

Characteristics of Students with Emotional Disturbance Manifesting Internalizing Behaviors: A Latent Class Analysis

Nicholas A. Gage

University of Florida

Abstract

Students receiving special education services for emotional disturbance (ED) present school personnel with many challenges and those challenges can typically be described as the manifestation of externalizing and internalizing behaviors. To date, most research has focused on students exhibiting externalizing behaviors. This study addresses this gap by examining (a) prevalence of students receiving special education services for ED manifesting internalizing behaviors, (b) characteristics of students with ED manifesting internalizing behaviors, and (c) the relationship between teachers' perceptions of internalizing behaviors and students' self-report of internalizing behaviors. Using a nationally representative sample of students receiving special education services for ED, a latent class analysis was conducted to identify an internalizing class of students. Characteristics of those students were then examined based on a direct assessment of students. Results indicate that approximately 7% of students receiving special education services for ED manifest internalizing behaviors and that teachers' ratings of in-class behavior were accurate when compared with students' self-report of internalizing behaviors.

Students receiving special education services for Emotional Disturbance (ED) present many challenges to educators and have among the worst short and long term outcomes among students with high incidence disabilities, including poor academic achievement, high suspension and drop-out rates, and increased risk for arrest (Bradley, Doolittle, & Bartolotta, 2008; Bradley, Henderson, & Monfore, 2004). Complicating intervention efforts is heterogeneity within the disability category (Wagner, Kutash, Duchnowski, Epstein, & Sumi, 2005), with students manifesting a myriad of social, emotional, and behavioral characteristics and profiles. ED heterogeneity can be mediated via two classes of behavior: externalizing and internalizing behaviors (Kauffman & Landrum, 2012). Externalizing behaviors are characterized

Preparation of this article was supported, in part, by a grant from the Institute for Education Sciences, U.S. Department of Education (R324B080007). Opinions expressed herein do not necessarily reflect the position of the U.S. Department of Education, and such endorsements should not be inferred.

Correspondence regarding this article should be addressed to Nicholas A. Gage; e-mail: gagenicholas@gmail.com

as behavioral patterns directed outwardly toward the social environment and include aggression, disruption, and hyperactivity. Internalizing behaviors are characterized as behavior patterns directed inwardly towards oneself and include depression, social-withdrawal, obsessive-compulsive behaviors, and selective mutism (Gresham & Kern, 2004; Morris, Shah, & Morris, 2002). Students with internalizing behaviors may have difficulty making and sustaining friendships absent of concomitant aggressive behavior, display depressed affect consistently across multiple settings, or be extremely reluctant to speak even though they have the requisite skills (Kendziora, 2004).

Typically, prevalence statistics for students receiving special education services for ED, including the annual Individuals with Disabilities Education ACT (IDEA, 2004) Reports to Congress (U.S. Department of Education, 2008) and research estimates (e.g., Forness, Freeman, Paparella, Kauffman, & Walker, 2012), aggregate all students within the ED category together, regardless of emotional or behavioral profile (i.e., internalizing, externalizing, comorbid). Studies have examined differential characteristics of students receiving special education services for ED (e.g., Wagner et al., 2005a), but to date, no study has examined behavioral profiles at the national level. This study addresses this gap by examining teacher reported behavior problems of students with ED using the Special Education Longitudinal Study (SEELS; see Wagner et al., 2005b) to identify the percentage of students manifesting externalizing, internalizing and comorbid behaviors in school. Additionally, this study extends the research by (a) examining characteristics of students with ED by behavioral profile (i.e., internalizing or externalizing) and (b) comparing teacher reports of student behaviors to students' self-reports to support behavioral profile classifications.

Characteristics of Students with Emotional Disturbance

Students receiving special education service for ED represent less than 1% of all students in U.S. schools (U.S. Department of Education, 2008). Although research suggests the presence of emotional and/or behavioral disorders is much greater than numbers for special education eligibility for ED would suggest (Forness et al., 2012); this study focuses exclusively on students receiving special education services for ED. A large body of epidemiological research exists examining behavioral profiles of children and youth presenting internalizing and externalizing behaviors (e.g., Child and Adolescent Health Measurement Initiative, 2012), but less is known about students in school receiving special education services for ED.

Studies have examined characteristics of students with ED using national databases. For example, Wagner et al. (2005a) examined demographic characteristics and functional skills, including cognitive and academic functioning, of students with ED using both SEELS and the National Longitudinal Transition Survey-2 (NLTS-2; see Wagner et al., 2005b), finding that students with ED are significantly more likely to be male, African-American, and living in poverty. Students with ED also have lower social skills and poor academic functioning compared with other students with high incidence disabilities. However, Wagner et al. (2005a) did not report on the prevalence of internalizing and externalizing behaviors of students and only included parent report of social skills, not teacher report of in-school performance. This distinction (parent versus teacher report) is important because special education services are based on school performance, not out-of-school performance (Yell, 2012).

