



CORPUS PUBLISHERS

Current Research in Psychology and Behavioral Science (CRPBS)

ISSN: 2833-0986

Volume 4, Issue 2, 2023

Article Information

Received date : 27 January, 2022

Published date: 07 February, 2023

*Corresponding author

Rachel L Oates, College of Education and Human Development, University of Delaware, 25 North College Avenue, Willard Hall Education Building, Suite 113, Newark, DE, USA

DOI: 10.54026/CRPBS/1086

Key Words

Disability stigma; Teacher Screening; Teacher support

Distributed under: Creative Commons CC-BY 4.0

Review Article

Disability Stigma Perspectives and Practice Survey: Predicting Teacher Efforts to Improve Acceptance

Rachel L Oates*

College of Education and Human Development, University of Delaware, USA

Abstract

Disability stigma is widespread in society, particularly in schools, and generates long-term risks for those stigmatized. Research shows teachers are well situated to address stigma in classrooms through empirically-based practices. However, disconnect arises between empirical recommendations and practical applications in schools. Therefore, further research was required to identify causes and potential implications. Survey measures for related constructs of teachers' backgrounds, personal attitudes towards students with disabilities, awareness of public stigma, use of empirically-based practices to increase acceptance, and perceived barriers to implementing practices were not yet developed. In this study, the Disability Stigma Perspectives and Practice (DSPP) survey was created for these purposes. The composite scales were adapted from existing measures or developed originally, and analyzed to ensure reliability, validity, and accurate factor structure. In initial analyses, 142 participants recruited through a school district completed the survey electronically. Correlational, item, and rasch analyses were conducted to revise the scales. Subsequently, 330 additional participants from the school district responded to finalized measures. Further correlational analyses, and exploratory and confirmatory factor analyses were completed. Implementation of the DSPP allows for examination of teachers' impacts on use of empirically-based practices to address disability stigma, and informs methods of supporting teachers for this purpose.

Disability Stigma Perspectives and Practice Survey: Predicting Teacher Efforts to Improve Acceptance

In the current educational environment, teachers retain the critical and complex responsibility to work towards increasing the acceptance of students with disabilities [1-5]. However, many teachers experience challenges in implementing initiatives for this purpose [6-10]. To investigate the disconnect between the research and legislative recommendations and the practical classroom application, it is necessary to examine the relationships of teachers' backgrounds, personal attitudes, perceptions of public stigma, classroom behavior, and reported barriers in order to identify methods to support teachers in the imperative practice of addressing disability stigma. Initially, surveys would be most effective as a means of collecting foundation data for this purpose due to feasibility and practicality. However, survey measures for the constructs of teachers' personal stigma, perceptions of public stigma, employment of empirically based practices to address disability stigma, and barriers to addressing disability stigma are lacking in the current field of research. Furthermore, many of the measures that have been developed use other methods of data collection rather than surveys, address different types of stigma or practices, were created for dissimilar respondent groups, or focus on specific disabilities.

Measures of personal attitudes

Stigma measures present in the literature are not specifically focused on gathering survey data from teachers regarding their personal attitudes towards students with disabilities. Some existing measures instead employ different methods of data collection rather than survey, such as interviews [11] or qualitative surveys in which respondents select descriptive adjectives that may characterize individuals with disabilities [12]. Alternatively, previously developed measures may focus on other types of stigma. For example, the Personal Stigma for Seeking Career Counseling scale gathers information about stigma towards counseling rather than stigma towards individuals with disabilities in general [13]. Several existing measures for personal stigma were developed for diverse respondent groups that are not relevant to the current study, including children [14], adolescents [9], or parents [15,16]. Finally, some measures focus on personal stigma towards individuals with particular disabilities, such as the Placement and Services Survey (PASS), which gathers data on respondents' attitudes towards students identified with various specific disabilities including autism and intellectual disabilities [8].

Measures of perceptions of public stigma

The existing measures used to study perceptions of public stigma have similar limitations. In particular, some such measures employ methods of data collection apart from surveys, including written responses to open-ended questions, as in the teacher impression journals [10]. Some measures examine different types of public stigma, such as that for seeking counseling or psychological support [13,17]. There are also measures in the literature that were developed to assess perceptions of public stigma through surveying respondent groups aside from teachers, such as individuals with disabilities themselves [18]. Finally, some measures of public stigma focus on specific disabilities, including the ADHD Stigma Questionnaire (ASQ; [19,20]).

