these instruments as they are not dovetailed to the unique teaching context in self-contained full-time special education classrooms and the administrative support and feedback provided to these self-contain SETs as well as their professional growth (Close et al., 2020; Aldosiry, 2022). Prompted by the reauthorization of acts such as *Federal Title I* and *Title II-A programs under the Elementary and Secondary Education Act* (ESEA, 1965), these evaluation procedures have been revamped, but they still leave gaps in the usefulness by administrators' feedback and support to SETs who teach in self-contained settings serving students with low-incidence disabilities (Billingsley & Bettini, 2019). These federal initiatives aim to ensure fair access to quality education for all students, including students with disabilities, while enhancing the rigor of educator evaluations. Notably, the landscape shifted with the reauthorization *of the No Child Left*

✉ Correspondence
E-mail address: jlesh@lynn.edu

Behind Act (NCLB, 2001) into *the Every Student Succeeds Act* (ESSA) in 2015, granting states more autonomy in shaping their academic standards, assessments, and accountability systems under Title I. Title II-A further allocates funding for initiatives to bolster educator effectiveness and equitable evaluation (Troppe et al., 2017).

However, these evaluation and accountability measures have not been without their challenges (Close et al., 2020). The Marzano Teacher Evaluation System is one such teacher evaluation tool that several school districts have embraced as a response to these state mandates (Walkinshaw, 2022). This observation assessment system was developed with the aim of providing continual pedagogical improvement for the teachers, direction for administrators to assist teachers, and professional developers regarding the evaluation process (Kellett, 2006). Hence, it is intended to provide the supervisors with a platform for the other processes in the classroom regarding content pedagogy and classroom management improvement. Evolving into the Marzano Focused Teacher Evaluation Model precisely aligned with state evaluation measures (Marzano, 2017), it was designed in partnership between Robert J. Marzano and Learning Sciences International to address emerging needs aligned with standards-based instruction (Carbaugh et al., 2017).

Teacher accountability is one of the main underpinnings of ESSA, which requires all teachers to be evaluated. Special education teachers who teach in self-contained settings are included in this accountability requirement. The evaluation of special education teachers who teach in self-contained settings should be considered unique due to the demands and characteristics of these settings. Even with these apparent demands, no current separate evaluation tool is being used to evaluate special education teachers who teach in self-contained settings (Snyder & Pufpaff, 2021; Jones et al., 2022). Glowacki and Hackmann (2016) argue that state statutes may need more guidance for evaluating special education teachers. Self-contained special education teachers need valuable feedback and specific support to grow as professional educators. Exceptional education teacher pedagogy growth support and feedback should be reinforced through a tailored evaluation process (Billingsley & Bettini, 2019). How can self-contained special education teacher pedagogy be supported by administrators and purposeful feedback given to those teachers using an evaluation tool that has been designed for general and inclusive education settings?

This study addresses this question by investigating school principals' and special education teachers' perceptions of the efficacy of current teacher evaluation processes for self-contained special education teachers. Furthermore, the study sought to identify the components necessary for a specialized evaluation tool designed to cater to the distinctive requirements of these self-contained special education teachers. Beyond the realm of professional growth, the study recognizes the broader implications-ensuring accountability and effective evaluation measures for students with profound cognitive disabilities within this unique educational context.

1.1. Conceptual Framework

The Marzano model is a scientific-behavioral evaluation system. In their white paper, *The Marzano Focused Teacher Evaluation Model, Toth and Marzano* (2017) explain the focused evaluation model as agnostic in that it recognizes effective instruction with student evidence as the critical factor. The Marzano Teacher Evaluation Model is coined as a valid, reliable, defensible model that is based on four domains: (1) standards-based planning, (2) standards-based instruction, (3) conditions for learning, and (4) professional responsibilities. These domains encompass twenty-two elements.

The following is a summary description of the procedures that occur in the evaluation of teachers:

Implementation: Observation Types

1**.** Formal Observation type is administered for about thirty to fifty-five minutes and is evaluative. It can also be up to when the assessment or the lesson is completed. There will be a notification for each of the employees each week of the observations made formally. Notification will be made earlier at the most convenient time and not later than the ultimate workday of the previous week. Also, there will be a Pre-Conference that must be done in face-to-face meetings. There will be a post-conference that will be done face-to-face. This post-conference will provide feedback and coaching from the administrator. The observations will be scheduled based on the teacher's specific date and time. There Written feedback will be sent to the teacher. Typically, the formal evaluation the final evaluation and is conducted after the informal observation.

2. Informal Observation type is administered for twenty to forty minutes. The informal observation uses the same tool as the formal. This type may also be announced or unannounced to the teacher. This observation is not encouraged to occur immediately before or after Thanksgiving, winter, and spring breaks. Actionable and written feedback must be provided to the teacher.

It is recommended that administrators conduct walkthroughs during the earlier fifteen teacher working days. They are non-evaluative observations. Hard copy feedback will be issued in the form of an electronic model. It is expected that the Category 1A teachers, who are new to the school district, will be provided with four observations for the entire year and two other observations. Below is Figure 1, a visual presentation of a typical school district's Observation Implementation Requirements Reference Sheet.

Implementation Requirements	
Formal	• 30-55 minutes (or until the completion of a lesson). • Must be scheduled between the teacher and their observer for a specific date and time. • Pre-Conference must be a face-to-face meeting. • Post-Conference must be a face-to-face meeting. • Actionable written feedback must be provided through the observation instrument in iObservation.
Informal	• 20-40 minutes. • May be announced or unannounced. • Informal observations taking place immediately before or after the Thanksgiving, winter and spring breaks are strongly discouraged. • Actionable written feedback must be provided through the observation instrument in iObservation.
Walkthroughs	• Shall not be used for evaluative purposes.
Data Marks	• All 22 elements are NOT required to be observed or scored in one academic year. • Only dominant elements should be coded during an observation. • At least one (1) Element must be coded from each of the 4 Domains during the evaluation period. • No elements are to be scored during the first 3 weeks of employment.