A handful of studies have explicitly examined behavioral profiles of students receiving special education services for ED (e.g., Gresham, Lane, MacMillan, & Bocian, 1999; Montague, Enders, & Castro, 2005). For example, a series of studies by Cullinan and colleagues (e.g., Cullinan, Evans, Epstein, & Ryser, 2003; Cullinan & Sabornie, 2004) examined characteristics of students receiving special education services for ED using standardized teacher report measures, finding that girls and Caucasian students had higher levels internalizing behaviors. Studies also found high levels of comorbidity (externalizing and internalizing) across the samples.

Gresham et al. (1999) examined similarities and differences between students identified as at-risk for school failure manifesting internalizing or externalizing behaviors. Unlike previous studies, Gresham and colleagues first classified students as manifesting internalizing or externalizing behaviors using the *Social Skills Rating System-Teacher* (SSRS-T; Gresham & Elliott, 1990), a standardized measure of student social behavior, and then compared the groups across academic achievement and social/behavioral measures, including social and academic self-concept. Overall, findings indicated that students manifesting internalizing behaviors were more likely to be female and have more absences than students with externalizing behaviors, but were similar across most academic and social/behavioral outcomes. Although this study explicitly examined similarities and differences between students with internalizing and externalizing behaviors, the students in the study were not identified as receiving services for ED, only as at-risk for school failure.

Purpose

To date, prevalence and characteristics of students with ED manifesting internalizing behaviors have not been fully delineated in the literature. Therefore, this study was designed to examine a national sample of students receiving special education services for ED and (a) identify the percentage of students in the ED category manifesting primarily internalizing behaviors (b) replicate the Gresham et al. (1999) study by examining characteristics of students manifesting internalizing and externalizing behaviors, and (c) examine the relationship between teachers' perceptions of students exhibiting internalizing behaviors and student self-report of internalizing behaviors.

Method

Sample

This study used the Special Education Elementary Longitudinal Study (SEELS) dataset for all analyses. The SEELS dataset is a nationally representative sample of students receiving special education services in all IDEA disability categories. The SEELS study used a two-stage sampling procedure, enabling generalization to the national population of U.S. students with disabilities (see SRI International, 1998 for more details about the study design). Data were collected in three waves starting in the fall of 2000. Data for this study are from the first wave (2000-01 school year) because (a) the first wave contained the largest number of students with ED and (b) the research questions did not necessitate analysis of student outcomes across time therefore choosing the largest sample was the most appropriate choice for the analyses.

The wave 1 SEELS dataset included 1,176 students receiving special education services for ED (weighted sample ~200,000 students). This sample included students in grades 1st through 9th, with the majority of students (61%) in 3rd through 6th grade. Fifty-eight percent of the sample was Caucasian, 30% was African-American, 11% was Hispanic, and a majority of the sample was male (78%). Forty-seven percent of the students' family household income was \$25,000 or less, while only 15% were from households earning \$50,000 or more. Forty-two percent of students lived in urban settings, 48% of the students lived in suburban settings, and only 10% of students lived in rural settings.

Measures

Only information from the SEELS Teacher Survey, completed by each students' primary language arts teacher (general or special education), and the SEELS Direct Assessment Battery were included.

Table 1
11-Item Student In-Class Behavior Survey

How often does [the] student do each of the following in [your] class? (1) Never, (2) Sometimes, (3) Always	
Internalizing Items	Externalizing Items
<ul style="list-style-type: none"> • Appear lonely • Joins group activities without being told to • Have low self-esteem • Make friends easily • Act sad or depressed • Start conversations rather than waiting for others to talk first 	<ul style="list-style-type: none"> • Argue with others • Avoid situations that are likely to result in trouble • Control his or her temper in conflict situations with other students • Fight with others • Follow your directions

Teacher survey. The SEELS Teacher Survey included a variety of questions about the teacher, the services provided to each student, and the teachers' perceptions of individual students' ability and behavior. Included in this study were questions from the in-class behavior assessment, a 24-item scale asking respondents to rate how often students exhibited specific behaviors in their classroom on a 3-point Likert scale (1) *Never*, (2) *Sometimes*, or (3) *Very Often*. Upon review of the 24-items and preliminary analyses indicating limited contribution to model fit in the latent class analysis (see below for more details); the 24 items were reduced to 11 items focusing on the manifestation of internalizing or externalizing behaviors. Table 1 presents each item used.

Student direct assessment. To replicate Gresham and colleagues' (1999) study and examine the relationship between teacher and student report of behavior, student direct assessment items were included. The SEELS direct assessment battery included an abbreviated version of the Student Self-Concept Scale (SSCS; Gresham, Elliott, & Evans-Fernandez, 1992). The SSCS is a multidimensional norm-referenced self-report measure of self-concept. Students rated their self-confidence in response to academic and social questions. Two subscales were used in this study, the academic and the social self-concept subscales. The Cronbach's alpha, a measure of reliability, for the abbreviated SSCS was 0.73 for the academic self-concept subscale and 0.68 for the social self-concept subscale. The maximum value possible on both the social and academic self-concept subscales was 15, with the maximum value indicating high self-concept.