Measures of practices to address disability stigma

Measures of teachers' practices to address disability stigma in the classroom did not exist in the literature. Some studies have collected data on teachers' practices for similar purposes, such as enhancing inclusion of students with disabilities or implementing behavioral or academic interventions [6,8,10,21]. However, many used alternative methods of data collection, including observation or written responses to open-ended queries [6,10,21].



Measures of barriers to addressing disability stigma

Previously developed measures of barriers to teachers' employment of practices in the classroom have either utilized data collection methods apart from surveys or examined practices for purposes distinctive from addressing stigma. Specifically, some studies have collected data on barriers through observation [6] or brief summary ratings of practice acceptability or feasibility [9]. Other measures in the literature focus on barriers to implementing interventions for other purposes such as successful inclusion [8] or behavioral support for students with disabilities [22]. Thus far, measures for the intended constructs are severely limited.

Methods

Measures

The Disability Stigma Perspectives and Practice (DSPP) survey was partially adapted from several existing and empirically supported measures. The five-part questionnaire focuses on teachers' personal attitudes towards students with disabilities, recognition of public stigma, attempts to address stigma, barriers to addressing stigma, and demographics (see Appendix). In the first section of the DSPP examining teachers' personal attitudes, participants were asked to read a vignette describing behaviors and test scores of a fictional special education student who would be in their class the following year. Following the vignette, they were asked to respond to questions about their feelings towards the fictional student. Both the vignette and subsequent questions were adapted from a section of the PASS, which was adapted and tested in a study conducted by Segall [8]. This particular section of the PASS was previously adapted from the Parental Attitudes Toward Children with Handicaps (PATCH) scale [16]. The PASS includes several vignettes about students with different diagnostic labels and students without a diagnostic label. Based on the findings from the study, there was no difference between responses for the various diagnostic labels. Therefore, the vignette has been altered to describe the fictional student as "a special education student with an IEP," to generalize the responses. The questions were drawn from a section in the PASS focusing on teacher attitudes towards the fictional student. Research supports the analysis of responses to this section by construct [8,16]. Demographics of the samples used for this testing were geographically, racially, and professionally representative, with one study involving teachers and another sampling parents. Reliability was found to be adequate with an overall coefficient alpha of 0.88 [16].

The validity of measured attitudes was evaluated by analysis of responses based on respondents' previous contact with individuals with disabilities, which is known to improve attitudes towards such individuals. It was found that participants who personally knew someone with a disability had significantly higher measured attitudes than those who did not (28.9 versus 26.6) [16]. Three of the six constructs were included in the present survey: affective attitudes (items 1 to 5), cognitive attitudes (items 6 to 10), and behavioral intentions (items 11 to 16). The responses were on six-point Likert scales ranging from "strongly disagree" to "strongly agree," as in the original PASS.

The second section of the survey gathered information about respondents' views of public perceptions, or their recognition of public stigma. Items for this section were adapted from the ASQ. The psychometric properties of the ASQ were evaluated in a study conducted by Kellison, et al. [19,20]. Although this study involved sampling of adolescent respondents, other demographics such as gender and racial background were more representative. Confirmatory factor analysis supported a three-factor structure, including disclosure concerns, negative self-image, and concern with public attitudes (goodness-of-fit index=0.96; root mean square error of approximation=0.06). Internal consistency reliability was found to be adequate for both the overall measure ($\alpha=0.92$) and for each construct (disclosure concerns: $\alpha=0.83$; negative self-image: $\alpha=0.80$; concern with public attitudes: $\alpha=0.84$). The test-retest reliability was found to be adequate with a correlation of 0.71 when given at two-week intervals. Construct validity was also determined to be supported by related constructs such as clinical

maladjustment, depression, self-esteem, and emotional symptoms. For the current study, the thirteen items in the factor of concern with public attitudes from the ASQ were selected. Items were adjusted to ask about "children" rather than "people" and about "disabilities" rather than "ADHD." Also, one item on jobs was removed because of its irrelevance to children in schools, leaving twelve items. The responses were on four-point Likert scales from "strongly disagree" to "strongly agree," as on the original ASQ.

