Special Education and Communication
Disorders Faculty Publications

Department of Special Education and
Communication Disorders

8-30-2022

Characteristics of emotional disturbance of female and male students in elementary, middle, and high school

Douglas Cullinan

Matthew C. Lambert

Michael Epstein

RESEARCH ARTICLE

Characteristics of emotional disturbance of female and male students in elementary, middle, and high school

Douglas Cullinan[1] | Matthew C. Lambert[2] | Michael H. Epstein[2]

[1]Department of Teacher Education and Learning Sciences, North Carolina State University, Raleigh, North Carolina, USA

[2]Department of Special Education and Communication Disorders, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

Correspondence
Douglas Cullinan, Department of Teacher Education and Learning Sciences, North Carolina State University, Raleigh, NC, USA.
Email: dacullin@ncsu.edu

Abstract

Provide data on the five characteristics of emotional disturbance (ED). For 503 students with ED and 2016 without disabilities, teachers rated the characteristics (*Inability to Learn*; *Relationship Problems*; *Inappropriate Behavior*; *Unhappiness or Depression*; *Physical Symptoms or Fears*), plus *Socially Maladjusted*. We applied a 2 (ED, without disabilities) × 2 (female, male) × 3 (elementary, middle, high school) covariance analysis, with follow-up comparisons. Students with ED showed greater problems than students without disabilities on all five characteristics, and *Socially Maladjusted*. On *Inability to Learn*, among students with ED genders did not differ at elementary but males had greater problems at middle school. On *Inappropriate Behavior* and *Physical Symptoms or Fears*, students with ED varied across school levels but students without disabilities did not. All five characteristics discriminated students with ED from those without disabilities. Differences between genders and school levels varied across characteristics.

KEYWORDS
age, characteristics, emotional disturbance, gender, school

1 | INTRODUCTION

The Individuals with Disabilities Education Act (IDEA) is a US law that guarantees the right to appropriate education for students with disabilities. The IDEA names several categories of education disability, describes each category in general terms, and authorizes federal funding for some of the excess cost of educating a student with a disability. To receive such funding, states must affirm that the student qualifies under one of the IDEA categories of disability. A student with behavioral, personal, social, or emotional problems may qualify under the IDEA's emotional disturbance (ED) category if an assessment process determines that he or she meets the criteria described in the IDEA definition:

(i) *Emotional disturbance* means a condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree that adversely affects a child's educational performance: (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. (E) A tendency to develop physical symptoms or fears associated with personal or school problems.

(ii) ED includes schizophrenia. The term does not apply to children who are socially maladjusted, unless it is determined that they have an ED [under (i) above]. (Federal Register, 2006; p. 46756)

To restate briefly, this definition says that to qualify as ED a student must exhibit at least one of the characteristics (A) inability to learn, (B) relationship problems, (C) inappropriate behavior, (D) unhappiness or depression, or (E) physical symptoms or fears, in an enduring, intensive way that adversely affects her or his educational performance. Clause (ii) emphasizes that students with schizophrenia or who are socially maladjusted may or may not qualify for ED, depending on whether they meet the criteria stated in clause (i).

Students with ED are at high risk of maladaptive outcomes both later in school and in postschool functioning (e.g., Sanford et al., 2011; U.S. Department of Education, 2018). For example, compared to students without disabilities or with disabilities other than ED, students with ED are more likely to get suspended from school and leave school before graduating. Once out of school, large percentages of former students with ED experience employment problems and incarceration. Regrettable as such postschool outcomes are, they are not surprising in light of strong evidence that behavioral and emotional problems during childhood often predict mental disorders in adulthood, along with substance abuse, incarceration, frequent job loss, residential instability, and other maladaptive functioning (e.g., Copeland et al., 2015; Rutter et al., 2006).

Adequately measuring the five characteristics in the definition of ED is a critical part of an appropriate assessment for students suspected of qualifying for the ED disability. In the first place, such assessment is required by the IDEA to enable the multidisciplinary decision as to whether a student qualifies for the ED category by reason of exhibiting one or more of the characteristics to a sufficient extent. Relatedly, if a student with ED were to show improvement sufficient to consider removing the ED status, such a decision also ought to take into account the results of measuring the characteristics again. Third, measuring the characteristics has implications for intervention, even though the ED definition states the characteristics too generally to serve as objectives or to specify interventions for any particular student. For instance, suppose a student with ED shows notably maladaptive functioning on one particular characteristic. That would suggest an initial direction for practitioners to implement specific objectives and intervention practices for that student. Moreover, the characteristics have intervention policy implications. For example, school professionals ideally want to serve students with ED who experience any of the characteristics, but evidence-based school interventions appear to be limited for some characteristics, such as (D) unhappiness or depression (e.g., Arora et al., 2019; Corrieri et al., 2014; Feiss et al., 2019). To address this problem, potentially useful school-based interventions can be identified and organized according to the five characteristics. The ready catalog of intervention options may facilitate better planning and delivery of the intervention.

There is regrettably little research on the five qualifying characteristics in the IDEA definition of ED (Cullinan, 2004; Lambert, Cullinan, et al., 2021b). However, one can make inferences about them by considering

research on young people with diagnosed or otherwise confirmed behavioral and emotional disorders that seemingly correspond to characteristics (A)–(E) in the definition. For that reason, it is worth considering research on problematic school learning (e.g., Benner et al., 2002; DuPaul & Langberg, 2015), maladaptive social interaction and friendship (e.g., Bukowski et al., 2019), aggression and defiance (e.g., McMahon & Frick, 2019), depression (e.g., Fristad & Black, 2018; Keyes et al., 2019), and anxiety (e.g., Ghandour et al., 2019; Hammen et al., 2014). On the other hand, such results based on studies of children and adolescents with diagnosed mental disorders, suggestive as they may be, have uncertain generality to students identified with ED by their schools. Many US minors with a confirmed mental disorder (estimated to be at least 7 million; Ghandour et al., 2019) are not identified as a student with ED (numbering about 360,000; U.S. Department of Education, 2021); conversely, only a proportion of students with ED have any diagnosed mental disorder (e.g., Mattison, 2014).

Two studies examined the relative functioning of students with ED and students without disabilities on a teacher rating scale that measures the five characteristics of ED and a sixth variable, *Socially Maladjusted*, a concept named in the IDEA definition of ED. In Cullinan et al. (2003), elementary school level students with ED were found to be significantly higher (more problematic) than students without disabilities on all five characteristics and on *Socially Maladjusted*. Also, there were significant interactions of category and gender for *Relationship Problems* (for students with ED, no female–male difference, but for students without disabilities, males exceeded females) and *Physical Symptoms or Fears* (for students with ED, females exceeded males; for students without disabilities, no female-male difference). Cullinan and Sabornie (2004) studied middle and high school level students, again finding that students with ED were significantly higher on all five characteristics and on *Socially Maladjusted*. There were significant interactions of category and school level for three characteristics of ED: *Relationship Problems* (for students with ED, the middle school exceeded high school, but for students without disabilities, no differences across levels); *Unhappiness or Depression* (for students with ED, no difference across levels, but for students without disabilities, middle school was exceeded by high school); and *Physical Symptoms or Fears* (same pattern as for *Unhappiness or Depression*).

These two studies show that the characteristics of ED may be exhibited differentially not only by the category of a student but according to student demographic variables such as gender, level in school, and race or ethnicity. Similarly, research on mental disorders of young people has clearly shown, for some disorders, substantial female-male differences (e.g., Burstein et al., 2014; Hartung & Lefler, 2019) and age-level differences (e.g., Burstein et al., 2014; Weis, 2020), and has suggested differences by race, ethnicity, and socioeconomic status (Kessler et al., 2012). Thus, it is important to explore the possibility of interactions involving category, gender, and level in school on the five qualifying characteristics of ED.

In consideration of the foregoing research needs, the present study addresses gaps in knowledge about the characteristics of ED. We collected and analyzed data derived from a teacher rating scale specifically designed to measure six main aspects of the IDEA definition of ED, that is, characteristics (A)–(E) as earlier stated, and *Socially Maladjusted*. This allowed exploration of whether and in what ways students with ED differ from peers without disabilities on the six variables. Moreover, we subdivided the student participants by female and male gender and three age-related levels of schooling, to study how differences between students with ED and those without disabilities might vary by student gender and level in school.