Consistent with Gershman et al. (1999), we included academic achievement. Academic achievement in reading and mathematics were assessed with the research edition of the Woodcock-Johnson Test of Achievement-3rd Edition (WJ-III; Woodcock, McGrew, & Mather, 2001). The study included letter word identification and passage comprehension subtests for reading and applied problems and calculations subtests for mathematics. Test-retest reliabilities are reported to range from 0.76 to 0.93 across the four subtests (Schrank, McGrew, & Woodcock, 2001). The WJ-III provides standard scale scores ($M = 100$, $SD = 15$) for each subtest, which were used in this study.

The SEELS direct assessment battery also included a measure of both friendship and loneliness, which was culled from a survey developed by Asher, Hymel, and Renshaw (1984). The original survey had 24-items, but 22 of the 24 were control items, and were dropped from the direct assessment battery by SRI. Students were presented with two statements, "I can find a friend when I need one", and "I am lonely at school", and asked to respond *yes* (1), *sometimes* (2), or *no* (3).

Data Analysis

Missing data. Missing data are a significant concern when using large national data sets, such as SEELS, and is a critical concern in quantitative research (Little & Rubin, 2002). To address missing data, the Expectation-Maximization (EM) multiple imputation procedure was used to impute missing values for each variable. The EM algorithm attempts to find a value for theta which maximizes $g(y \mid \theta)$ given an observed y , and does so by making use of the associated family $f(x, y \mid \theta)$ (Dempster, Laird, & Rubin, 1977). First, all items were removed with 75% or more missing values (contact author for list of deleted items); then the 75% criterion was defined, subjectively, based on prior research (Gorelick, 2006). EM imputation was then conducted with sets (or families) of related variables (i.e. same source, related items from survey) for construct validity with associated family items. Only imputed values were retained for families meeting Little's Missing Completely at Random (MCAR) test ($\chi^2 < .05$) (Little, 1988). This approach retained the full sample; no students were deleted from the analysis sample. All EM imputation was conducted using SPSS 19.0 Missing Value Add-in software.

Latent class analysis. The purpose of this study was to identify behavioral profiles (i.e., internalizing, externalizing, comorbid) of students with ED based on in-school performance. One way to define student profiles is to identify classes of students based on item-clustering on a measure of student performance, in this case, teacher reported in-class behavior. Underlying the assumption of

item clustering is the presence of latent classes (e.g. externalizing and internalizing) of students based on their teachers' perceptions of in-class behavior. That is, each observation (student) is a member of one, and only one, of T latent, or unobserved, classes (Magidson & Vermunt, 2004). The basic idea of latent class analysis is that some parameters of a hypothesized statistical model differ across unobserved subgroups (i.e. classes) and these subgroups form the categories of a categorical latent variable (Vermunt & Magidson, 2003). The goal of latent class analysis is the identification of the best fitting model (both statistically *and* conceptually), which results in T latent classes, and to identify the magnitude of effect (discrimination) for each indicator (item). Using the parameter estimates and the best-fitting model, the outcome is the generation of latent class profiles (% of students within each class, generally, and within items, specifically) and latent class membership for each subject in the dataset. The latent class analysis was conducted with Latent GOLD 4.5 (Vermunt & Magidson, 2005), a software program designed specifically for latent class analysis, utilizing maximum likelihood and posterior mode estimation and Bayes constants to eliminate boundary solutions. To address grade-level heterogeneity, each student's age was included as a covariate across all latent class models. A total of nine models were calculated and compared for best fit. The determination of the best model and the final number of latent classes was made based on the minimization of the Bayesian information criteria (BIC), the Akaike information criterion (AIC), the minimization of classification error, Entropy R^2 , and the proportion of total variance explained (R^2). Entropy R^2 indicates how well the model predicts class membership, with values closer to one indicating better prediction.

MANCOVA. Once the best fitting model was identified, each student's classification was used as the fixed factor in a multivariate analysis of variance (MANCOVA) model with student perceptions of themselves (social and academic self-concept, friendship and loneliness) and their mathematics and reading achievement, as dependent variables, controlling for student-level characteristics (ethnicity, gender, SES, and urbanicity). The MANCOVA model included Helmert contrasts and univariate tests with Tukey HSD post hoc comparisons performed after obtaining a significant multivariate effect. The MANCOVA was conducted to assess whether or not latent classes of students (i.e., manifesting internalizing or externalizing behaviors), were significantly different across dependent variables by identified latent class. The MANCOVA was conducted in SPSS 19.0.