Figure 1. Observation implementation requirements reference sheet.

1.2. Additional Relevant Literature

The complexity of self-contained settings, designed to cater to students with disabilities, relies on specially designed individualized instruction aligned with Individualized Education Plans (IEPs) (CEC, 2015; Petersen, 2016).

Special education teachers operating within these contexts face unique challenges, often unaddressed by evaluation tools tailored for general education teachers. Research literature has explored teacher evaluation systems extensively, but studies focusing on evaluating special education teachers within self-contained settings still need to be completed (Holdheide et al., 2010; Jones et al., 2022).

Emerging evaluation models in the United States must account for the distinctive challenges of evaluating special educators, including English Language Learner (ELL) specialists. Researchers (Holdheide et al., 2010; Jones et al., 2022) note the need for more literature evaluating special educators. Many states lack a distinct evaluation tool for special education teachers in self-contained settings (Billingsley & Bettini, 2019). This gap, combined with the limited research on specialized evaluation tools, underscores the importance of this study.

1.3. Specially Designed Instruction in Self-Contained Special Education Classrooms

The delivery of specially designed instruction is the core job responsibility of special education teachers. IDEA regulations define "specially designed instruction" as "adapting, as appropriate to the needs of an eligible child under this part, the content, methodology or delivery of instruction (i) to address the unique needs of the child that result from the child's disability; and (ii) ensure access of the child to the general curriculum, so that the child can meet the educational standards within the jurisdiction of the public agency that apply to all children." (34 CFR Sec. 300.39(b)(3)). Billingsley et al. (2020) explain that practical terms, "Specially Designed instruction is what makes special education 'special'."

Settings as such are unique and require instruction as dictated by the individual education plans of students that make up these settings. The special education teacher is a provider of services within the self-contained setting. Instruction is standards-based, well-planned, organized, and meaningful if delivered in an explicit and systematic way.

Teacher advocate Keely Swartzer (2017) proposes thoughts from three teachers regarding suitable characteristics of specially designed instruction in self-contained classes. She states that a teacher of a self-contained setting should (1) develop clear and consistent routines and procedures. These structured methods can decrease behaviors and minimize the students' cognitive load, allowing for more space/brain power to learn content. (2) Take time to reflect on everything your students do from when they arrive until they leave. Develop routines for each step of the day. (3) Train paraprofessionals on the importance of supporting the routines in the classroom. (4) Be Flexible- Every single child is an individual. There must be a willingness to look at EACH child's strengths, needs, and interests when we develop programming to help students succeed! Moreover, lastly, (4) Use a structured work system, as these systems teach the student to work independently, and independence.

2. Method

This study employed a quantitative and qualitative approach through a cross-sectional survey research design (Creswell & Creswell, 2018). This design enabled the researchers to collect perceptions of self-contained special education teachers and the administrators who evaluate these special education professionals using the Marzano evaluation tool. The sample was drawn from state certified special education professionals and administrators from a certification department of education in the southeastern United States. The survey was created using the Council for Exceptional Children's *Professional Standards Advanced Specialty Set Standards for Special Education Developmental Disabilities and Autism Spectrum Disorder Specialist* (CEC, 2015). The researchers underwent through a rigorous vetting and institutional review board process prior to collecting data.

2.1. Work Group

The survey was sent to 962 participants who fit the inclusion criteria: 479 administrators and 483 self-contained special education teachers. One hundred and seven participants started the survey but still needed to complete it. Participants were kept in the data set if they completed 75% or more of the questions; those participants were given the middle answer (neither agree nor disagree). Four participants fit the missing data category.

Seventy-nine participants completed the survey; the breakdown of their position and by level are reported below. While analyzing the results by teachers, the researcher removed data responses of thirty teachers who identified themselves as general education teachers and teachers who identified as others. The researcher assumed that these participants needed to fit the original inclusion criteria. Out of the seventy-nine participants, as shown in Figure 2, 36 were special education teachers, and 43 were administrators.

Figure 3 provides percentages of the special education teachers. The number of special education teachers that participated by school level is reported below. Twenty-four special education teachers were elementary, five special education teachers were K-12, five were secondary, and two special education teachers were identified as others. All special education teachers worked in self-contained "varying exceptionalities" (VE) classrooms. These VE classrooms have students with autism, or emotional behavioral disorders, or intellectual disabilities, or language impairment or any combination of two or more of these disabilities.

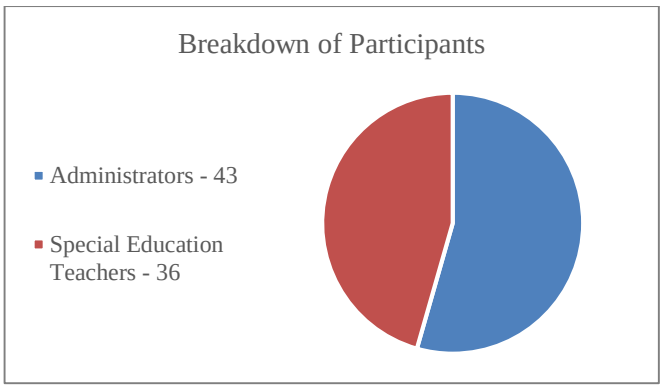

Figure 2. Breakdown of participants.