2 | METHOD

2.1 | Participants

2.1.1 | Students

Student participants were 2519 US students aged 6–18 years. These participants were drawn from a larger sample of 2984 students rated by educators during the process of establishing norms and determining reliability and

validity for the Scales for Assessing Emotional Disturbance (3rd edition) Rating Scale (SAED-3 RS), one of four instruments in the SAED-3 assessment package (Epstein et al., 2020). To compare students with ED to those without disabilities, we excluded from analysis the 372 students who either had an IDEA disability other than ED (e.g., Autism, Specific Learning Disability) or who had an exceptional student-like condition (e.g., gifted, ADHD), even though they were not identified with an IDEA disability. Moreover, we excluded the 93 students younger than age 6, in part because in many school districts attendance is not required of children younger than age 6 years. As a result, participants were 503 students with ED (159 females, 31.6%) and 2016 students without disabilities (820 females, 40.7%).

Age, race-ethnic data, and geographic information on the two categories of student participants were as follows. Among the participants with ED: the mean age was 13.43 (SD = 3.31); race-ethnic statuses (U.S. Department of Education, 2018; terminology) were American Indian/Alaska Native, 1.0%, Asian/Pacific Islander, 2.9%, Black, 30.0%, Hispanic, 8.5%, White, 52.1%, Two or more races, 5.4%. Participants with ED came from all four census regions of the United States (West, South, Midwest, Northeast) including at least 16 states and 103 zip codes (state and/or zip code information was omitted for several dozen participants). Among participants without disabilities: the mean age was 12.31 (SD = 3.62); race-ethnic statuses were American Indian/Alaska Native, 0.6%, Asian/Pacific Islander, 2.1%, Black, 16.1%, Hispanic, 17.1%, White, 62.5%, two or more races, 1.6%. Participants without disabilities also came from all four census regions, including at least 29 states and 186 zip codes.

The resulting pool of 2519 students already differed by Category (with ED, without disabilities) and by Gender (female, male). We wanted study findings to be relevant to a very wide age range of students, yet we anticipated that results involving category, gender, or category × gender might apply at some age levels but not others. Therefore, we subdivided students by their ages into Levels of schooling: elementary, 6–11 years; middle school, 12–14 years; high school, 15–18 years.

2.1.2 | Raters

Rater participants were each student's teacher or, very infrequently, another educator who knew the student well. In addition to rating each student on the SAED-3 RS, these educators provided basic information about each student, including IDEA educational disability (if any), gender, and age in years.

2.2 | MEASUREMENT

2.2.1 | Instrument

The SAED-3 RS is a standardized, norm-referenced rating scale of 45 behavior and emotion problem items that operationalizes the five characteristics of ED stated in the IDEA definition of ED, plus the concept of "socially maladjusted" as found in part (ii) of that definition (Federal Register, 2006). The SAED-3 RS was developed in a stepwise way using input from teachers, assessors, and researchers with expertise related to students with ED, aided by statistical analyses of successive attempts to create separate scales for each of the five characteristic of ED, and socially maladjusted (Epstein & Cullinan, 1998, 2010; Epstein et al., 2002). The instrument has demonstrated good reliability and validity, of various kinds, in studies based on the original normative and psychometric data (e.g., Epstein & Cullinan, 1998) as well as new norming and psychometric data collected for the SAED-3 RS (Epstein et al., 2020; Lambert, Cullinan, et al., 2021a; Lambert, Cullinan, et al., 2021b; Lambert, Martin, Epstein, and Cullinan, 2021; Lambert, Martin, Epstein, Cullinan, and Katsiyannis, 2021). For example, Cronbach's coefficient α for each of the SAED-3 RS subscales (Epstein et al., 2020) were as follows: *Inability to Learn*, .92;

Relationship Problems, .86; *Inappropriate Behavior*, .91; *Unhappiness or Depression*, .83; *Physical Symptoms or Fears*, .79; *Socially Maladjusted*, .85.

To employ the SAED-3 RS, a teacher who knows the student well rates each of the 45 items on a 0–3 scale. A higher rating indicates that the behavior described by that item is present to a more problematic extent. Scoring the SAED-3 RS according to simple directions produces five subscales corresponding to the IDEA definition's five characteristics: *Inability to Learn*, *Relationship Problems*, *Inappropriate Behavior*, *Unhappiness or Depression*, and *Physical Symptoms or Fears*.

Inability to Learn consists of items that address problems that interfere with school success (e.g., "Gets distracted; doesn't pay attention to teachers or work"). *Relationship Problems* items assess trouble in establishing and maintaining relationships with peers and teachers (e.g., "Does not work well in group activities"). *Inappropriate Behavior* items relate to disruptive, defiant, or aggressive behaviors, such as "Disruptive, loud, or rowdy." *Unhappiness or Depression* subscale assesses the student's negative mood, affect, and thinking; for instance, "Experiences little pleasure or joy." The items in *Physical Symptoms or Fears* address the extent of the student's anxiety and physical distress (e.g., "Anxious, worried, tense").

Scoring also yields a sixth subscale score, corresponding to the IDEA definition's concept of socially maladjusted. The SAED-3 RS's *Socially Maladjusted* subscale items (e.g., "Vandalizes property in the community") measure the extent to which a student performs antisocial behaviors outside of school. Socially maladjusted is not one of the characteristics of ED and therefore cannot be used to justify identifying a student with ED, but it has been of enduring interest (e.g., Forness & Kavale, 2000; Grosenick & Huntze, 1980; Hanchon & Allen, 2013; Kauffman & Landrum, 2017). We analyzed this dependent variable to provide some empirical data that may help clarify this infrequently researched topic.

2.3 | PROCEDURE

2.3.1 | Rating students

In the process of creating national norms for the SAED-3 RS, large numbers of educators, nearly all teachers, agreed to rate one or several of their students. These ratings took place in 2016, 2017, and 2018. Raters were asked to complete the SAED-3 RS on either every one of their students, or a subset selected as follows: "First, decide how many students you will rate. Then start at one end (top or bottom) of your roll, and rate each student in order. Do not skip any student unless you have known that student less than 2 months. When you have completed scales on the number of students you decided, stop rating."

2.3.2 | Calculating subscale scores

From these completed SAED-3 RS ratings, we calculated subscale raw scores by summing the ratings (0–3) of the items on each subscale (Epstein et al., 2020). In this way, every student received a raw score for six dependent variables: *(A) Inability to Learn*, *(B) Relationship Problems*, *(C) Inappropriate Behavior*, *(D) Unhappiness or Depression*, *(E) Physical Symptoms or Fear*, and *Socially Maladjusted*.

2.4 | Data analysis

The purpose of the present study was to explore whether and in what ways students with ED and students without disabilities differ on the six subscales of the SAED-3 RS, especially as these two Categories may

intersect with two potential moderator variables, Gender and Level. Therefore, we used Stata v13 (StataCorp, 2013) to analyze data within a multivariate analysis of covariance framework to examine differences between students with ED and students without disabilities across the six SAED-3 subscale scores, while allowing Category to interact with Gender and Level, and while controlling for any group differences in years of age and race-ethnic status.

Student age (within each school level) and race-ethnic status are not of substantive interest in this study, but both were included in the analyses as control variables because of the research evidence that young people's maladaptive behavioral and emotional functioning may vary by age and race-ethnic status. Specifically, student age in years was entered into the analyses of covariance as a continuous covariate, while race-ethnic status was entered as a categorical factor with six mutually exclusive levels: American Indian/Alaska Native, Asian/Pacific Islander, Black, Hispanic, White, and Two Or More Races.

Even though the only effects of interest in the present study are those that involve Category, we included all main effects in the analysis, all two-way interactions involving Category, Gender, and Level, and the three-way interaction among Category, Gender, and Level. The statistical significance of multivariate effects was evaluated at the 0.013 level, which constitutes a conservative adjustment for the four substantive tests (i.e., main effect of Category, Category × Gender, Category × Level, Category × Gender × Level).

After interpreting the multivariate effects, factorial analysis of covariance models was estimated to examine the univariate effects for each of the six SAED-3 subscale scores. Because of multiple follow-up analyses, the Benjamini-Hochberg false discovery rate adjustment (Benjamini & Hochberg, 1995) was used to control the Type I error rate for the univariate tests. The false discovery rate was set at 0.01, so, in this case, univariate tests with p-values less than or equal to .0050 were statistically significant.