Design weights. As noted above, all estimates were based on the weighted sample to generalize results to the national sample of

students with ED. SEELS utilized a stratified sampling procedure, not a random sampling procedure, therefore parametric statistics using the sample weights must utilize special software that adjust the standard errors, accounting for stratification, clustering and the sampling weight. Latent Gold 4.5 includes advanced options to incorporate survey designs into the analysis model and corrects the standard errors and Wald statistics (Vermunt & Madigson, 2005). Similarly, the MANCOVA was conducted in SPSS 19.0 using the Complex Samples Add-In to correct all standard errors and p -values. Therefore, all results are based on the weighted sample to generalize estimates to the population of students with ED.

Results

Latent Class Model Identification

To identify latent classes (i.e., behavioral profiles) of students with ED based on their teachers' perceptions of in-class behavior, a latent class analysis was conducted. A total of nine models were examined to identify the best fitting model (contact author for complete table of model fit statistics by model). Across all models, the 4-class model was the best fitting model because it had the least amount of classification error, smaller BIC and AIC values, and explained most of the variance in the model. The 4-class model also had the same Entropy R^2 as the 5-9 class models (.97), while the 1-3 class models were all smaller. The 1-3 class models were run multiple times with different start values, all resulting in the same model fit statistics. The classification error was less than 2% and the model explained ~94% of the variance. Approximately 67% of the students were classified into class 1, 14% were classified into class 2, 12% were classified into class 3, and 7% were classified into class 4. To increase interpretability of the classes, the percentage of students within each class was examined across each item on the in-class behavior assessment. Specifically, the percentage of students receiving the extreme values on the internalizing (e.g. 3 for "appears lonely") and externalizing subtest items (e.g. 3 for "fights with others") were examined (see Figures 1 and 2).

Review of the proportion of students within each class receiving extreme values on the internalizing subtest indicated that class 4 consistently received the extreme values for all six internalizing items. Review of the externalizing subtest items indicated that students in class 2 consistently received the most extreme values for each item, with the exception of the "follows directions" item, which indicated that all students with ED followed directions at least some of the time. Further review indicated that class 4 had low percentages of students at the extreme values on the externalizing subtest, with

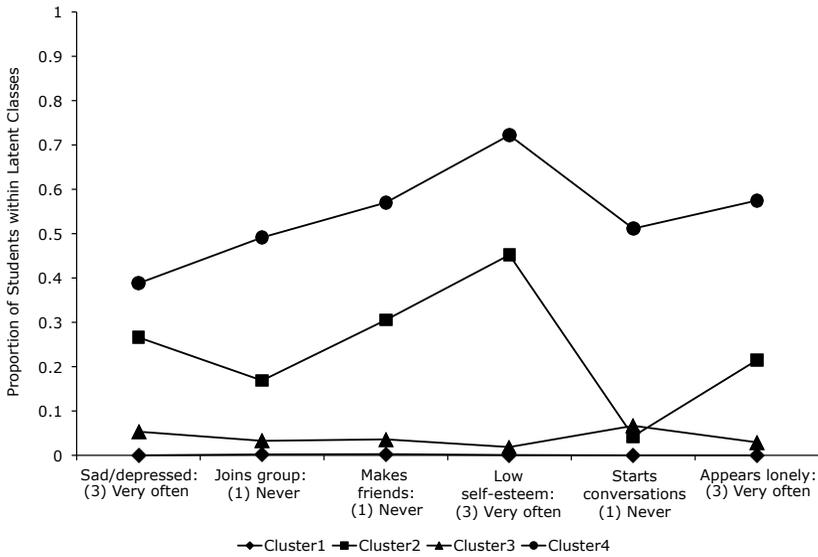


Figure 1. Proportion of students within classes receiving extreme values on the internalizing subtest

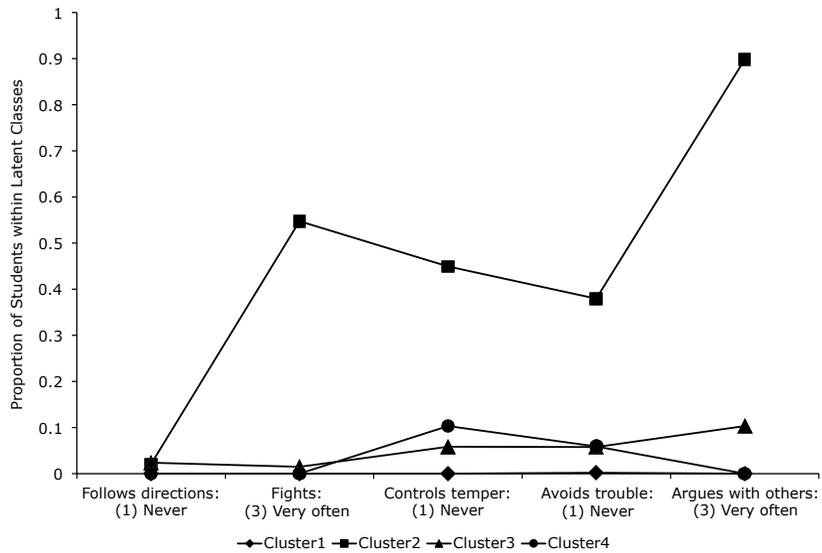


Figure 2. Proportion of students within classes receiving extreme values on the externalizing subtest.

the exception of “controls temper”. Class 2, which had the largest percentage of students at the extreme externalizing items, also had large percentages of students at the extreme values of the internalizing items. Class 1 always had the lowest percentage of students at the extreme values, while students in class 3 did not consistently or regularly receive extreme values on either the internalizing or externalizing subtests.