Teachers were asked about their attempts to address stigma in the third section of the DSPP, which was developed for this study. The respondents were provided with a list of empirically based practices [7] and asked to indicate for each if they had used it in the past year, how effective they thought it would be on a four-point Likert scale from "not effective" to "very effective," and if they would use it if they had "more time, training, resources, etc." Twenty-three empirically based practices identified in the literature as sufficiently research supported were included [7], as well as an opportunity for respondents to write in additional practices (see items in Appendix).

In the fourth section of the DSPP, respondents were asked about their perceived barriers to addressing stigma. Items for this section were adapted from the Barriers to Implementing Evidence-Based Interventions in the Classroom survey [22]. In a study conducted by McGoey, et al. [22] involving a sample of classroom teachers varied in training and experience, this survey was found to be adequately reliable, with a Cronbach alpha of 0.90. All nineteen items from the original survey were included, although six items were slightly altered to refocus on interventions for reducing stigma of students with disabilities. Instructional wording was also modified to improve respondents' understanding. Items 1 and 21 were added pertaining particularly to this study. Responses were on seven-point Likert scales from "not at all" to "extremely," as on the original survey. The fifth section gathered information on participants' background and demographics, including their school, gender, age, ethnicity, highest degree, graduation year, special education certification, teacher title, years of teaching experience, position, experience with individuals with disabilities, school structure, school wide campaigns, and involvement therein.

Participants and Procedure

The participants were recruited through a large school district in south Florida, with the approval of the appropriate district officials. In the first phase of the study, 142 teachers were surveyed and 330 different teachers were surveyed in the second phase. Based on an a priori power analysis completed using G*Power 3.1 software with overall power set to 0.80 and an anticipation of a medium effect size [23] ($f=0.25$), a sample size of at least ninety six participants was required [24]. The intended sample size was larger to ensure more accurate results. The eligibility criteria included those currently working in schools as teachers at the elementary, middle, or high school level. Information about the study and invitations to participate were distributed through email to the eligible members of the school district. The email invited potential participants to use an included link to participate in the survey via the online survey software Qualtrics or to contact the researchers with any questions. Consent forms were included in the online survey as an access step to the survey measure. The consent procedures were followed as required for research with human subjects by the University of Delaware Institutional Review Board and the school district Department of Research and Evaluation. Current demographic information on the teaching workforce was obtained and effort was made to approximate these demographics in the participant sample in order to increase the external validity of the findings. When compared with a sample of the teaching workforce gathered by the National Center for Education Statistics [25], the participants of both the first and second phases of this study were found to be similar in gender, age, ethnicity or race, highest degree earned, years of teaching experience, and grades taught, as quantities of all characteristics were within 12 percentage points for the samples. The demographics of the participants from each phase are presented alongside those of the national teaching workforce in (Table 1).

**Table 1:** Demographic characteristics of participants and teaching workforce.

Characteristic	Pilot Study Participants		Research Study Participants		Teaching Workforce ^a	
	n	% ^b	n	% ^b	n (in thousands)	% ^b
Total	142	100	330	100	3,385	100
Gender						
Female	116	82	269	82	2,584	76
Male	25	18	59	18	802	24
Other	1	1	2	1	0	0
Age						
20-39	59	42	135	41	1,497	44
40-59	63	44	162	49	1,632	48
60 or more	12	8	30	9	256	8
Did not provide	9	6	3	1	0	0
Ethnicity/Race						
White	108	76	268	81	2,773	82
Black or African American	12	9	31	9	231	7
American or Alaska Native	1	1	1	0	17	1
Asian or Asian American	1	1	3	1	61	8
Multiracial	9	6	13	4	35	1
Other	11	8	14	4	0	0
Highest Degree						
High School diploma	0	0	1	0	0	0
Associate's degree	1	1	1	0	128	4
Bachelor's degree	64	45	172	52	1,350	40
Master's degree	69	49	136	41	1,614	48
Specialist's degree	3	2	9	3	257	8
Doctorate degree	5	3	11	3	37	1
Years of Teaching Experience						
01-Sep	47	33	123	37	1,433	43
Oct-20	60	42	125	38	1,232	36
More than 20 Grades Taught	35	25	82	25	720	21
Elementary	61	43	169	51	1,726	51
Secondary	81	57	161	49	1,659	49

^aThe figures provided for the "Teaching Workforce" were located in a report published by the National Center for Education Statistics [25].