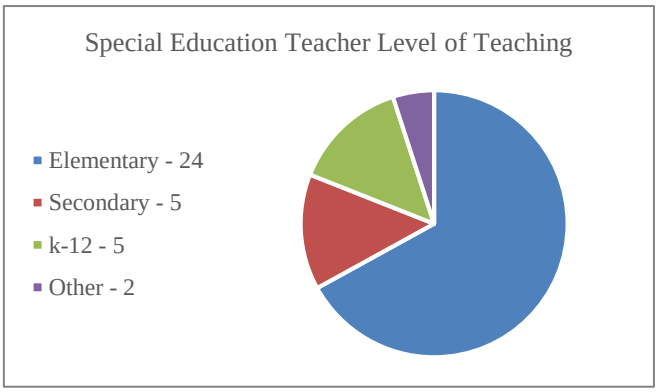

Figure 3. Breakdown of special education teachers level of teaching.

Figure 4 depicts the demographics of administrators who participated, including their work level. Twenty elementary administrators, seventeen secondary administrators, three K-12 administrators, and three others (such as "district administrators") completed the survey. The others included a district specialist, a pre-k administrator, and a pre-K-8 administrator.

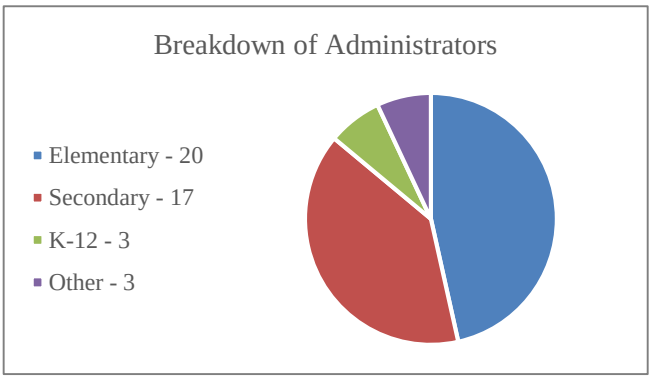

Figure 4. Breakdown of administrators.

2.2. Recoding of Variables

The variables "strongly agree" and "agree" were recoded. The variables "disagree" and "strongly disagree" were also recoded as one variable. Justification for this recoding allowed

the opportunity to measure a more distinct difference between the variables.

2.3. Analysis of Survey Items

To analyze survey questions 2, 3, 4, 5, 6, 7, and 8, the researcher ran descriptive statistics and a Chi-Square of Independence using Statistical Package for Social Sciences 28.0 (SPSS), a statistical analysis software. The survey data was gathered using Survey Monkey. The researcher imported results from Survey Monkey into SPSS. As mentioned above, the study's research participants were seventy-nine professionals, including thirty-six special education teachers and forty-three administrators.

3. Results

Survey Question #2 - *My current evaluation tool/system is aligned to the following teacher evaluation framework: Marzano Framework, Danielson Framework, District Created Framework, or Another Framework.*

This question allowed participants to select which framework to which their teacher evaluation is aligned. Participants' results are shown below in Figure 5. Frequency command produced normal distribution and no skewness. Fifty-two percent of the participants identified Marzano as their evaluation-aligned framework. Twenty-nine percent of the participants identified Danielson as their aligned evaluation-associated framework. Eighteen percent of the remaining were identified as having a district-created framework, and the other 1% chose "another".

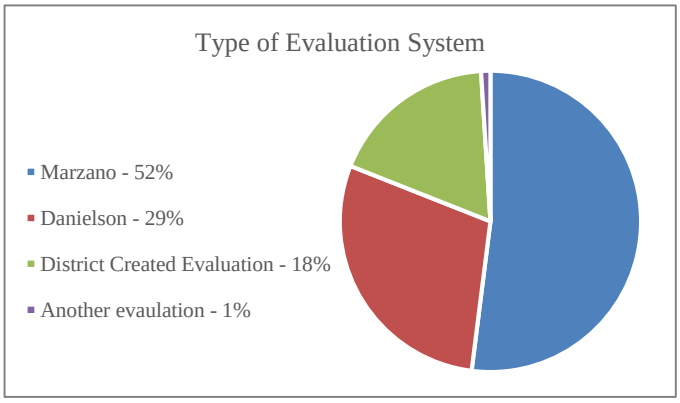

Figure 5. Participants identified an aligned framework.

Survey Question #3 - *My teacher evaluation tool/system adequately includes evidence-based practices for teaching students with low-incidence disabilities such as Intellectual Disabilities and Autism Spectrum Disorder*.

As shown below in Figure 6, participants' responses were 34% agreed that their identified teacher evaluation tool/system adequately includes evidence-based practices for teaching students with low-incidence disabilities such as Intellectual Disabilities and Autism Spectrum Disorder. Fifty-two percent disagreed that their identified teacher evaluation tool/system adequately includes evidence-based practices for teaching students with low-incidence disabilities such as Intellectual Disabilities and Autism Spectrum Disorder. The remaining fourteen percent neither agreed nor agreed. Frequency command produced normal distribution and no skewness.

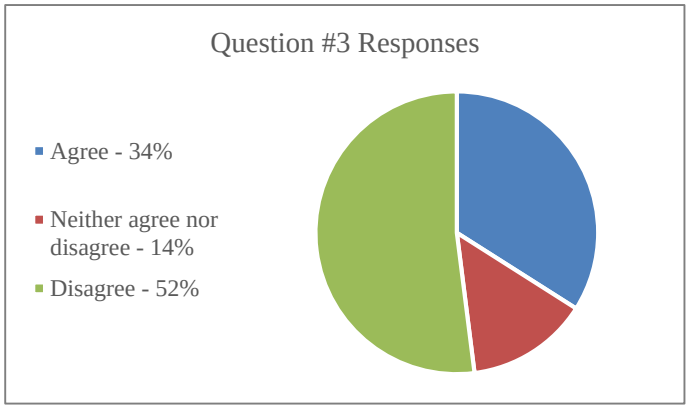

Figure 6. Question #3 responses.