Statistically significant interactions involving Category were then probed to evaluate the nature of the interaction by computing *simple effects* for Category at each level of the moderator (i.e., Gender, Level) (Aiken & West, 1991). To illustrate, suppose there were a significant Category x Gender interaction; if follow-up found a significant Category simple effect, it would mean that the Category effect applied to one gender but not the other. The statistical significance of simple effects was obtained using the *contrast* command in STATA. Simple effects were evaluated at the 0.05 per-test significance level (Keppel & Wickens, 2004).

Effect sizes (ESs) were computed for significant main effects, interaction effects, and simple effects. Partial eta-squared (η_p^2) was computed for each multivariate and univariate main effect. η_p^2 indicates the proportion of variance attributable to the effect after accounting for other effects in the model. For multivariate effects, η_p^2 was computed from the degrees of freedom and the F ratio. For univariate effects (main, interaction, and simple effects), η_p^2 was computed based on the effect sum of squares and the residual sum of squares for the full analysis of covariance model. According to Cohen's (1988) guidelines, η_p^2 of 0.01–0.05 is a *small* ES, 0.06–0.13 is *medium*, and ≥0.14 is *large* (η_p^2 values < 0.01 are trivial).

For significant interactions, Cohen's f^2 was computed. Cohen's f^2 indicates the proportion of *explained* variance attributable to the interaction term. Instead of Cohen's general guidelines about f^2, we utilized a review of applied psychological research (Aguinis et al., 2005) which found that the mean interaction ES was $f^2 = 0.009$, the median ES was $f^2 = 0.002$, and that 75% of interactions were smaller than $f^2 = 0.0053$. Therefore, we considered f^2 values less than 0.002 as *small*, those ranging from 0.002–0.009 as *medium*, and ≥ 0.009 as *large*.

For significant univariate main effects and simple effects (with a single df), Cohen's d was computed using model-adjusted means and unconditional standard deviations, as recommended by the What Works Clearinghouse (2020). Cohen's d indicates the mean difference between the two groups being compared scaled in standard deviation units. We characterized each Cohen's d estimate according to Cohen's suggested ranges (Cohen, 1988): d values ranging from 0.20–0.49 are *small*, 0.50–0.79 are *medium*, and ≥0.80 are *large* (d values < 0.20 are trivial).

3 | RESULTS

The results from the multivariate analysis are reported in Table 1. The multivariate main effect of Category was statistically significant ($F_{6, 2496}$ = 140.47, p < .0001, η_p^2 = 0.252 [*large* ES]) indicating that, on the combination of six SAED-3 RS subscales, students with ED had significantly higher (more problematic) scores than their peers without disabilities. As noted above, the ES was *large*. The multivariate two-way interactions between Category and Gender ($F_{6, 2496}$ = 3.08, p = .0053, f^2 = 0.007 [*medium* ES]) and Category and Level ($F_{12, 4994}$ = 7.85, p < .0001, f^2 = 0.019 [*large* ES]) were statistically significant, indicating that the effect of Category varied across both Gender and Level. The three-way interaction between Category and Gender and Level was also statistically significant ($F_{12, 4994}$ = 5.37, p < .0001, f^2 = 0.013 [*large* ES]) indicating that at least one two-way interaction was inconsistent across the third grouping variable.

3.1 | Inability to Learn

Table 2 lists the results of the univariate analysis for the *Inability to Learn* subscale score. The main effect of Category was statistically significant ($F_{1, 2501}$ = 330.99, p < .0001, η_p^2 = 0.117 [*medium* ES], d = 0.728 [*medium* ES]) indicating that the mean score was significantly higher for students with ED than students without disabilities.

In addition, the three-way interaction of Category, Gender, and Level was statistically significant ($F_{2, 2501}$ = 5.32, p = .0050, f^2 = 0.004 [*medium* ES]) indicating that at least one two-way interaction was inconsistent. In this case, the Category × Gender interaction varied across Levels: that two-way interaction was statistically significant at the elementary level ($F_{1, 2501}$ = 4.19, p = .0407, f^2 = 0.002 [*medium* ES]) and the middle level ($F_{1, 2501}$ = 6.46, p = .0111, f^2 = 0.003 [*medium* ES]), but nonsignificant at the high school level ($F_{2, 2501}$ = 0.06, p = .8022). The simple effects follow-up revealed that at elementary school level (see Figure 1) there was no significant mean difference between male and female students with ED ($F_{1, 2501}$ = 0.06, p = .8016), but male students without disabilities had a significantly higher mean than female students without disabilities ($F_{1, 2501}$ = 26.79, p < .0001, d = 0.361 [*small* ES]). For middle school students (see Figure 2), male students with ED had a significantly higher mean than female students with ED ($F_{1, 2501}$ = 12.31, p = .0005, d = 0.636 [*medium* ES]), but there was no significant difference between male and female students without disabilities ($F_{1, 2501}$ = 2.27, p = .1321).

TABLE 1 Multivariate results

	Pillai's trace	df$_1$	df$_2$	F	p Value
Category	0.2524	6	2496	140.47	<.0001*
Gender	0.0225	6	2496	9.57	<.0001*
Level	0.0292	12	4994	6.17	<.0001*
Age	0.0034	6	2496	1.42	.2018
Race/ethnicity	0.0655	30	12,500	5.53	<.0001*
Category × Gender	0.0073	6	2496	3.08	.0053*
Category × Level	0.0370	12	4994	7.85	<.0001*
Gender × Level	0.0249	12	4994	5.25	<.0001*
Category × Gender × Level	0.0255	12	4994	5.37	<.0001*

*p ≤ .013.

TABLE 2 Univariate results for *Inability to Learn*

	SS	df	MS	F	p Value
Category	11,789.37	1	11,789.37	330.99	<.0001***
Gender	1208.35	1	1208.35	33.92	<.0001***
Level	292.26	2	146.13	4.10	.0166
Age	0.79	1	0.79	0.02	.8814
Race/ethnicity	276.24	5	55.24	1.55	.1707
Category × Gender	10.61	1	10.61	0.30	.5852
Category × Level	23.44	2	11.72	0.33	.7196
Gender × Level	157.06	2	78.53	2.20	.1105
Category × Gender × Level	378.89	2	189.44	5.32	.0050***
Category × Gender at Level					
Elementary	149.24	1	149.24	4.19	.0407**
Middle	230.10	1	230.10	6.46	.0111**
High school	2.14	1	2.14	0.06	.8022
Gender at Category (Level = Elementary)					
With ED	2.14	1	2.14	0.06	.8016
Without disabilities	954.23	1	954.23	26.79	<.0001**
Gender at Category (Level = Middle)					
With ED	438.47	1	438.47	12.31	.0005**
Without disabilities	80.85	1	80.85	2.27	.1321

Abbreviations: ED, emotional disturbance; MS, mean square; SS, sum of squares.

**p ≤ .05; ***p ≤ .005.

3.2 | Relationship Problems

Table 3 lists the results of the univariate analysis for the *Relationship Problems* score. The main effect of Category was significant ($F_{1, 2501}$ = 555.00, p < .0001, η_p^2 = 0.182 [*large* ES], d = 0.943 [*large* ES]). The mean *Relationship Problems* score was significantly higher for students with ED than students without disabilities, and this ES was *large*. No interaction effects were significant at the 0.005 adjusted significance level.

3.3 | Inappropriate Behavior

Table 4 presents the results of the univariate analysis for *Inappropriate Behavior*. The main effect of Category was significant ($F_{1, 2501}$ = 577.93, p < .0001, η_p^2 = 0.188 [*large* ES], d = 0.962 [*large* ES]): students with ED had a higher mean than students without disabilities.