Naming Classes

To increase interpretability, each class was given a name based on response patterns on the in-class behavior assessment. It was clear that class 4 had the highest percentage of students at the extreme values on the internalizing subtest and generally smaller percentages of students at the extreme values on the externalizing subtest; therefore, class 4 was named the “internalizing class”, and represented 7% of sample. Similarly, class 2 had the largest percentage of students at the extreme values on the externalizing subtest, and was subsequently named the “externalizing class”, and represented 14% of the sample. Class 1 consistently had no students at the extreme values on either the internalizing or externalizing subtests, and was defined as the “control class” because these students did not vary on any of the in-class measures and represented 67% of the sample. Class 3 had small percentages of students at the extreme values on both the internalizing and externalizing subtests and as a result was defined as an “other” class as the students’ teachers reported that they had some significant in-class behavior problems, but neither extreme nor consistent, and represented 12% of the sample.

The largest percentage of students with ED was in the control class, indicating that the majority of students with ED did not consistently receive extreme values on the in-class behavior assessment. The second largest class was the externalizing class, providing evidence that most students in the ED category are those exhibiting externalizing behaviors in school. Similarly, the smallest percentage of students with ED was those manifesting internalizing behaviors in schools. Of note is that the externalizing class also had large percentages of students in the extreme value on the internalizing class, indicating that the students exhibiting the most extreme externalizing behaviors also exhibited the most extreme internalizing behaviors, suggesting comorbidity for the externalizing class.

Age as Covariate

As noted above, student age was used as a covariate in all models. Results from the 4-class model indicated that the parameter estimate for age for each class was statistically significant, which may

Table 2
Student Characteristics by Class

	(1) Control	(2) Externalizing	(3) Other	(4) Internalizing	4X1 Effect Sizes	4X2 Effect Sizes
Age	10.82	10.60	10.67	10.93	0.07	0.19
(1) Caucasian	54	68.2	55.9	75	0.57	0.20
(2) African American	30.2	29.4	33.8	17	-0.45	-0.43
(3) Hispanic	12.9	1.6	9.3	7.2	-0.39	0.95
(4) Asian / Pacific Islander	1.9	0	0.5	0.8	-	-
(5) American Indian / Alaska Native	0.9	0.8	0.4	0	-	-
(6) Multi/Other	0.1	0	0	0	-	-
Gender						
(1) Male	80.6	77.2	73.2	75	-0.2	-0.07
(2) Female	19.4	22.8	26.8	25		
SES						
(1) \$25,000 and Under	52.1	35.9	36.7	27.6	-0.64	-0.23
(2) \$25,001 to \$50,000	32.4	52.6	50.9	57.8	0.64	0.13
(3) Over \$50,000	15.5	11.5	12.4	14.6	0	0.17
Urban						
(1) Rural	9.0	18.7	12.5	6.7	-0.19	-0.71
(2) Suburban	48.9	42.3	45.3	57.9	0.22	0.38
(3) Urban	42	38.9	42.2	35.4	-0.17	-0.09

Note. All figures are percentages of students. Effect sizes are Hedges' g calculated using the dichotomous variable calculation used by What Work Clearinghouse. 4X1 is the comparison of internalizing to control and 4X2 is the comparison of internalizing to externalizing.

have been related to the large sample size. The parameter for the control condition was 0.012, the parameter for the externalizing condition was -0.093, the parameter for the internalizing class was 0.109, and the parameter for the other class was -0.028. The parameters indicate that students in the internalizing class were significantly more likely to be older students, while students in the externalizing class were significantly more likely to be younger in age.

Student Characteristics by Class

Student-level characteristics were compared and contrasted across identified classes. Table 2 provides the percentage of students by characteristics by class. What Works Clearinghouse's (WWC) dichotomous outcome effect sizes, which are similar to Hedges' g (see WWC, 2011 for computational description), were used to further examine differences by class. As such, dichotomous effect sizes at or above 0.25 are considered substantively important. Table 2 provides effect sizes for comparisons between the internalizing and the control classes (4X1) and between the internalizing and externalizing classes (4X2). Students in the internalizing class were significantly more likely to be Caucasian ($g = .57$) and significantly less likely to be African-American ($g = -.45$) and Hispanic ($g = -.39$) than the control class. Students in the internalizing class were significantly less likely to be African-American ($g = -.43$), but significantly more likely to be Hispanic ($g = .95$) than the externalizing class.