Survey Question #4 - *My teacher evaluation tool/system adequately measures a self-contained special education teacher's effectiveness within the various roles and responsibilities as a teacher of students with low-incidence disabilities such as Intellectual Disabilities and Autism Spectrum Disorder.*

This question allowed participants to share their responses to the evaluation tool/system, adequately measuring their effectiveness. Participants' results are shown below in Figure 7. Sixty-seven percent of the participants disagreed that their evaluation tool/system adequately measured their effectiveness, and twenty-three percent agreed. The remaining ten percent neither agreed nor disagreed. Frequency command produced normal distribution and no skewness.

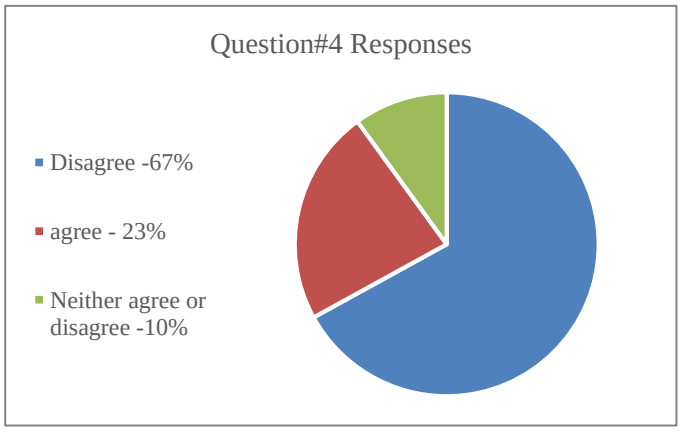

Figure 7. Responses for Question #4.

Survey Question #5 - *I have received adequate training on using my current evaluation tool/system to evaluate teachers of students with low-incidence disabilities. (If you are a teacher,*

have you received adequate training on interpreting your evaluation).

Survey Question number five allowed participants to share their responses on whether they had received adequate training related to their framework. Participants' results are shown below in Figure 8. Forty percent of the participants agreed that they had received adequate training. Forty-two percent of the participants disagreed. The remaining eighteen percent neither agreed nor disagreed.

Frequency command produced normal distribution and no skewness.

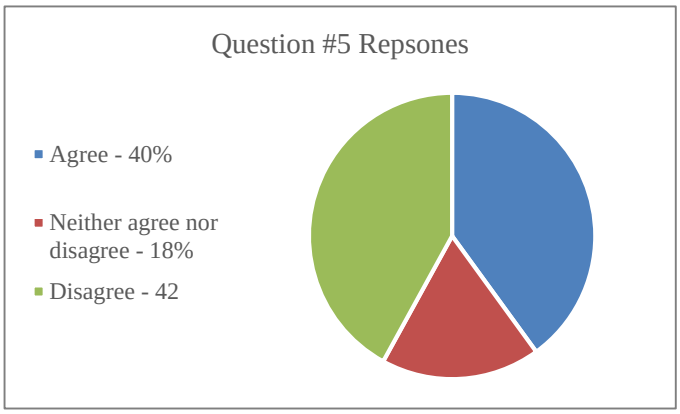

Figure 8. Responses for Question #5.

Survey Question #6 - *My teacher evaluation tool/system effectively and fairly uses students' growth data as part of the teacher evaluation system for teachers of students with low-incidence disabilities*.

Question number six allowed participants to respond to whether student growth data is used effectively in their identified evaluation system. Approximately Forty-seven percent of the participants disagreed as to whether student growth data is used in their identified evaluation system effectively, while thirty-six percent agreed. The remaining seventeen percent neither agreed nor disagreed. Results are displayed in Figure 9.

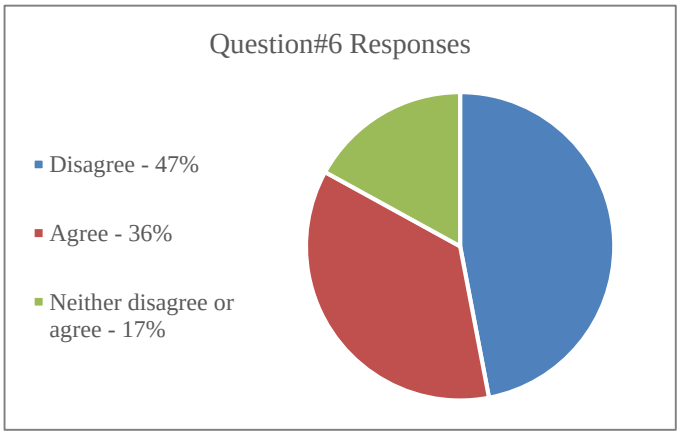

Figure 9. Responses for Question #6.

Survey Question #7 - My current evaluation tool/system effectively ties the teacher evaluation process to professional development opportunities for teachers of students with low-incidence disabilities.

Participants' responses to Question Seven related to whether their evaluation tool/system effectively ties to professional development opportunities for teachers. As shown in Figure 10, Fifty-eight percent disagree with the question, while 24% agree that their current evaluation effectively ties the evaluation process to professional development for teachers. Frequency command produced normal distribution and no skewness.