In addition, the interaction of Category × Level was significant ($F_{1, 2501}$ = 5.88, p = .0028, f^2 = 0.005 [*medium* ES]). This interaction represents the simple effect of Level differing between students with ED and students without disabilities. Specifically, there was a significant difference at the 0.05 significance level for the Level factor for both students with ED ($F_{2, 2501}$ = 7.10, p = .0008, η_p^2 = 0.006 [*trivial* ES]) and students without disabilities ($F_{2, 2501}$ = 3.32,
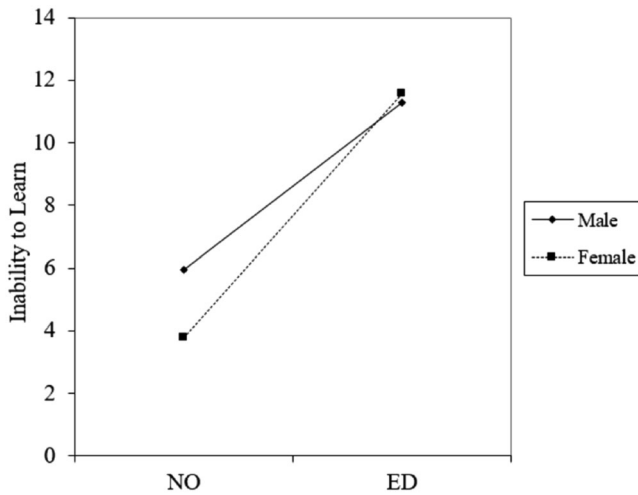
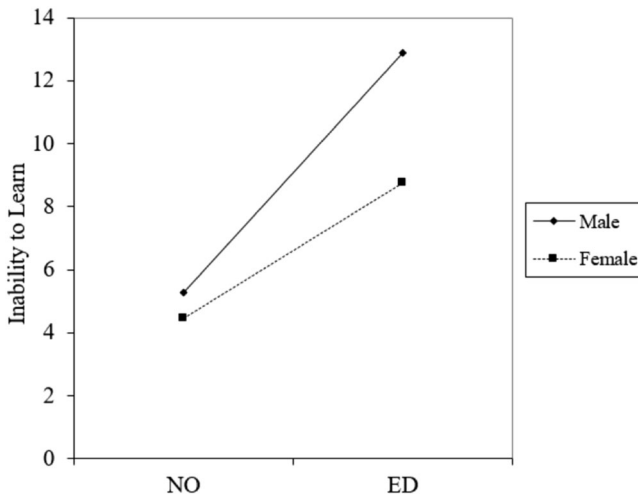**FIGURE 1** *Inability to Learn* Category × Gender interaction for elementary school students

FIGURE 2 *Inability to Learn* Category × Gender interaction for middle school students

$p = .0365$, $\eta_p^2 = 0.003$ [trivial ES]), so we examined details of Levels via follow-up pairwise comparisons. Among students with ED, the means for elementary ($F_{1, 2501} = 7.92$, $p = .0049$, $d = 0.296$ [*small* ES]) and middle school students ($F_{1, 2501} = 13.09$, $p = .0003$, $d = 0.317$ [*small* ES]) differed significantly from the mean for high school students (see Figure 3). However, among students without disabilities, there were no significant differences between pairs of school levels (see Figure 3).

3.4 | Unhappiness or Depression

Table 5 lists the results of the univariate analysis for *Unhappiness or Depression*. The main effect of Category was significant ($F_{1, 2501} = 477.15$, $p < .0001$, $\eta_p^2 = 0.160$ [*large* ES], $d = 0.873$ [*large* ES]): the mean score was significantly higher for students with ED than for students without disabilities. No interaction effects were statistically significant.

TABLE 3 Univariate results for *Relationship Problems*

	SS	df	MS	F	p Value
Category	5116.40	1	5116.40	555.00	<.0001***
Gender	27.59	1	27.59	2.99	.0837
Level	28.62	2	14.31	1.55	.2119
Age	2.35	1	2.35	0.26	.6134
Race/ethnicity	93.44	5	18.68	2.03	.0718
Category × Gender	13.03	1	13.03	1.41	.2345
Category × Level	80.30	2	40.15	4.36	.0129
Gender × Level	31.58	2	15.79	1.71	.1805
Category × Gender × Level	81.96	2	40.98	4.45	.0118

Abbreviations: MS, mean square; SS, sum of squares.
**p ≤ .05; ***p ≤ .005.

TABLE 4 Univariate results for *Inappropriate Behavior*

	SS	df	MS	F	p Value
Category	15,200.07	1	15,200.07	577.93	<.0001**
Gender	80.10	1	80.10	3.05	.0811
Level	416.72	2	208.36	7.92	.0004**
Age	13.28	1	13.28	0.51	.4773
Race/ethnicity	894.94	5	178.98	6.81	<.0001**
Category × Gender	176.13	1	176.13	6.70	.0097
Category × Level	309.17	2	154.58	5.88	.0028**
Gender × Level	253.86	2	126.93	4.83	.0081
Category × Gender × Level	32.68	2	16.34	0.62	.5373
Level at Category					
With ED	373.48	2	186.74	7.10	.0008***
Without disabilities	174.64	2	87.32	3.32	.0365***

Abbreviations: ED, emotional disturbance; MS, mean square; SS, sum of Ssquares.
**p ≤ .05; ***p ≤ .005.

3.5 | Physical Symptoms or Fears

Table 6 shows the results of the univariate analysis for the *Physical Symptoms or Fears* subscale. The Category main effect was significant ($F_{1, 2501}$ = 519.87, $p < .0001$, η_p^2 = 0.172 [*large* ES], d = 0.912 [*large* ES]) indicating that the mean score was significantly higher for students with ED than for their peers without disabilities.

In addition, the interaction of Category × Level was significant ($F_{1, 2501}$ = 10.66, $p < .0001$, f^2 = 0.008 [*medium* ES]). This arose because the simple effect of Level differed between students with ED and students without disabilities. That is, there was a significant difference between school levels for students with ED ($F_{2, 2501}$ = 12.12, $p < .0001$, η_p^2 = 0.007 [trivial ES]), but not for students without disabilities ($F_{2, 2501}$ = 2.24, p = .1068). Following up
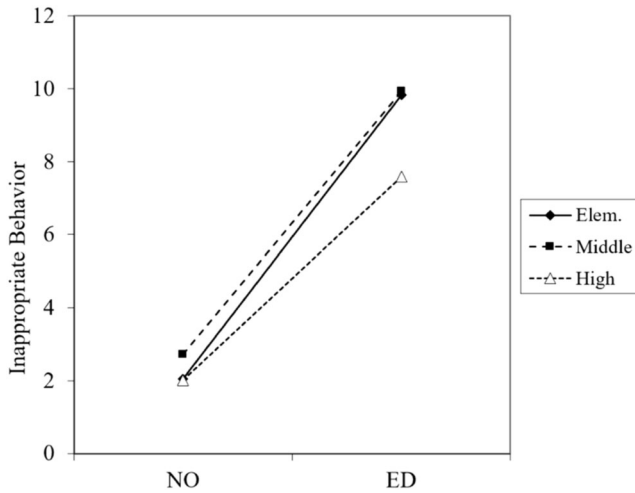
FIGURE 3 *Inappropriate Behavior* Category × School level interaction

TABLE 5 Univariate results for *Unhappiness or Depression*

	SS	df	MS	F	p Value
Category	4392.00	1	4392.00	477.15	<.0001***
Gender	9.69	1	9.69	1.05	.3049
Level	48.56	2	24.28	2.64	.0717
Age	13.96	1	13.96	1.52	.2181
Race/ethnicity	110.14	5	22.02	2.39	.0356
Category × Gender	1.12	1	1.12	0.12	.7264
Category × Level	30.34	2	15.17	1.65	.1926
Gender × Level	19.08	2	9.54	1.04	.3547
Category × Gender × Level	11.45	2	5.72	0.62	.5370

Abbreviations: MS, mean square; SS, sum of squares.

***$p \leq .005$.

on the simple effect finding for students with ED, the *Physical Symptoms or Fears* mean for elementary students was significantly greater than the means for students in middle ($F_{1, 2501} = 24.98$, $p < .0001$, $d = 0.459$ [*small* ES]) and high school ($F_{1, 2501} = 22.31$, $p < .0001$, $d = 0.461$ [*small* ES]), whereas the middle school and high school students with ED did not significantly differ (see Figure 4).
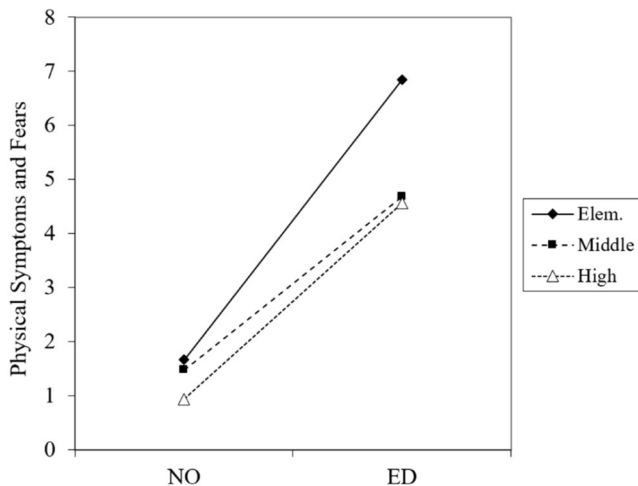
3.6 | Socially Maladjusted

Table 7 lists the results of the univariate analysis for the Socially Maladjusted subscale score. It revealed that the main effect of Category was significant ($F_{1, 2501} = 119.31$, $p < .0001$, $\eta_p^2 = 0.046$ [*small* ES], $d = 0.439$ [*small* ES]), indicating that the mean score was significantly higher for students with ED than for students without disabilities.