No significant differences were found between the internalizing class and the control and externalizing classes on gender. The internalizing class was significantly less likely to be low SES than the control class ($g = -.64$) and more likely from families earning \$25,001 to \$50,000 ($g = .64$). Students in the internalizing class were significantly less likely to be from rural settings ($g = -.71$) and significantly more likely to be from suburban settings ($g = .38$) compared with students in the externalizing class. Overall, review of the student characteristics by class indicated that students in the internalizing class were more likely to be Caucasian, middle class and live in suburban settings when compared with students in the control and externalizing classes.

Differences by Class on Social/Emotional and Academic Outcomes

In order to further describe the characteristics of students classified in the internalizing class and contrast differences with the externalizing class and the control class, differences across social/emotional and academic outcomes were examined. The MANOVA controlling for student-level characteristics was statistically significant, with a

Table 3
Covariate Adjusted Means and Observed Standard Deviations for Classes on
Social/Emotional and Academic Outcomes

Variable	Control	SD	Externalizing	SD	Other	SD	Internalizing	SD
Academic Self Concept	12.65	1.09	12.68	1.18	12.45	1.18	12.60	1.28
Social Self Concept	12.36	1.09	12.46	1.18	12.26	1.18	12.13	1.09
WJ-Applied Problems	89.61	9.25	89.74	9.23	89.41	9.00	90.13	9.12
WJ-Calculations	89.76	8.70	89.02	8.75	90.43	8.76	91.82	8.94
WJ-Letter-Word	86.22	10.33	86.11	10.17	86.72	10.19	89.67	10.22
WJ-Passage Comprehension	84.60	9.79	84.70	9.94	85.48	9.95	88.26	9.85
I can find a friend	1.38	0.54	1.37	0.24	1.43	0.24	1.49	0.36
I am lonely	2.57	0.54	2.63	0.47	2.58	0.47	2.48	0.36

Note. WJ is Woodcock-Johnson III Tests of Cognitive Ability (Woodcock, McGrew, & Mather, 2001).

Table 4
Class Contrasts Based on Covariate Adjusted Means for
Social/Emotional and Academic Outcomes

Variable	Internalizing X Control	Externalizing X Control	Internalizing X Externalizing
Academic Self-Concept	-0.05	0.02	-0.06
Social Self Concept	-0.21	0.10	-0.29
WJ-Applied Problems	0.06	0.01	0.04
WJ-Calculations	0.24	-0.08	0.32
WJ-Letter-Word	0.33	-0.01	0.35
WJ-Passage Comprehension	0.37	0.01	0.36
I can find a friend	0.20	-0.03	0.42
I am lonely	-0.18	0.10	-0.34

Note. All figures are Hedges' g standardized mean difference effect sizes based on covariate-adjusted means and sample standard deviations; WJ is *Woodcock-Johnson III Tests of Cognitive Ability* (Woodcock, McGrew, & Mather, 2001).

Wilk's $\Lambda = 746.16$. Because of the significant overall finding, Helmert contrasts and univariate tests with Tukey HSD post hoc comparisons were conducted for all included outcomes. Results for all univariate tests indicate that differences between all classes (fixed factor in model) across all measures were statistically significant. Therefore, to examine the practical importance of the univariate ANCOVA results, effect sizes were calculated for each comparison.

The covariate adjusted means and observed standard deviations for all outcome measures are presented in Table 4, and calculations of group differences by class are presented in Table 5. Using the Hedges' g criterion of .25 or above for substantively important difference, a number of class differences were evident. Students in the internalizing class reported significantly less social self-concept than students in the externalizing class ($g = -0.29$). The Friendship and Loneliness items further confirmed this; students in the internalizing class reported that they could not friend a friend at school and they often felt lonely when compared with the externalizing class (0.42 and -0.38 respectively). Social/emotional differences between the internalizing class and the control class followed a similar pattern, but the group differences (as defined by the effect sizes) were not as large. The results on the academic outcomes indicated a number of significant differences by class. The internalizing class performed significantly higher on three of the four academic outcomes (calculations, passage

comprehension, and letter-word identification), particularly on both reading measures when compared with the externalizing class. Similarly, the internalizing group performed higher on both reading measures (passage comprehension and letter-word identification) when compared with the control class.

Discussion

This study was designed to (a) identify prevalence of students in the ED category manifesting primarily internalizing behaviors, (b) replicate Gresham and colleagues' (1999) study by examining characteristics of students manifesting internalizing behaviors, and (c) examine the relationship between teachers' perceptions of students exhibiting internalizing behaviors and student self-report of internalizing behaviors.