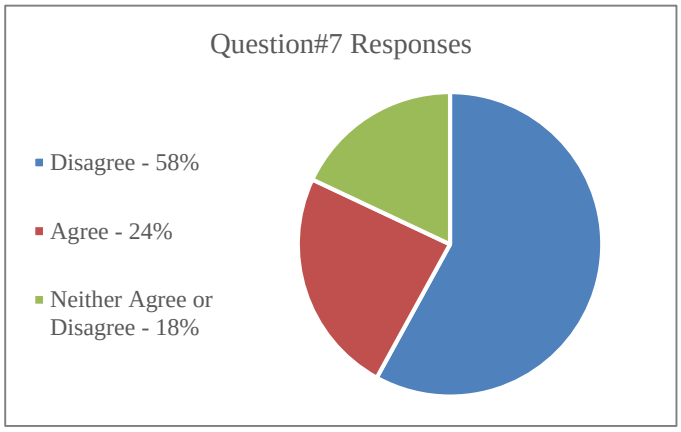

Figure 10. Responses for Question #7.

Survey Question #8 - "My current evaluation tool/system adequately provides feedback to teachers of students with low-incidence disabilities about their teaching practices to assist them in determining areas for improvement and developing their skills."

Below, Figure 11 shows results provided by participants to question 8 related to feedback to teachers to determine areas for improvement and developing their skills.

Forty-nine percent of the participants disagreed with the question, while 29% agreed. The remaining twenty-two percent neither agreed nor disagreed. Frequency command produced normal distribution and no skewness.

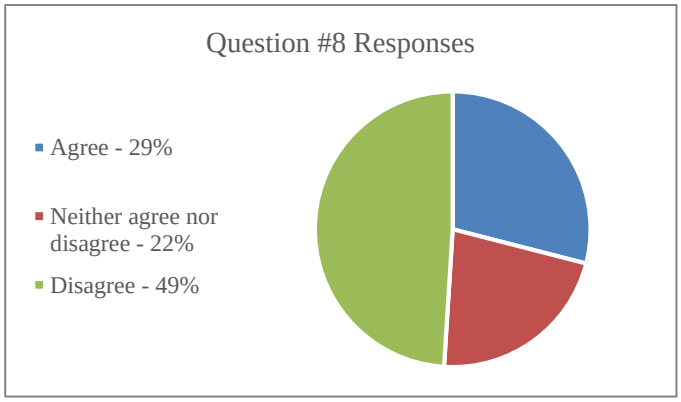

Figure 11. Responses for Question #8.

3.1. Chi-Square Analysis

A chi-square test of independence was calculated comparing the frequency of participants aligned teacher evaluation tool/system, as identified in survey question number 2: "*My current evaluation tool/system is aligned to the following teacher evaluation framework: Marzano Framework, Danielson Framework, District Created Framework, or Another Framework*" was individually compared to questions 3, 4, 5, 6, 7 and 8.

A chi-square test of independence was run between question number two: "*My current evaluation tool/system is aligned to the following teacher evaluation framework: Marzano Framework, Danielson Framework, District Created Framework, or Another Framework*" and question three: "*My teacher evaluation tool/system adequately includes evidence-based practices for teaching students with low incidence disabilities such as Intellectual Disabilities and Autism Spectrum Disorder*". A significant interaction was found (c2(1)=1.086, *p*>.000). Over half of the participants (56%) regardless of their aligned framework (25.3% Marzano; 18.8% Danielson; 10.23% district created; 1.26% Another framework) disagreed with their evaluation tool adequately including evidence-based practices for teaching students with low incidence disabilities.

A chi-square test of independence was run between question number two: "*My current evaluation tool/system is aligned to the following teacher evaluation framework: Marzano Framework, Danielson Framework, District Created Framework, or Another Framework*" and question number four: "*My teacher evaluation tool/system adequately measures a self-contained special education teacher's effectiveness within the various roles and responsibilities as a teacher of students with low-incidence disabilities such as Intellectual Disabilities and Autism Spectrum Disorder*". A significant interaction was found (c2(1)=.800, *p*>.000). Fifty-nine percent of the participants regardless of their aligned framework (26.58% Marzano; 20.25% Danielson; 11.39% district created; 1.26% Another framework) disagreed with their evaluation tool adequately measuring a self-contained special education teacher's effectiveness within the various roles and responsibilities as a teacher of students with low-incidence disabilities.

A chi-square test of independence was run between question number two: "*My current evaluation tool/system is aligned to the following teacher evaluation framework: Marzano Framework, Danielson Framework, District Created Framework, or Another Framework*" and question number five: "*I have received adequate training on how to use my current evaluation tool/system to evaluate teachers of students with low-incidence disabilities (If you are a teacher, have you received adequate training on how to interpret your evaluation)*". A significant interaction was found (c2(1)=1.076, *p*>.000). Fifty-nine percent of the participants, regardless of their aligned framework (24% Marzano; 10.13% Danielson; 5.06% district created; 0% Another framework), disagreed with their evaluation tool being aligned to adequate training on how to use and interpret their current evaluation tool/system.

A chi-square test of independence was run between question number two: "*My current evaluation tool/system is aligned to the following teacher evaluation framework: Marzano Framework, Danielson Framework, District Created Framework, or Another Framework*" and question number six: "*My teacher evaluation tool/system effectively and fairly uses students' growth data as part of the teacher evaluation system for teachers of students with low-incidence disabilities*". A significant interaction was found (c2(1)=1.076, *p*>.000). Forty-seven percent of the participants, regardless of their aligned framework (27% Marzano; 10.13% Danielson; 8.86% district created; 1% Another framework), disagreed with their evaluation tool effectively and somewhat using students' growth data as part of the teacher evaluation system for teachers.