TABLE 6 Univariate results for *Physical Symptoms or Fears*

	SS	df	MS	F	p Value
Category	5004.49	1	5004.49	519.87	<.0001***
Gender	0.85	1	0.85	0.09	.7652
Level	140.37	2	70.18	7.29	.0007***
Age	2.54	1	2.54	0.26	.6073
Race/ethnicity	312.58	5	62.51	6.49	<.0001***
Category × Gender	6.66	1	6.66	0.69	.4054
Category × Level	205.18	2	102.59	10.66	<.0001***
Gender × Level	90.91	2	45.45	4.72	.0090
Category × Gender × Level	30.90	2	15.45	1.61	.2011
Level at Category					
With ED	233.34	2	116.67	12.12	<.0001**
Without disabilities	43.12	2	21.56	2.24	.1068

Abbreviations: ED, emotional disturbance MS, mean square; SS, sum of squares.

**p ≤ .05; ***p ≤ .005.

FIGURE 4 *Physical Symptoms or Fears* Category × Level interaction

In addition, the Category × Gender interaction was significant, $F_{1,2501} = 14.68$, $p = .0001$, $f^2 = 0.006$ [*medium* ES], meaning that the simple effect of Gender differed between students with ED and students without disabilities. Specifically, the *Socially Maladjusted* mean was significantly higher for female students with ED than male students with ED ($F_{1,2501} = 8.95$, $p = .0028$, $d = 0.189$ [trivial ES]), but there was no significant female–male difference for students without disabilities ($F_{1,2501} = 3.29$, $p = .0696$) (see Figure 5).

Also, the two-way interaction between Category and Level was statistically significant ($F_{1,2501} = 8.99$, $p = .0001$, $f^2 = 0.007$ [*medium* ES]). This interaction is explained as the simple effect of Level differing between students with ED and students without disabilities. That is, the Socially Maladjusted scores of the three school

TABLE 7 Univariate results for Socially Maladjusted

	SS	df	MS	F	p Value
Category	398.13	1	398.13	119.31	<.0001***
Gender	21.50	1	21.50	6.44	.0112
Level	25.43	2	12.71	3.81	.0222
Age	6.04	1	6.04	1.81	.1786
Race/ethnicity	99.76	5	19.95	5.98	<.0001***
Category × Gender	48.99	1	48.99	14.68	.0001***
Category × Level	59.96	2	29.98	8.99	.0001***
Gender × Level	36.21	2	18.10	5.43	.0045***
Category × Gender × Level	39.23	2	19.61	5.88	.0028***
Gender at Category					
With ED	29.87	1	29.87	8.95	.0028**
Without disabilities	10.98	1	10.98	3.29	.0696
Level at Category					
With ED	40.51	2	20.26	6.07	.0023**
Without disabilities	7.40	2	3.70	1.11	.3292
Category × Gender at Level					
Elementary	0.83	1	0.83	0.25	.6157
Middle	29.43	1	29.43	8.82	.0030**
High	73.38	1	73.38	21.99	<.0001**
Gender at Category (Level = Middle)					
With ED	37.64	1	37.64	11.28	.0008**
Without disabilities	0.10	1	0.10	0.03	.8636
Gender at Category (Level = High)					
With ED	59.03	1	59.03	17.69	<.0001**
Without disabilities	14.58	1	14.58	4.37	.0366**

Abbreviations: ED, emotional disturbance MS, mean square; SS, sum of squares.
**$p \leq .05$; ***$p \leq .005$.

levels differed significantly for students with ED ($F_{2,2501} = 6.07$, $p = .0023$, $\eta_p^2 = 0.005$ [trivial ES]), but did not differ for students without disabilities ($F_{2,2501} = 1.11$, $p = .3292$). Following up on the students with ED, the elementary school mean was significantly lower than the means for middle school ($F_{1,2501} = 9.56$, $p = .0020$, $d = -0.349$ [small ES]) and high school ($F_{1,2501} = 10.27$, $p = .0014$, $d = -0.313$ [small ES]) (see Figure 6).

The three-way interaction between Category × Gender × Level was also statistically significant, $F_{1,2501} = 5.88$, $p = .0028$, $f^2 = 0.005$ [medium ES], indicating that at least one of the two-way interactions was inconsistent. In this case, the Category × Gender interaction was inconsistent: it was not significant at the elementary level ($F_{1,2501} = 0.25$, $p = .6156$), but it was significant at both the middle ($F_{1,2501} = 8.82$, $p = .0030$, $f^2 = 0.004$ [medium ES]) and high school levels ($F_{1,2501} = 21.99$, $p < .0001$, $f^2 = 0.009$ [large ES]). Follow-up testing identified the specifics of these Socially Maladjusted findings at middle and high school levels. For middle school students, females with ED
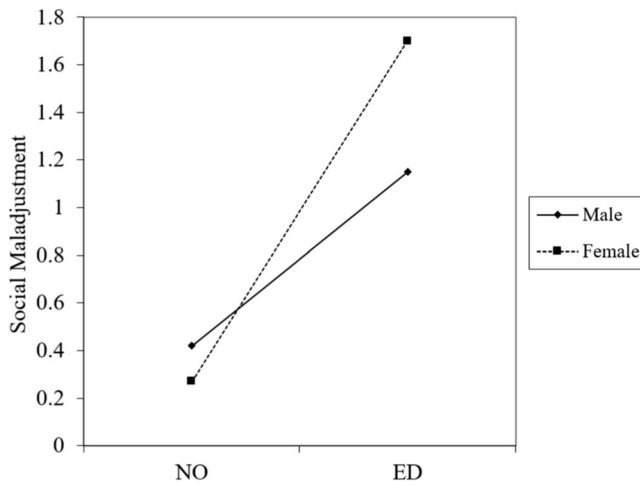
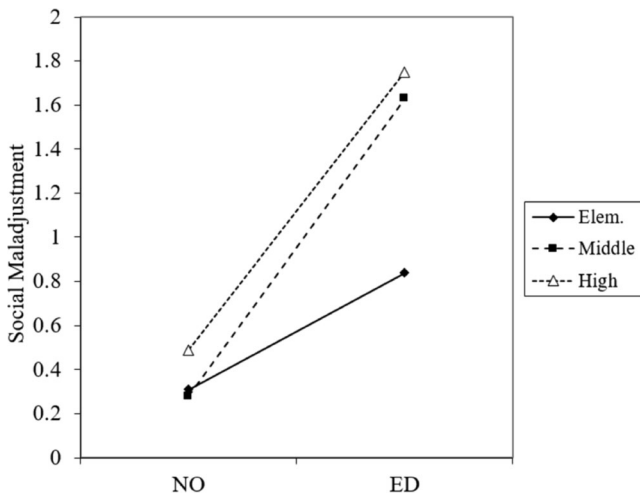**FIGURE 5** *Socially Maladjusted* Category × Gender interaction

FIGURE 6 *Socially Maladjusted* Category × School Level interaction

significantly exceeded males with ED ($F_{1, 2501}$ = 11.28, p = .0008, d = 0.427 [*small* ES]), but there was no difference between female and male students without disabilities ($F_{1, 2501}$ = 0.03, p = .8636). For high school students as well, female students with ED had a significantly higher mean than male students with ED ($F_{1, 2501}$ = 17.69, p < .0001, d = 0.305 [*small* ES]), but unlike the middle school findings, females without disabilities had a lower mean than males without disabilities ($F_{1, 2501}$ = 4.37, p = .0366, d = −0.137 [*trivial* ES]).