^bPercentages are rounded to the nearest number and therefore may not add to 100.

^cThese figures were not provided in the NCES report [25].

Data Analysis

Analysis and development of the DSPP was completed as part of the first phase. Internal consistency reliability analysis was performed using Cronbach's alpha on SPSS. Additionally, item and Rasch analysis was conducted on the Bond & Fox Winsteps software by analyzing the sample's responses to see if the survey items fit together. Based on these analyses, the most appropriate items were selected to be retained. Specifically, items were excluded from the final measures if they decreased the internal consistency reliability of their measure as observed through low item-total correlations or demonstrated exceptionally low Z-value statistics as determined through Rasch analysis. Following the additional data collection, further analysis on the DSPP was completed. For each scale, internal consistency reliability was again measured. Subsequently, Mplus software was used to conduct Exploratory Factor Analysis (EFA) in order to explore the factor structure of each scale. Data was then analyzed for each scale using Confirmatory Factor Analysis (CFA) on Mplus software to test the proposed factor models.

Results

Phase one

Internal consistency reliability: Internal consistency coefficients, or Cronbach's alpha, for each of the four experimental scales were calculated to determine the reliability of the measures and evaluate the consistency of results across items in each of the scales. George, et al. [26] provide the following rules for Cronbach's alpha scores: >0.9=Excellent, >0.8=Good, >0.7=Acceptable, >0.6=Questionable, >0.5=Poor, and <0.5 =Unacceptable. Item-total correlations, which measure the correlation between an item and the sum of the remaining items from a scale, were also analyzed to assess the construct validity of the measures. If a small item-total correlation is found, this suggests an item is not measuring the same construct as the remaining items and that if the item were removed, the scale's internal reliability would improve. For the analysis of the first scale, Part 1: Personal Attitudes, all sixteen items were examined, including items in the constructs of affective attitudes, cognitive attitudes, and behavioral intentions. The internal consistency reliability analysis revealed that the Cronbach's alpha of the first scale was acceptable ($\alpha=0.779$). Many items correlated well with the test as a whole; however, those from the construct of Cognitive Attitudes including items 6 to 10 had item-total correlations of only 0.187 to 0.310. If these items were all eliminated, the internal consistency coefficient would increase to 0.843, which is in the good range. The internal consistency reliability of each of the factors was also assessed. The affective attitudes construct, which included items 1 to 5, was found to have an acceptable internal consistency coefficient ($\alpha=0.757$). All items were well correlated with the factor as a whole. For the factor of cognitive attitudes, including items 6 to 10, the internal consistency coefficient was found to be 0.628, which is in the questionable range. Two items in this factor did not correlate well with the factor. These included items 6 and 7, with item-total correlations of 0.126 and 0.198, respectively. If these items were removed, the internal consistency of cognitive attitudes would increase to the acceptable range ($\alpha=0.770$). The internal consistency coefficient for the Behavioral Intentions construct, which included items 11 to 16, was found to be 0.775, which is in the acceptable range. All items were correlated well with the factor as a whole.

In the analysis of the second scale of the DSPP, Part 2: View of Public Perceptions, all twelve items were included. The internal consistency coefficient of this scale was excellent ($\alpha=0.912$). All items correlated well with this scale.

The seventy-two items of the third scale, Part 3: Attempts to address stigma, including the items in the subscales of use of practices, perceived effectiveness of practices, and likelihood to use if additional resources are provided, were included in the analysis of this scale. The Cronbach's alpha of this scale overall was excellent ($\alpha=0.963$). Most items correlated fairly well with the test as a whole, although items from the use of practices subscale tended to have lower item-total correlations. Item 11 had a particularly low item-total correlation of 0.167. The internal consistency would not change if this item were deleted. The use of practices subscale, which included items 1 to 24, was found to have an internal consistency coefficient of 0.884, which is in the good range. Again, most items were well correlated with the subscale, with the exception of item 11 ($\alpha=0.217$). However, if this item were deleted, the internal consistency of the subscale would not change. For the perceived effectiveness subscale, including items 25 to 48, the internal consistency coefficient was found to be 0.964, which is in the excellent range. All items correlated well with the subscale. The internal consistency coefficient for the likelihood to use if additional resources subscale, including items 49 to 72, was excellent ($\alpha=0.968$). All items were correlated well with the subscale as a whole.