A chi-square test of independence was run between question number two: "*My current evaluation tool/system is aligned to the following teacher evaluation framework: Marzano Framework, Danielson Framework, District Created Framework, or Another Framework*" and question number seven: "*My current evaluation tool/system effectively ties the teacher evaluation process to professional development opportunity for teachers of students with low-incidence disabilities*". No statistically significant interaction was found (c2(1)=.358, *p*>.865). The participants' aligned teacher evaluation tool/system and whether it effectively tied the teacher evaluation process to professional development opportunities for teachers of students with low-incidence disabilities appeared to be independent variables.

A chi-square test of independence was run between question number two: "*My current evaluation tool/system is aligned to the following teacher evaluation framework: Marzano Framework, Danielson Framework, District Created Framework, or Another Framework*" and question number eight: "*My current evaluation tool/system adequately provides feedback to teachers of students with low-incidence disabilities about their teaching practices to assist them in determining areas for improvement and developing their skills*". A significant interaction was found (c2(1)=1.078, *p*>.000). Forty-eight percent of the participants regardless of their aligned framework (25.31% Marzano; 10.12% Danielson; 11.39% district created; 1.26% Another framework) disagreed with whether their evaluation tool adequately provides feedback to teachers of students with low-incidence disabilities about their teaching practices to assist them in determining areas for improvement and developing their skills.

3.2. Open-Ended Question Analysis

"What recommendations/suggestions, if any, you might change, add, or delete to your current evaluation tool/system to meet the needs of teachers of students with low-incidence disabilities? Please feel free to write as much as you like".

The researcher organized the open-ended responses, coded them into groups, and eventually generated themes. In analyzing the responses of participants for question number nine, a few overarching themes were identified. The researcher identified the following themes: usage of an inappropriate tool, non-support of teacher growth mindset, and the importance of professional development. These themes were evident in the responses of both special education teachers and administrators.

3.2.1. Inappropriate tool

Self-Contained Special Education Teacher:

"It (evaluation tool) must include a growth element for all grade levels. Some grade levels are judged on proficiency, and others graded on growth are not equitable.

Student IEP goals should be a part of the evaluation tool. Students who are reaching their goals and students who are not should impact the teacher's evaluation".

Administrator:

"These evaluations should take into consideration how these classrooms and teacher/student interactions may look differently. The types of questions or strategies the teacher uses may be different than just a general education setting. For example, a teacher may ask what may appear to be a lower-level question, but to a student with special needs, this question really may be higher-order to them or maybe needed to lead them to higher-order thinking. Student growth in these classrooms should also be taken into consideration when connecting their performance to teacher evaluations".

3.2.2. Individualized growth mindset

Special Education Teacher:

"Evaluators need to understand that instruction can and often should look different for students with low-incidence disabilities. If students are being instructed on a modified curriculum, evaluators need to understand access points and how that should look for various groups in a classroom setting. I spoke with a principal today that did not realize the students had any kind of standards they were learning".

Self-contained Special Education Teacher:

"We need adequate training on preparing lessons, task boxes, etc., that appropriately engage all students (even students that are working at a participatory level)".

Self-contained Special Education Teacher:

"Engagement and growth should be an expectation for every child as measured by an appropriate tool that is sensitive enough to document that growth".

Administrator:

"(It should include) more constructive criticism; more informal evaluations to allow room for improvement and more explicit ways on how to meet the evaluation criteria".

3.2.3. Professional development

Self-contained Special Education Teacher:

"People in the district and schools should be on the same page.

Some of those involved with these children are not trained sufficiently to evaluate or even help us".

Administrator:

"Be able to give specific feedback to full-time ESE teachers so they can perfect their craft".

3.3. Summary of Research Questions

Research Question # 1 - How are self-contained special education teachers in self-contained classes currently being evaluated? This question is addressed by the responses of the survey data derived from survey question number two, which asks participants to answer which framework their current evaluation tool/system is aligned to. The survey results yielded that 52.56% and 41 participants state the Marzano framework, 28.3%, and 22 participants state the Danielson framework, 17.9%, and 14 participants state a district created framework and 1.28% and 1 participant states that another framework is being used.

Research Question # 2 - What are the distinctive characteristics experienced by special education teachers in self-contained settings that create a need for a unique self-contained special education teacher evaluation tool? This question is addressed by the responses of the survey derived from survey question number nine, which allowed participants to provide recommendations/suggestions if any, that they might change, add, delete to their current evaluation tool/system to meet the needs of teachers of students with low-incidence disabilities. Responses from teacher participants include the following:

3.3.1. Special education teacher responses

"The evaluation system needs to take these disabilities growth into account, realizing that the growth may be different for every child".

"We need a completely different evaluation tool that doesn't penalize us if we don't engage in "higher-order questioning." Most ID (Intellectual Disabled), not all my students are not

capable of answering higher-order questions. This is just one example of how the Danielson rubric doesn't work for the teachers of low- incidence".

"To evaluate a teacher on the data of students with learning disabilities, it should focus more on the individual student's growth rather than a comparison to the district benchmark or expectation for all students. The bar should be attainable".

"We use a Value-added model that relies heavily on state standardized testing, which many of my students have trouble showing mastery or growth. There need to be multiple measures for my students to show mastery".

"People in the district and schools should be on the same page. Some of those involved with these children are not trained sufficiently to evaluate or even help us".

"More constructive criticism. More informal evaluations to allow room for improvement. More explicit ways on how to meet the evaluation criteria".

"Another tool to measure learning gains, other than a grade-level assessment. If a child is 3-4 grade levels behind, they could still make a year's growth, but still fail the grade-level assessment, showing no growth".