4 | DISCUSSION

4.1 | Multivariate results

The multivariate analysis examined a combination of the six subscale scores from the SAED-3 RS. This initial analysis provided the basis for the subsequent univariate analyses. The multivariate test indicated a significant main

effect of Category and significant Category × Gender, Category × Level, and Category × Gender × Level interactions, which were probed using univariate models.

As an aside, the IDEA definition of ED calls for each characteristic to be considered separately in practical decisions about student identification. That is, a student may qualify for ED by sufficiently exhibiting one or more of the five characteristics, but should not qualify based on a combination of multiple characteristics when each characteristic alone is exhibited to an insufficient extent.

4.2 | Univariate results

One obvious study result is that, compared to students without disabilities, students with ED showed significantly higher (more maladaptive) levels of all five emotional and behavior problem characteristics found in the IDEA definition of ED. For these five Category comparisons (students with ED vs. students without disabilities) the ESs were *medium* or *large*. Present results generally replicate earlier studies of the SAED-RS with other samples of students with ED and without ED (Cullinan et al., 2003; Cullinan & Sabornie, 2004).

These overall comparisons of students with ED versus without disabilities support the IDEA definition of ED because if we had found, for even one of the five characteristics, no significant difference between the two categories of students, the definition's validity would have been lessened. Instead, results indicate that the IDEA definition's five characteristics of ED do have validity in discriminating students with ED from students without disabilities, and that measuring the characteristics is relevant to understanding ED as an educational disability. Thus, the present findings suggest that the five characteristics can and should be assessed and considered in the process of determining whether a student qualifies for the ED category of special education.

Results provided clarifications and context for three characteristics of ED showing significant interaction involving Category and either Gender, Level, or both. On the *Inability to Learn* characteristic there was a significant three-way interaction among Category, Gender, and Level indicating that the effect of Category depended on both Gender and Level. Among students with ED, females and males did not differ at elementary or high school, but at middle school, females showed significantly less of this problem than males. In contrast, among students without disabilities, females showed significantly less of this problem than males at elementary school but the genders did not differ at middle or high school.

There were significant Category × Level interactions for both *Inappropriate Behavior* and *Physical Symptoms or Fears*, but these two characteristics showed different patterns among the Category × Level subgroups. Among students with ED, on *Inappropriate Behavior*, both elementary and middle school students showed significantly greater problems than high school students, whereas on *Physical Symptoms or Fears*, elementary students showed significantly greater problems than either middle or high school students. In contrast, among students without disabilities, elementary, middle, and high school students did not differ on either *Inappropriate Behavior* or *Physical Symptoms or Fears*.

No Category × Gender interaction was detected for any characteristic of ED, which suggests that each of the five characteristics is applicable to both female and male students. One finding, the Category × Gender × Level interaction for *Inability to Learn*, does imply that student gender should not be entirely ignored in considerations of the characteristics of ED. As noted earlier, female and male students with ED did not differ on *Inability to Learn* at elementary or high school, but did differ at middle school.

Turning to the *Socially Maladjusted* variable, students with ED showed significantly more problematic functioning than their peers without disabilities, again disregarding specifics of gender and level in school. This finding is compatible with research demonstrating that adolescent students with ED are likely to exhibit antisocial behaviors outside of school (e.g., Wagner, 1995), and that a high percentage of detained juveniles shows various mental disorders (Beaudry et al., 2021; Teplin et al., 2002). In addition, the Category × Gender × Level interaction provided interesting context. (a) Among elementary school students, females and males did not differ regardless of

category. (b) Among middle schoolers, females with ED exceeded males with ED while females and males without disabilities did not differ. (c) Among high school students, females with ED again exceeded males with ED, but females without disabilities had significantly lower scores than males without disabilities.

We are curious about the finding that among students with ED at both middle and high school levels, females significantly exceeded males on *Socially Maladjusted* (engaging in antisocial actions outside of school). The first possible explanation, of course, is that adolescent females with ED do exceed their male peers in performing antisocial acts in the community. Another potential explanation is that females actually do not perform such acts more than males, but teachers are more aware of *Socially Maladjusted* behaviors of the females (e.g., females with ED may be more likely than males with ED to speak about their antisocial behaviors). A third possibility is that for any given out-of-school antisocial act, teachers are more alarmed about its performance by their female students.

4.3 | Study limitations

Analyses for the present research utilized data collected via teacher ratings of students, and relied on results from the SAED-3 RS, a rating scale designed to measure the IDEA definition's five characteristics of ED and socially maladjusted. Some of the present results could be specific to the use of the SAED-3 RS to collect the data. Had we employed other methods (e.g., a different teacher rating scale, target behavior recording of observed classroom behavior, or student self-report of her or his own behaviors, emotions, and thoughts) to measure those aspects of the IDEA definition, we might have obtained different findings.

Although both categories of students included females and males, ages 6–18, six race-ethnic statuses, and all US census regions, it remains unknown how representative the present data are. This is because neither the students nor the educators who rated them were randomly selected. Instead, the present data were obtained from hundreds of educators who contributed their time and effort in response to our request to complete the SAED-3 RS and provide context information (e.g., student age and gender). But we made the same request of many additional educators who were unable or unwilling to so contribute. If teachers who did contribute and teachers who did not are systematically different in a way that affects student ratings, our data could be unrepresentative in unknown ways.

Relatedly, not all teachers rated the same number of students. Using our rating instructions, some teachers rated one student, some rated several, and others rated all students on their class rolls. This situation calls for a multilevel analysis in which students are nested within teachers, but because of the necessary anonymity of data collection, it was not feasible to identify which teacher rated each student. Therefore, single-level analysis models were used as if each student were rated by a different teacher, which results in downwardly biased standard errors and inflation in the Type I error rate.

4.4 | Practical implications

Results of the present study underscore the need for practitioners to be ready to intervene for various kinds of emotional and behavioral problems. Those problems very often represent one or more of the five characteristics in the IDEA definition of ED. It is recommended (e.g., National Association of School Psychologists, 2016) that interventions for those problems be delivered within a multitiered system of support (MTSS). The MTSS concept features three levels of support ("tiers") for learning and/or behavior functioning in school (e.g., Crone et al., 2015; Kern et al., 2016; McIntosh & Goodman, 2016; Mitchell et al., 2019). Tier 1 applies proactive instruction and management practices to all students to create positive relationships among students and staff members. A large majority of students will prosper in tier 1, but a small proportion will need extra interventions, often in small groups (tier 2). Even less often, assistance is delivered individually (tier 3), sometimes with help at home or in the

community from outside agencies. All three tiers feature evidence-based teaching and management practices together with movement across tiers based on each student's functioning and need for intervention intensity. Many students with ED need higher intensity interventions, delivered skillfully and with fidelity (Peacock Hill Working Group, 1991; Zaheer, et al., 2019) in school and, if appropriate, additional settings. Beyond these general points, there are more specific implications for practice.

4.4.1 | Assess characteristics for intervention

Results showed that all five characteristics of the definition discriminate students with ED from students without disabilities. This supports not only the validity of the characteristics but just as significantly, the wisdom of assessing them to help multidisciplinary team participants decide whether a student qualifies for special education under the ED category. Psychologists and others tasked with educational assessment therefore should have available various methods for assessing all five characteristics. Using assessment methods for which norms exist may assist in ED qualification decisions.

4.4.2 | Intervention guidance

Results of measuring the characteristics may help in planning intervention. For example, if the characteristics are measured via a teacher rating scale, the rater can be interviewed thereafter, perhaps with an emphasis on one characteristic that was rated especially high, or on particular items rated especially problematic. Such follow-up may yield details that suggest directions for intervention efforts. Other methods of measurement would require other forms of follow-up, but may likewise suggest how to proceed with intervention.

4.4.3 | Availability of interventions

Individual teachers, multidisciplinary team leaders, and other school and district professionals who will be involved in serving students with ED should have available various school-based interventions for each of the characteristics of ED, as well as interventions for other student problems that fall outside the five characteristics. One way to facilitate this would be to catalog interventions that reportedly address one or more of the characteristics. Additionally, asking relevant questions about each possible intervention can add important context. Some examples might include (a) what is the scientific evidence basis for this intervention?; (b) to what extent is this intervention feasible and otherwise appropriate for school implementation?; (c) does this intervention's relevance vary by student grade level, gender, education setting, or other context?; (d) how much training will I (or my staff) need to skillfully implement the intervention?