For the analysis of the fourth scale, Part 4: Barriers to Addressing Stigma, the twenty-one items were examined. The Cronbach's alpha for the scale was excellent ($\alpha=0.941$). Most items showed acceptable item-total correlations, with the exceptions of items 1 ($\alpha=0.078$) and 7 ($\alpha=0.387$). The internal consistency would increase to .949 if these items were deleted.

Item and rasch analysis: Rasch analysis was used to determine if the items in the DSPP measured the intended constructs. The Infit Z-value statistic was utilized to determine the model fit for people and items. Z-values between -2.0 and 2.0 are acceptable [27]. For Part 1: Personal attitudes, fourteen of the sixteen items fell within the acceptable range. Items 4 and 5 were found to have Z-values of 2.8 and 2.7, respectively, and were therefore outside of the acceptable range. When examining Z-values for each factor of the first scale, it was determined that all items of the factor of Affective Attitudes fell in the acceptable range. In the factor of Cognitive Attitudes, items 6, 7, 8 and 10 fell outside of the acceptable range with Z-values of 4.5, 2.2, -2.5, and -4.3. For the factor of Behavioral Intentions, item 13 displayed a Z-value of 2.2 and was therefore not in



the acceptable range. Items 6 and 10 are most problematic, as they are farthest from the pathway.

Z-values for six of the twelve items on the second scale, Part 2: View of public perceptions, were in the acceptable range. Items 3, 6, 7, 8, 10 and 11 fell outside the acceptable range with Z-values ranging from 2.2 to 3.9. Items 8 and 10 fall farthest outside the pathway with Z-values of 3.9; therefore, it was considered whether these should be retained.

On the third scale, Part 3: Attempts to Address Stigma, Z-values for fifty-four of the seventy-two items fell in the acceptable range. The items that fell outside this range included 3, 5, 6, 10, 13, 14, 15, 18, 27, 28, 30, 33, 46, 48, 55, 57, 64 and 71 which had Z-values ranging from 2.1 to 6.9. In evaluating the Z-values for each factor, all items in use of practices fell in the acceptable range. In the factor of perceived effectiveness, eighteen of the twenty-four items fell in the acceptable range. Those outside the acceptable range included items 28, 35, 41, 43, 46 and 48 with Z-values from 2.3 to 8.2. For the factor of likelihood to use if additional resources, twenty-two of the twenty-four items were in the acceptable range; items 56 and 72 fell outside of this range with Z-values of -2.8 and 3.5. Overall, item 48 was farthest outside of the pathway, both for the scale and the Use of practices factor, and was considered for deletion.

For Part 4: Barriers to addressing stigma, eleven of the twenty-one items had Z-values that fell inside the acceptable range. Items 1, 3, 4, 5, 6, 7, 8, 12, 13 and 16 fell outside this range with Z-values ranging from 2.1 to 8.8. Items 1 and 7 were farthest from the pathway with respective Z-values of 8.8 and 5.7 and were considered for removal.

Item selection: Part 1: Personal attitudes had the lowest internal consistency reliability alpha level, although it was within the acceptable range. It was determined that items from the factor of cognitive attitudes, items 6 to 10, were most problematic. Three of these items, 6, 7, and 10 had particularly low item-total correlations. Additionally, the internal consistency of the cognitive attitudes factor was in the questionable range and two of the items, 6 and 7 had very low item-total correlations within the factor. The Rasch analysis on this factor demonstrated that the Z-values for four of the five items, 6, 7, 8, and 10 fell outside of the acceptable range. If the five items from this factor were deleted, it was determined that the cronbach's alpha for the scale would increase to the good range and the Z-values would be improved. Based on these findings, the five items in the cognitive attitudes factor on the first scale were removed.

The internal consistency reliability for Part 2: View of public perceptions was in the excellent range. There were a few items with Z-values outside of the acceptable range. However, the item-total correlations were all in the acceptable range and the internal consistency would decrease if any item were removed. Therefore, all items in this scale were retained.