Results from the nationally representative sample of students receiving special education services for ED indicate that an internalizing class of students, based on teacher perceptions, was identified and accounted for 7% of students with ED. Pragmatically, this result suggests that targeted intervention services for students exhibiting internalizing behaviors may only be necessary for a very small group of students. Yet, the high comorbidity for the externalizing students (i.e. high percentages of students at extreme values on the internalizing items) indicates that students in the externalizing class may also benefit from interventions targeting internalizing behaviors.

Unlike earlier research (e.g. Cullinan et al., 2003), students manifesting internalizing behaviors were statistically significantly older than other students with ED, while students manifesting externalizing behaviors were significantly younger. Still, this result should be interpreted with caution because the effect size difference in age between students in the externalizing and internalizing class of behaviors was small ($g = .19$). Results for gender found slight differences, with students in the internalizing class more likely to be female supporting Cullinan and colleagues (2003) and Gresham et al.'s (1999) results, but differences were minor and should be considered tenuous at best. The largest differences were in ethnicity, income, and the size/type of city the students lived in. Students with ED in the internalizing class were significantly more likely to be Caucasian and Hispanic, less likely to be from the lowest income group, and more likely to live in suburban settings. This finding was different from the Gresham et al. (1999) study; however, their sample was from five southern California school districts and was not nationally representative.

Class differences on academic, social, and behavioral outcomes were evident, particularly between the externalizing and internalizing classes. Students in the externalizing class were significantly

more likely to report high social self-concept, which was not wholly unexpected based on the theory of positive illusory bias (Gresham, Lane, MacMillan, Bocian, & Ward, 2000), or an overly positive view of oneself despite contradictory external indices. This finding was similar to the Gresham et al. (1999) that found students with internalizing behaviors report lower academic and social self-concept. Students in the internalizing class generally performed better across most of the academic measures, which was different from the Gresham et al. (1999) study that found no differences between the groups. Finally, students in the internalizing class reported that they were lonelier in school and that they could not find a friend when they needed one. Taken together, this study provides evidence of clear academic, social, and emotional differences between students primarily manifesting internalizing behaviors and other students with ED.

Another goal of this study was to examine the relationship between teacher perceptions of students with ED manifesting internalizing behaviors and those students' self-reported perceptions of internalizing behaviors (i.e., loneliness in school and the ability to make a friend). The results support teachers' ability to identify students that both manifest internalizing behaviors and self-report feeling lonely and an inability to find a friend. These results are encouraging and provide evidence of teachers' ability to accurately refer these students. Although students in the externalizing class also demonstrated elevated comorbidity with internalizing behaviors, particularly when compared with the control and other classes, students in the externalizing class did not self-report either low social self-concept or low on the loneliness and friendship measures. This result suggests that the relationship between teacher perceptions and student self-report for students manifesting externalizing behaviors may be more complex, given both the positive illusory bias and that these students, based on their teachers' report, are the most problematic in their classrooms. More research is needed to further describe and understand this class of students.

Limitations

This study used a nationally representative sample of students with ED, however a number of limitations warrant highlighting. Although the analysis adjusted standard errors and all p -values for accurate estimates, the results are just that, estimates of the population. Like all samples, generalization without further replication should be noted as exploratory. Similarly, the presence of missing data and the analysis with imputed values further reinforces that the findings are estimates of the population and not absolute values. Additional issues

with the dataset include the observed variables. The self-concept measure was adapted for SEELS from its original form, therefore direct comparison between the results from the database and the standardization sample of the original measure should not be made. The loneliness and friendship measures were problematic because they represented just one question on a 3-point Likert scale. A standardized measure of student self-reported depression and/or anxiety would have strengthened the results.

Regardless of the types of behaviors exhibited by students with ED, it is clear from both past research and personal experience in schools that these students need help. The students identified in this study as manifesting internalizing behaviors reported that they are lonely in school and have difficulty developing and maintaining friendships. This should concern us all and push us to continue developing efficient and effective interventions and supports for these student.

References

- Asher, S. R., Hymel, S., & Renshaw, P. D. (1984). Loneliness in children. *Child Development, 55*, 1456-1464.
- Bradley, R., Doolittle, J., & Bartolotta, R. (2008). Building on the data and adding to the discussion: The experiences and outcomes of students with emotional disturbance. *Journal of Behavioral Education, 17*, 4-23.
- Bradley, R., Henderson, K., & Monfore, D. A. (2004). A national perspective on children with emotional disorders. *Behavioral Disorders, 29*, 211-223.
- Child and Adolescent Health Measurement Initiative. (2012). *National profile of complementary and alternative medicine use for children with emotional, mental, or behavioral conditions or problems (2-17 years)*. Retrieved from www.Childhealthdata.org
- Cullinan, D., Evans, C., Epstein, M. H., & Ryser, G. (2003). Characteristics of emotional disturbance of elementary school students. *Behavioral Disorders, 28*, 94-110.
- Cullinan, D., & Sabornie, E. J. (2004). Characteristics of emotional disturbance in middle and high school students. *Journal of Emotional and Behavioral Disorders, 12*(3), 157-167. doi: 10.1177/10634266040120030301
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, 39*, 1-38.