4.5 | Research needs

4.5.1 | Study replication

We need to know more about the ways in which students with ED experience learning difficulties, social disabilities, defiance and aggression, depression, and disabling anxiety. One important direction is for studies that replicate the present research, using the SAED-3 RS and other assessment methods to quantify and evaluate the five characteristics of ED. This is important because different insights about the characteristics may arise from studies that measure students' perspectives about themselves, about their school peers, that directly record classroom target behaviors relevant to the characteristics of ED, or measure in other ways.

4.5.2 | Increase data on characteristics

A limitation noted earlier is the possibility that participating and nonparticipating teachers might systematically differ as to how they would rate the same student. Perhaps highly persuasive colleagues or powerful rewards would induce greater participation in a replication of the present study. Another possibility would take advantage of the fact that the IDEA requires regular reports about how that law has been implemented, including various data about students with ED (e.g., U.S. Department of Education, 2018). However, data on the five characteristics of ED have not been presented in these reports, despite the fact that characteristics are a significant legal basis for identifying a student with ED. For future reports, it may be desirable that IDEA authorities collect data on the emotional and behavioral functioning of students with ED, including their functioning on the five characteristics.

4.5.3 | Improve data analysis

Our instructions to raters were designed to increase the number of teachers who elected to participate, but another result was that some teachers rated all their students, others rated several students, and some teachers rated only one. We recommend that in future replications, researchers take steps to allow for nested data analyses. That may involve, for example, directing all teachers to rate a set number of students and linking each rated student to the teacher who rated.

4.5.4 | Characteristics as marker variables

It hardly needs stating that students identified into the ED category show diverse emotional and behavioral problems. This obviously presents challenges to their educators, but also to consumers of research on school interventions because the diversity of emotional and behavioral functioning often makes it unclear how much similarity is shared by the "students with ED" from study to study. This external validity problem might be reduced if persons researching students with ED were encouraged to measure and report the functioning of their study participants on the five characteristics of ED. In that case, the characteristics of ED would become "marker variables" (Keogh et al., 1978) that help clarify the comparability and generalizability of results from different investigations of students with ED.

4.5.5 | Other research on assessment and intervention

Although findings shed light on the extent to which students exhibit the characteristics of ED, significant questions remain about the relationship of those characteristics to assessment and intervention. For instance, (a) are some characteristics more influential than others in a teacher's decision to refer a student to a multidisciplinary team for special education consideration, or in that team's decision as to whether the student qualifies for the ED category? (b) Are there situations in which characteristics of ED should be used as one basis for judging intervention outcome? In other words, if a school intervention improves behaviors, learning, emotions, and other measured aspects of student functioning, is it reasonable to anticipate that one or more characteristics of ED also will show improvement? (c) Can we discover interactions between certain school interventions and certain characteristics of ED? That is, does some "Intervention X" tend to produce greater improvement among students who are extreme on one particular characteristic than among students who are extreme on a different characteristic?

4.6 | A final point having research and practical implications

Child and adolescent mental health during COVID-19 greatly worsened in the United States (and globally), especially in regard to depression and anxiety, and especially among girls (Murata et al., 2021; Racine et al., 2021). As schooling in the United States recovers key aspects of its status before the pandemic (e.g., extensive in-person teaching, close interpersonal interaction), many young people may continue to experience such maladaptive behaviors, emotions, and thoughts. Does this portend a surge of students who qualify as ED under characteristics (D) unhappiness or depression, or (E) physical symptoms or fears? Will the students who do qualify tend to demonstrate those two characteristics even more extremely than students who qualified before COVID? The data for the present study, collected before COVID, can help answer such questions by indicating baseline levels against which to compare post-COVID student functioning on the five characteristics. In any event, for students who are identified with ED due in substantial part to their depression or anxiety, we hope that evidence-based, high-intensity school interventions will be readily available.

CONFLICTS OF INTEREST

D. C. and M. H. E. developed and revised the Scales for Assessing Emotional Disturbance (3rd edition), of which one instrument (Rating Scale) was used in the present study. They receive royalties from sales of those instruments. The remaining author declares no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings will be available from the authors upon reasonable request, following an embargo from the date of publication to allow for the commercialization of research findings.

ORCID

Douglas Cullinan http://orcid.org/0000-0002-5817-173X

REFERENCES

Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology*, 90(1), 94–107. https://doi.org/10.1037/0021-9010.90.1.94

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.

Arora, P. G., Collins, T. A., Dart, E. H., Hernandez, S., Fetterman, H., & Doll, B. (2019). Multi-tiered systems of support for school-based mental health: A systematic review of depression interventions. *School Mental Health*, 11(2), 240–264. https://doi.org/10.1007/s12310-019-09314-4

Beaudry, G., Yu, R., Långström, N., & Fazel, S. (2021). An updated systematic review and meta-regression analysis: Mental disorders among adolescents in juvenile detention and correctional facilities. *Journal of the American Academy of Child and Adolescent Psychiatry*, 60(1), 46–60. https://doi.org/10.1016/j.jaac.2020.01.015

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society (Series B)*, 57(1), 289–300. https://www.jstor.org/stable/2346101

Benner, G. J., Nelson, J. R., & Epstein, M. H. (2002). Language skills of children with EBD: A literature review. *Journal of Emotional and Behavioral Disorders*, 10, 43–56. https://doi.org/10.1177/106342660201000105

Bukowski, W. M., Laursen, B., & Rubin, K. H. (Eds.). (2019). *Handbook of peer interactions, relationships, and groups* (2nd ed.). Guilford.

Burstein, M., Beesdo-Baum, K., He, J., & Merikangas, K. R. (2014). Threshold and subthreshold generalized anxiety disorder among US adolescents: Prevalence, sociodemographic, and clinical characteristics. *Psychological Medicine*, 44(11), 2351–2362. https://doi.org/10.1017/S0033291713002997

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Copeland, W. E., Wolke, D., Shanahan, L., & Costello, E. J. (2015). Adult functional outcomes of common childhood psychiatric problems: A prospective, longitudinal study. *JAMA Psychiatry*, 72(9), 892–899. https://doi.org/10.1001/jamapsychiatry.2015.0730

Corrieri, S., Heider, D., Conrad, I., Blume, A., Konig, H.-H., & Riedel-Heller, S. G. (2014). School-based prevention programs for depression and anxiety in adolescence: A systematic review. *Health Promotion International*, *29*(3), 427–441. https://doi.org/10.1093/heapro/dat001

Crone, D. A., Hawken, L. S., & Horner, R. H. (2015). *Building positive behavior support systems in schools: Functional behavioral assessment*. Guilford.

Cullinan, D. (2004). Research issues in classification and definition of emotional and behavioral disorders. In R. B. Rutherford Jr., M. M. Quinn, & S. R. Mathur (Eds.), *Handbook of research in behavioral disorders* (pp. 32–53). Guilford.

Cullinan, D., Evans, C., Epstein, M. H., & Ryser, G. (2003). Characteristics of emotional disturbance of elementary school students. *Behavioral Disorders*, *28*(2), 94–110. https://doi.org/10.1177/019874290302800201

Cullinan, D., & Sabornie, E. J. (2004). Characteristics of emotional disturbance in middle and high school students. *Journal of Emotional and Behavioral Disorders*, *12*, 157–167. https://doi.org/10.1177/10634266040120030301

DuPaul, G. J., & Langberg, J. M. (2015). Educational impairments in children with ADHD. In R. A. Barkley (Ed.), *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (4th ed., pp. 169–189). Guilford.

Epstein, M. H., & Cullinan, D. (1998). *Scale for assessing emotional disturbance*. PRO•ED.

Epstein, M. H., & Cullinan, D. (2010). *Scales for assessing emotional disturbance* (2nd ed.). PRO•ED.

Epstein, M. H., Cullinan, D., Pierce, C., Huscroft-D'Angelo, J., & Wery, J. (2020). *Scales for assessing emotional disturbance* (3rd ed.). PRO-ED.