For Part 3: Attempts to address stigma, the internal consistency reliability was found to be in the excellent range. The internal consistency reliability for the use of practices subscale was in the good range, and in the excellent range for the perceived effectiveness and likelihood to use if additional resources subscales. Although there was an item in the use of practices subscale with questionable item-total correlations with the scale and with this subscale, the internal consistency reliability would not change if this item were removed. The Rasch analysis revealed that the Z-values for the item "Other (please list)." In the subscales of perceived effectiveness and likelihood to use if additional resources fell farthest from the acceptable range. Based on the results from these analyses and the literature on empirically based practices to address stigma, it was determined that all of the listed items would be retained for the second phase of data collection as they could all yield valuable, unique information and were not a detriment to the internal consistency or summary statistics of the scale. However, it was decided that the items "Other (please list)." would be analyzed as separate, descriptive items as they did not fit the same pathways as the remaining items.

Part 4: Barriers to addressing stigma had an excellent internal consistency reliability alpha. Two items, 1 and 7, had notably lower item-total correlations than the others and their Z-values fell farthest outside the acceptable range, based on the Rasch analysis. If these items were removed, the internal consistency reliability would improve. Additionally, the face validity of these items as compared to the rest of the scale was questionable. Therefore, it was determined that items 1 and 7 would be removed from the scale for the second phase of data collection.

Phase two

Internal consistency reliability: The data collected in the second phase of surveys was analyzed to further examine the internal consistency reliability of each of the DSPP scales. Cronbach's alpha was again used to evaluate the consistency across items [26] and item-total correlations were analyzed to assess the construct validity of the measures. The eleven items of the first scale, Part 1: Personal Attitudes, including items in the constructs of affective attitudes and behavioral intentions were included in the analysis. The internal consistency reliability analysis revealed that the cronbach's alpha of the first scale was good ($\alpha=0.851$). Items correlated well with the test as a whole. For the affective attitudes construct, which included items 1 to 5, the internal consistency coefficient was found to be 0.704, which is in the acceptable range. The internal consistency coefficient for the behavioral intentions construct, which included items 6 to 11 (items 11 to 16 on the original measure), was 0.796, in the acceptable range. All items were correlated well with their respective factors.

The analysis of the second scale, Part 2: View of Public Perceptions, included all twelve items. The internal consistency coefficient of this scale was good ($\alpha=0.891$). All items correlated well with the scale.

For the third scale, Part 3: Attempts to address stigma, each of the 69 items in the subscales of use of practices, perceived effectiveness of practices, and likelihood to use if additional resources are provided, were included in the internal consistency reliability analysis. The Cronbach's alpha of this scale overall was 0.870, in the good range. Although the item-total correlations of some items were low for the scale as a whole, all items correlated well with their individual subscales. This is acceptable for this analysis, as the items would only be analyzed individually or summed into their subscales, rather than summed into the scale as a whole. The use of practices subscale, which included items 1 to 23, was found to have an internal consistency coefficient of 0.883, which is in the good range. The items were well correlated with the subscale. For the perceived effectiveness subscale, which included items 24 to 46 (previously items 25 to 47), the internal consistency coefficient was excellent ($\alpha=0.960$) and all items correlated well with the subscale. The internal consistency coefficient for the likelihood to use if additional resources subscale, which included items 46 to 69 (previously items 49 to 71), was found to be .960, in the excellent range. All items were correlated well with the subscale.

The nineteen items of the fourth scale, Part 4: Barriers to addressing stigma were included in the internal consistency reliability analysis. The Cronbach's alpha for the scale was excellent ($\alpha=0.953$). All items showed acceptable item-total correlations.

Factor analysis: To investigate the factor structure of each of the four experimental scales, EFA and then CFA were conducted on the Mplus software. EFA was conducted first in order to identify integral constructs in the scales, as research has shown that it is more effective in uncovering first-order factors than CFA [28,29]. Subsequently, CFA was employed to further test the factors identified in the EFA due to its strength in assessing emergent structural evidence [30,31].

For the EFA, full-information maximum likelihood estimation was used to uncover underlying factors [32]. To estimate communalities, the observed covariance matrix was examined [28]. The assumption was made that retained factors were correlated due to both theoretical and empirical reasoning. Accordingly, promax rotation was used [33]. Several rules were employed to decide the quantity of retained factors, including eigenvalues greater than 1.0 [34], scree test [35], Glorfeld's [36] parallel analysis [37], and Minimum Average Parcels (MAP) [38]. Although the most commonly used rule is that of eigenvalues, research shows that PA and MAP are the strongest methods for determining the number of factors to retain and that the scree test is valuable as well [36,38,39].