"I believe that these evaluation tools/systems do not work. I have witnessed not so great teachers getting HE (Highly Effective) solely because they are the principal favs and teachers that go above and beyond not receiving the credit that they get. Therefore, teachers are worn out, frustrated, and seeking other schools or professions. Why can't we just be the professionals that we are and have the trust of our district to do what we are EDUCATED to do every day. Instead of evaluating us and tearing us apart with false opinions, based on if your principals like you or not".

"The current evaluation tool/system should be revised to meet the needs of IND teachers".

3.3.2. Administrator responses

"Make the rubric more suitable for what is actually being evaluated".

"System needs to account for students with significant disabilities and have adjustments to the rubric for evaluating those teachers (i.e., high order questioning, student discussions, etc.)".

"The current system to become highly effective has a lot of verbiage that states student-led when a lot of the Special needs students need support and guidance and are not able to generate independent work".

"For the evaluation tool to be more of a professional growth tool instead of punitive".

"The Marzano Evaluation tool is too generic to be used with students who are outside of the general education classroom effectively. It relies on the evaluator to determine what the "desired effect" is for the higher-order elements of the tool. It also does not consider the difficulty in organizing these students to "interact." It is a very unfair tool to use and requires the observer to be very subjective in determining if the teachers meet the element".

"Be able to give specific feedback to full-time ESE teachers".

Research Question # 3 - Is training necessary to implement a unique self-contained special education teacher evaluation tool? Question number five of the survey completed by participants allowed them to answer the question if as an administrator if they might have received adequate training on how to use their current evaluation tool/system to evaluate teachers of students with low-incidence disabilities. If they are a teacher, have they received adequate training on how to interpret their evaluation. The survey results yielded that 44.87% agreed and 35 participants agreed, 15.38% and 12 participants neither agreed nor disagreed, 39.74% and 31 participants disagreed. A statistically significant interaction was found ($c2(1)=1.076$, $p>.000$) using a chi-square test of independence. The participants' results reveal that adequate training is necessary in order to implement any evaluation tool regardless of its framework.

Research Question # 4 - How should a unique self-contained special education teacher evaluation tool be implemented? Research question number four is also addressed by some of the responses derived from survey question number nine, which allowed participants to provide recommendations/suggestions if any, that they might change, add, delete to their current evaluation tool/system to meet the needs of teachers of students with low-incidence disabilities. The following responses specifically address this research question.

3.3.3. Participant responses

"I don't believe the evaluation system should be used as a way to determine a teacher's worth. It should be a tool to reinforce teaching strategies and increase student learning. Also, when administrators see that a teacher is not working to the standards, they need a strong leadership team to support them. The overall Marzano system is informative and informational, but it is subjective to the user".

"Portfolios- baseline data from multiple sources- quantitative and qualitative data- surveys- interviews- multiple ways to evaluate".

"It needs to include a growth element for all grade levels. Some grade levels judged of proficiency and others graded on growth are not equitable. Student IEP goals should be a part of the evaluation tool. Students who are reaching their goals and students who are not should impact the teacher's evaluation."

"The ability to identify the class with specificity".

"Have additional categories, or elements, that are tailored to more closely align with what effective instruction looks like in these classrooms".

"Evaluators need to understand that instruction can and often should look different for students with low-incidence disabilities. If students are being instructed on a modified curriculum, evaluators need to understand access points and how that should look for various groups in a classroom setting. I spoke with a principal today that did not realize the students had any kind of standards they were learning. Teachers need adequate training on preparing lessons, task boxes, etc. that appropriately engage all students (even students that are working at a participatory level). Engagement and growth should be an expectation for every child as measured by an appropriate tool that is sensitive enough to document that growth".

"These evaluations should take into consideration how these classrooms and teacher/student interactions may look differently. The types of questions or strategies the teacher uses may be different than just a general education setting. For example, a teacher may ask what may appear to be a lower level question, but to a student with special needs, this question really may be higher order to them or may be needed to lead them to higher-order thinking. Student growth in these classrooms should also be taken into consideration when connecting their performance to teacher evaluations".

4. Discussion

The survey results indicate significant gaps in the existing teacher evaluation systems used for special education teachers. There is clearly a demand for an inclusive evaluation tool. The majority of participants, including both special education teachers and administrators, expressed dissatisfaction with current evaluation tools, citing a lack of alignment with the unique needs and teaching environments encountered in self-contained classrooms for students with low-incidence disabilities. The recoding of variables in the survey, combining "strongly agree" with "agree" and "disagree" with "strongly disagree", revealed an apparent dichotomy in perceptions regarding the adequacy of these evaluation tools. This dichotomy points to an underlying issue of these tools needing to be tailored for the specific challenges and objectives of these low-incidence special education settings.

4.1. Implications for Policy and Practice

4.1.1. Evaluation frameworks

Over half of the survey participants indicated that the Marzano Framework was their primary evaluation tool, yet the majority disagreed that this framework adequately includes evidence-based practices for low-incidence disabilities. This mismatch suggests a need for re-evaluating and possibly redesigning existing frameworks to be more inclusive of special education contexts. The researchers suggest using the framework of the *Marzano or Danielson* evaluation tools using *the CEC Specialty Set of Standards for Developmental Disability and Autism Spectrum Disorder Specialist* (CEC, 2015) to design a specific evaluation tool for low-incidence special education teachers that is sensitive enough to give feedback to these professionals.

4.1.2. Professional development

The survey responses highlight a significant need for professional development tailored to evaluators and teachers in low-incidence special education settings. Training should focus on understanding the unique challenges these low-incidence special education teachers face and how to assess their effectiveness accurately.

4.1.3. Feedback and growth

The feedback mechanism within the current evaluation systems seems inadequate for these special education professionals. There is a need for a more constructive approach that focuses on growth, especially in environments where student progress may not align with conventional academic benchmarks.