Epstein, M. H., Cullinan, D., Ryser, G., & Pearson, N. (2002). Development of a scale to assess emotional disturbance. *Behavioral Disorders*, *28*, 5–22. https://doi.org/10.1177/019874290202800101

Federal Register. (2006). *Assistance to states for the education of children with disabilities and preschool grants for children with disabilities; Final Rule 34. CFR Parts 300 and 301, 71 Fed. Reg. 46540 (August 14, 2006)*. https://sites.ed.gov/idea/idea-files/aug-14-2006-71-fr-46540/

Feiss, R., Dolinger, S. B., Merritt, M., Reiche, E., Martin, K., Yanes, J. A., Thomas, C. M., & Pangelinan, M. (2019). A systematic review and meta-analysis of school-based stress, anxiety, and depression prevention programs for adolescents. *Journal of Youth and Adolescence*, *48*, 1668–1685. https://doi.org/10.1007/s10964-019-01085-0

Forness, S. R., & Kavale, K. A. (2000). Emotional or behavioral disorders: Background and current status of the E/BD terminology and definition. *Behavioral Disorders*, *25*, 264–269. https://doi.org/10.1177/019874290002500304

Fristad, M. A., & Black, S. R. (2018). Mood disorders in childhood and adolescence. In J. N. Butcher (Ed.), *APA handbook of psychopathology* (Vol. 2, pp. 253–278). American Psychological Association.

Ghandour, R. M., Sherman, L. J., Vladutiu, C. J., Ali, M. M., Lynch, S. E., Bitsko, R. H., & Blumberg, S. J. (2019). Prevalence and treatment of depression, anxiety, and conduct problems in U.S. children. *Journal of Pediatrics*, *206*, 256–267. https://doi.org/10.1016/j.jpeds.2018.09.021

Grosenick, J. K., & Huntze, S. L. (1980). *National needs analysis in behavior disorders: Severe behavior disorders*. Department of Special Education, University of Missouri-Columbia.

Hammen, C., Rudolph, K. D., & Abaied, J. L. (2014). Anxiety disorders. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (3rd ed., pp. 338–365). Guilford.

Hanchon, T. A., & Allen, R. A. (2013). Identifying students with emotional disturbance: School psychologists' practices and perceptions. *Psychology in the Schools*, *50*(2), 193–208. https://doi.org/10.1002/pits.21668

Hartung, C. M., & Lefler, E. K. (2019). Sex and gender in psychopathology: DSM–5 and beyond. *Psychological Bulletin*, *145*(4), 390–409. https://doi.org/10.1037/bul0000183

Kauffman, J. M., & Landrum, T. J. (2017). *Characteristics of emotional and behavioral disorders of children and youth* (11th ed.). Pearson

Keogh, B. K., Major, S. M., Reid, H. P., Gándara, P., & Omori, H. (1978). Marker variables: A search for comparability and generalizability in the field of learning disabilities. *Learning Disability Quarterly*, *1*(3), 5–11. https://doi.org/10.2307/1510932

Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Pearson Prentice Hall.

Kern, L., George, M. P., & Weist, M. (2016). *Supporting students with emotional and behavioral problems: Prevention and intervention strategies*. Paul H. Brookes.

Kessler, R. C., Avenevoli, S., Costello, J., Green, J. G., Gruber, M. J., McLaughlin, K. A., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Merikangas, K. R. (2012). Severity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Archives of General Psychiatry*, *69*(4), 381–389. https://doi.org/10.1001/archgenpsychiatry.2011.1603

Keyes, K. M., Gary, D., O'Malley, P. M., Hamilton, A., & Schulenberg, J. (2019). Recent increases in depressive symptoms among US adolescents: Trends from 1991 to 2018. *Social Psychiatry and Psychiatric Epidemiology*, *54*(8), 987–996. https://doi.org/10.1007/s00127-019-01697-8

Lambert, M. C., Cullinan, D., Epstein, M. H., & Martin, J. (2021a). Differences between students with emotional disturbance, learning disabilities, and without disabilities on the five dimensions of emotional disturbance. *Journal of Applied School Psychology*, *38*, 58–73. https://doi.org/10.1080/15377903.2021.1895399

Lambert, M. C., Cullinan, D., Epstein, M. H., & Martin, J. (2021b). Factor structure of the Scales for Assessing Emotional Disturbance-3 Rating Scale for students identified with emotional disturbance. *Behavioral Disorders*, *47*, 187–195. https://doi.org/10.1177/01987429211012328

Lambert, M. C., Martin, J., Epstein, M. H., & Cullinan, D. (2021). Differential item functioning of the Scales for Assessing Emotional Disturbance-3 for White and Hispanic students. *Journal of Emotional and Behavioral Disorders*, *29*, 252–260. https://doi.org/10.1177/1063426620988658

Lambert, M. C., Martin, J., Epstein, M. H., Cullinan, D., & Katsiyannis, A. (2021). Differential item functioning of the Scales for Assessing Emotional Disturbance-3 for White and African American students. *Psychology in the Schools*, *58*(3), 1–16. https://doi.org/10.1002/pits.22463

Mattison, R. E. (2014). The interface between child psychiatry and special education in the treatment of students with emotional/behavioral disorders in school settings. In H. M. Walker & F. M. Gresham (Eds.), *Handbook of evidence-based practices for emotional and behavioral disorders* (pp. 104–128). Guilford.

McIntosh, K., & Goodman, S. (2016). *Integrated multi-tiered systems of support: Blending RTI and PBIS*. Guilford.

McMahon, R. J., & Frick, P. J. (2019). Conduct and oppositional disorders. In M. J. Prinstein, E. A. Youngstrom, E. J. Mash, & R. A. Barkley (Eds.), *Treatment of disorders in childhood and adolescence* (4th ed., pp. 102–172). Guilford.

Mitchell, B. S., Kern, L., & Conroy, M. A. (2019). Supporting students with emotional or behavioral disorders: State of the field. *Behavioral Disorders*, *44*(2), 70–84. https://doi.org/10.1177/0198742918816518

Murata, S., Rezeppa, T., Thoma, B., Marengo, L., Krancevich, K., Chiyka, E., Hayes, B., Goodfriend, E., Deal, M., Zhong, Y., Brummit, B., Coury, T., Riston, S., Brent, D. A., & Melhem, N. M. (2021). The psychiatric sequelae of the COVID-19 pandemic in adolescents, adults, and health care workers. *Depression and Anxiety*, *38*(2), 233–246. https://doi.org/10.1002/da.23120

National Association of School Psychologists. (2016). *Integrated model of academic and behavioral supports* [Position statement]. https://www.nasponline.org/research-and-policy/policy-priorities/position-statements

Peacock Hill Working Group. (1991). Problems and promises in special education and related services for children and youth with emotional or behavioral disorders. *Behavioral Disorders*, *16*, 299–313. https://doi.org/10.1177/019874299101600406

Racine, N., McArthur, B. A., Cooke, J. E., Eirich, R., Zhuy, J., & Madigan, S. (2021). Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: A meta-analysis. *JAMA Pediatrics*, *175*, 1142–1150. https://doi.org/10.1001/jamapediatrics.2021.2482

Rutter, M., Kim-Cohen, J., & Maughan, B. (2006). Continuities and discontinuities in psychopathology between childhood and adult life. *Journal of Child Psychology and Psychiatry*, *47*(3/4), 276–295. https://doi.org/10.1111/j.1469-7610.2006.01614.x

Sanford, C., Newman, L., Wagner, M., Cameto, R., Knokey, A.-M., & Shaver, D. (2011). *The post-high school outcomes of young adults with disabilities up to 6 years after high school: Key findings from the National Longitudinal Transition Study-2 (NLTS2)* (NCSER 2011-3004). SRI International. https://ncser.ed.gov/pubs

StataCorp. (2013). *Stata statistical software: Release 13*. StataCorp LP.

Teplin, L. A., Abram, K. M., McClelland, G. M., Dulcan, M. K., & Mericle, A. A. (2002). Psychiatric disorders in youth in juvenile detention. *Archives of General Psychiatry*, *59*(12), 1133–1143. https://doi.org/10.1001/archpsyc.59.12.1133

U.S. Department of Education, National Center for Education Statistics. (2021). *Digest of education statistics, 2019 (NCES 2021-009)*. https://nces.ed.gov/pubsearch/pubsinfoasp?pubid=2021144

U.S. Department of Education, Office of Special Education Programs. (2018). *40th Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act, 2018*. http://www.ed.gov/about/reports/annual/osep

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.

Weis, R. (2020). *Introduction to abnormal child and adolescent psychology* (4th ed.). Sage.

What Works Clearinghouse. (2020). *What Works Clearinghouse standards handbook, version 4.1*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/handbooks

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