CFA was conducted to further evaluate the models as it provides a more rigorous inspection of the scales' factorial structures [30,31]. Full-information maximum likelihood estimation was employed to uncover underlying constructs [32]. Communalities were estimated from the observed covariance matrix [28]. As they were developed within different theoretical frameworks and focus on diverse aspects of fit, multiple measures of fit were considered including chi-square values, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Weighted Root Mean Square Residual (WRMR), Tucker-Lewis Index (TLI), and Comparative Fit Index (CFI) [40-42]. For chi-square values, non-significant values

or chi-square to degrees of freedom ratios of less than three to one suggest a good fitting model [43]. For RMSEA and SRMR, smaller values of 0.05 or less indicate a good fitting model and values up to 0.08 suggest reasonable errors of approximation [40]. For WRMR, values of 0.95 or higher indicate good fit [44]. TLI and CFI range between 0.00 and 1.0, with larger values reflecting better fit and values of .90 or greater providing evidence of good fitting models [41].

- a) **Factor analysis of part 1: Personal attitudes:** For the EFA of the first measure of the DSPP, part 1: personal attitudes, one-through six-factor solutions were rotated. Kaiser's eigenvalue criterion, the scree test, and PA all suggested that two factors be retained. MAP pointed to one factor. Both the one and two-factor solutions satisfied requirements for simple structure in that all variables showed appreciable factor loadings and each variable loaded on only one factor [45]. Based on these results, the two-factor solution was accepted. The two factors were interpreted according to the magnitude of their salient pattern coefficients within the rotated pattern matrix. All coefficients greater than or equal to 0.32 were considered appreciable. The first factor was characterized by high loadings on the first nine items in the scale and was named affective attitudes and negative behavioral intentions. The second factor was defined by appreciable loadings on the final two items and was named positive behavioral intentions. The correlation between the two retained factors was examined and the coefficient was moderate in magnitude ($r=0.52$), indicating association and common variance among factors. CFA was run to further examine the two-factor model suggested in the EFA results.

This model was compared to three alternative models: a single factor model, the theoretically proposed two factor model, in which items 1 to 5 made up the factor affective attitudes and items 6 to 11 made up the factor Behavioral Intentions, and a hierarchical two-factor model, based on the model suggested through the EFA. The measures of fit for each model are presented in Table 2. All indices most supported the retention of one of the two-factor models suggested by the EFA. As the criterion for these two models were similar, correlations between the factors were moderate in magnitude, and the presence of a higher order factor was theoretically based, the hierarchical two-factor model was accepted. Although these results differ slightly from theoretical expectations, the presence of a higher order factor, which was named Personal Attitudes, was predicted. Factor loadings linking item to respective factors were large (≥ 0.83).

Table 2: Measures of fit for part 1: Personal attitudes.

Fit Statistics	Models			
	One-Factor	Theorized Two-Factor	EFA Suggested Two-Factor	Hierarchical Two-Factor
χ^2	372.147*	365.200*	166.899*	166.899*
df	44	43	43	43
RMSEA	0.15	0.151	0.093	0.093
SRMR	0.078	0.079	0.053	0.053
CFI	0.777	0.781	0.916	0.916
TLI	0.721	0.72	0.892	0.892

Note: *Indicates χ^2 are statistically significant at $p < 0.001$.

- b) **Factor analysis of part 2: View of public perceptions:** EFA of the second measure, Part 2: View of public perceptions, was conducted to rotate one-through six-factor solutions. Kaiser's eigenvalue criterion and the scree test suggested that two factors be retained, whereas PA and MAP pointed to a one-factor model. As PA and MAP are known to be the most accurate criteria for determining the quantity of factors to retain, the one-factor model was accepted [36,38] (Verlicer et al. 1976). This solution satisfied requirements for simple structure. All coefficients were greater than or equal to 0.47 and considered appreciable. The sole factor, characterized by high loadings all items, was named view of public perceptions. CFA was run to further examine the one-factor model suggested in the EFA results. The measures of fit are presented in Table 3. Most of the criterion considered suggested the retention of the one-factor model suggested by the EFA, with the exceptions of chi-square value and RMSEA. These results aligned with the theoretical expectation that the items in this scale make up a single factor of view of public perceptions. Factor loadings were large (> 0.73).