4.2. Future Research

Evaluation tool development is paramount for the growth and retention of low-incidence special education teachers. Research should focus on developing and testing new evaluation tools specifically designed for special education contexts, such as the draft evaluation tool the researchers suggest. These tools should incorporate aspects of student IEP goals, classroom interaction dynamics, and teacher effectiveness in teaching students with low-incidence disabilities.

4.3. Longitudinal Studies

Future studies could explore the long-term impacts of implementing specialized evaluation tools such as the tool the researchers suggest on teacher performance and student outcomes in special education settings. Comparing the effectiveness of different evaluation frameworks (like Marzano and Danielson) using *the CEC Specialty Set of Standards for Developmental Disability and Autism Spectrum Disorder Specialist* (CEC, 2015) in special education contexts would provide deeper insights into their suitability and areas for improvement.

5. Conclusion

The survey reveals a critical need for rethinking and reforming the teacher evaluation process for special education full-time classrooms. The current systems largely fail to address

the unique challenges and objectives of teaching students with low-incidence disabilities (Jones et al., 2022). There is a pressing need for frameworks that not only include special education settings but also focus on fostering growth and providing meaningful feedback for these special education teachers. There very underpinning of special education is inclusion and here the field has fallen short. Transformations to the evaluation tool are required to for genuine inclusivity to be realized. Implementing these changes will require a collaborative effort involving special educators, special education directors, administrators, and policy-makers, guided by ongoing research and feedback from those directly involved in special education. This effort will improve the evaluation practice and enrich the overall quality of education provided to students with special needs.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Aldosiry, N. (2022). The Influence of support from administrators and other work conditions on special education teachers. *International Journal of Disability, Development and Education, 69*(6), 1873-1887. https://doi.org/10.1080/1034912X.2020.1837353

Billingsley, B., & Bettini, E. (2019). Special education teacher attrition and retention: A review of the literature. *Review of Educational Research, 89*(5), 697-744. https://doi.org/10.3102/0034654319862495

Billingsley, B., Bettini, E., Mathews, H. M., & McLeskey, J. (2020). Improving working conditions to support special educators' effectiveness: A call for leadership. *Teacher Education and Special Education, 43*(1), 7-27. https://doi.org/10.1177/0888406419880353

Carbaugh, B., Marzano, R., & Tooth, M. (2017). *The Marzano focused teacher evaluation model. A learning sciences Marzano centre monograph*. Learning Sciences International.

CEC. (2015). *Advanced specialty set: Special education developmental disabilities and autism spectrum disorder specialist*. Council for Exceptional Children (CEC). https://exceptionalchildren.org/sites/default/files/2020-07/Advanced%20Specialty%20Set%20%20Special%20Education%20Developmental%20Disabilities%20and%20Autism%20Spectrum%20Disorder%20Specialist.pdf

Close, K., Amrein-Beardsley, A., & Collins, C. (2020). Putting teacher evaluation systems on the map: An overview of state's teacher evaluation systems post–Every Student Succeeds Act. *Education Policy Analysis Archives, 28*(58), 1-31. https://doi.org/10.14507/epaa.28.5252

Creswell, J.W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches. (5th ed.)*. Sage.

ESEA (Elementary and Secondary Education Act). (1965). *Public Law 89-10, 79 Stat. 27*. https://www.govinfo.gov/content/pkg/COMPS-748/pdf/COMPS-748.pdf

Glowacki, H., & Hackmann, D. G. (2016). The effectiveness of special education teacher evaluation processes: Perspectives from elementary principals. *Planning and Changing, 47*(3/4), 191-209.

Holdheide, L. R., Goe, L., Croft, A., & Reschly, D. J. (2010). *Challenges in evaluating special education teachers and English language learner specialists*. National Comprehensive Center for Teacher Quality. https://files.eric.ed.gov/fulltext/ED520726.pdf

Jones, N. D., Bell, C. A., Brownell, M., Qi, Y., Peyton, D., Pua, D., Fowler, M., & Holtzman, S. (2022). Using classroom observations in the evaluation of special education teachers. *Educational Evaluation and Policy Analysis, 44*(3), 429-457. https://doi.org/10.3102/01623737211068523

Kellett, S. (2006). *A picture of creative entrepreneurship: Visual narrative in creative enterprise education report*. Internationalizing Entrepreneurship Education and Training Conference. Oestrich-Winkel.

Marzano, R. J. (2017). *The new art and science of teaching*. Solution Tree Press.

NCLB (No Child Left Behind Act). (2001). *Pub. L. No. 107-110, § 101, Stat. 1425*.

Petersen, A. (2016). Perspectives of special education teachers on general education curriculum access: Preliminary results. *Research and Practice for Persons with Severe Disabilities, 41*(1), 19-35. https://doi.org/10.1177/1540796915604835

Snyder, R. A., & Pufpaff, L. A. (2021). Current state of high stakes teacher evaluation for special education teachers. *The Journal of Special Education Apprenticeship, 10*(1), 1-18.

Swartzer, K. (2017). *Self reg in the classroom*. Teaching Channel with Learners Edge. https://www.teachingchannel.com/k12-hub/blog/social-emotional-skills-of-self-regulation/

Troppe, P., Milanowski, A., & Heid, C. (2017). *Implementation of title I and title II-A program initiatives: Results from 2013–14*. National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/projects/evaluation/other_titleI.asp

Walkinshaw, N. M. (2022). *The relationship between Florida accountability programs and merit- based pay in two Large urban based school districts*. ProQuest. www.proquest.com/dissertations-theses/relationship-between-florida-accountability/docview/2671709230/se-2