- Forness, S. R., Freeman, S. F. N., Paparella, T., Kauffman, J. M., & Walker, H. M. (2012). Special education implications of point and cumulative prevalence for children with emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders, 20*, 4-18. doi. 10.1177/1063426611401624
- Gorelick, M. H. (2006). Bias arising from missing data in predictive models. *Journal of Clinical Epidemiology, 59*, 1115-1123. doi. 10.1016/j.jclinepi.2004.11.029
- Gresham, F. M., & Elliot, S. N. (1990). *Social skills rating system*. Circle Pines, MN: American Guidance Service.
- Gresham, F. M., Elliott, S. N., & Evans-Fernandez, S. (1992). *Student self-concept scale*. Circle Pines, MN: American Guidance Service.
- Gresham, F. M., Lane, K. L., MacMillan, D. L., & Bocian, K. M. (1999). Social and academic profiles of externalizing and internalizing groups: Risk factors for emotional and behavioral disorders. *Behavioral Disorders, 24*, 231-245.
- Gresham, F. M., Lane, K. L., MacMillan, D. L., Bocian, K. M., & Ward, S. L. (2000). Effects of positive and negative illusory biases: Comparisons across social and academic self-concept domains. *Journal of School Psychology, 38*, 151-175.
- Gresham, F. M., & Kern, L. (2004). Internalizing behavior problems in children and adolescents. In R. B. Rutherford, M. M. Quinn & S. R. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 262-281). New York, NY: The Guilford Press. *Individuals with Disabilities Education Improvement Act*, Author, Pub. L. No. 108-446 (2004).
- Kendziora, K. T. (2004). Early intervention for emotional and behavioral disorders. In R. B. Rutherford, M. M. Quinn, & S. R. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 327-351). New York, NY: The Guilford Press.
- Kauffman, J. M., & Landrum, T. J. (2012). *Characteristics of emotional and behavioral disorders of children and youth* (10th ed.). Upper Saddle River, NJ: Pearson.
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*, 1198-1202.
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. Hoboken, NJ: Wiley.
- Magidson, J., & Vermunt, J. K. (2004). Latent class models. In D. Kaplan (Ed.). *The Sage handbook of quantitative methodology for the social sciences* (pp. 175-198). Thousand Oaks, CA: Sage Publications.

- Montague, M., Enders, C., & Castro, M. (2005). Academic and behavioral outcomes for students at risk for emotional and behavioral disorders. *Behavioral Disorders, 31*, 84-94.
- Morris, R. J., Shah, K., & Morris, Y. P. (2002). Internalizing behavior disorders. In K. L. Lane, F. M. Gresham, & T. E. O'Shaughnessy (Eds.), *Interventions for children with or at risk for emotional and behavioral disorders* (pp. 223-241). Boston, MA: Allyn & Bacon.
- Schrank, F. A., McGrew, K. S., & Woodcock, R. W. (2001). *Technical abstract* (Woodcock-Johnson III Assessment Service Bulletin No. 2). Itasca, IL: Riverside.
- SRI International. (1998). *What do stakeholders want and need to know?* Menlo Park, CA: Author.
- U.S. Department of Education. (2008). *Thirtieth annual report to Congress on implementation of the Individuals with Disabilities Education Act*. Washington, DC: Author.
- Vermunt, J., & Magidson, J. (2003). *Statistical Innovation. Encyclopedia of Social Science Research Methods*. Thousand Oaks, CA: Sage Publications.
- Vermunt, J. K., & Magidson, J. (2005). *Latent GOLD 4.0 user's guide*. Belmont, MA: Statistical Innovations. <http://www.statisticalinnovations.com/technicalsupport/UG.html>
- Wagner, M., Kutash, K., Duchnowski, A. J., Epstein, M. H., & Sumi, W. C. (2005a). The children and youth we serve: A national picture of the characteristics of students with emotional disturbances receiving special education. *Journal of Emotional and Behavioral Disorders, 13*, 79-96.
- Wagner, M., Kutash, K., Duchnowski, A. J., & Epstein, M. H. (2005b). The special education elementary longitudinal study and the national longitudinal transition study: Study designs and implications for children and youth with emotional disturbance. *Journal of Emotional and Behavioral Disorders, 13*, 25-41.
- What Works Clearinghouse. (2011). *What Works Clearinghouse: Procedures and standards handbook* (Version 2.1). Washington, DC: Author.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Examiner's manual. Woodcock- Johnson III Tests of Cognitive Ability*. Itasca, IL: Riverside Publishing.
- Yell, M. L. (2012). *The law and special education* (3rd ed.). Upper Saddle River, NJ: Pearson Education.