Table 3: Measures of fit for part 2: View of public perceptions.

Fit Statistics	Model
	One-Factor
χ^2	203.123*
df	54
RMSEA	0.091
SRMR	0.049
CFI	0.904
TLI	0.833

Note: *Indicates χ^2 are statistically significant at $p < 0.001$.

- c) **Factor analysis of part 3: Attempts to address stigma:** For the EFA of the third DSPP scale, Part 3: Attempts to address stigma, one- through six-factor solutions were rotated. Kaiser's eigenvalue criterion, the scree test, PA, and MAP suggested that a multitude of factors be retained: fourteen, six, five, and nine, respectively. However, when the items were examined based on their salient pattern coefficients within the rotated pattern matrix, only the theoretically based three-factor solution satisfied the requirements for simple structure. Based on these findings and the interpretation of the theoretical framework, the three-factor solution was accepted. The three factors were interpreted according to the magnitude and meaning of their salient pattern coefficients within the rotated pattern matrix. All coefficients greater than or equal to 0.39 were considered appreciable. The first factor was characterized by high loadings on the first twenty-three items in the scale and was named Use of Practices. The second factor was defined by appreciable loadings on items 24 to 46 and was named Perceived Effectiveness. The third factor was characterized by appreciable loadings on items 47 to 69 and was named likelihood to use if additional resources. The correlations between the three retained factors were small in magnitude ($r=0.24, 0.19, 0.41$), suggesting little association with one another.

CFA was run to further examine the three-factor model suggested in the EFA results and by the theoretical framework. This model was compared to two alternative models: a single factor model and a hierarchical three-factor model. The measures of fit for each of the models are presented in Table 4. All of the indices most supported the retention of either of the three-factor models. As the values for these two models were similar and the correlation between the factors was small in magnitude, the first order three-factor model was accepted. This model aligned with the theoretical expectation that the first third of the items in the scale made up the factor of use of practices, the second third made up the factor of perceived effectiveness, and the final third made up the factor of likelihood to use if additional resources. Factor loadings linking each item to its respective factor were large (≥ 0.88).

Table 4: Measures of fit for part 3: Attempts to address stigma.

Fit Statistics	Models		
	One-Factor	Theoretical Three-Factor	Hierarchical Three-Factor
χ^2	6500.613*	3058.707*	3030.963*
df	2277	2274	2275
RMSEA	0.075	0.032	0.032
WRMR	2.458	1.234	1.235
CFI	0.701	0.944	0.946
TLI	0.692	0.943	0.944

Note: *Indicates χ^2 are statistically significant at $p < 0.001$.

- d) **Factor analysis of part 4: Barriers to addressing stigma:** EFA of the fourth measure, Part 4: Barriers to addressing stigma, was conducted to rotate one-through six-factor solutions. Kaiser's eigenvalue criterion and MAP suggested that three factors be retained. The scree test and PA pointed to two factors. Both the two and three-factor solutions satisfied requirements for simple structure, although coefficients were more appreciable in the three-factor model. Based on these results considered with the meaning of the items, the three-factor solution was accepted. The three factors were interpreted according to the magnitude and meaning of their salient pattern coefficients within the rotated



pattern matrix. All coefficients greater than or equal to 0.49 were considered appreciable. The first factor was characterized by high loadings on items 1 to 4 in the scale and was labeled lack of training. The second factor was defined by appreciable loadings on items 5 to 9 and was named lack of resources. The third factor was characterized by appreciable loadings on items 10 to 19 and was designated other barriers. The correlations between the three retained factors were moderate to large in magnitude (r=0.58, 0.59, 0.60). This finding indicates that the dimensions are associated and share common variance.

CFA was run to further examine the three-factor model suggested in the EFA results. This model was compared to two alternative models: A single factor model and a hierarchical three-factor model. The measures of fit for each of the models are presented in Table 5. All of the indices most supported the retention of one of the three-factor models suggested by the EFA. As the criterion for these two models were similar, the correlations between the factors were moderate to large in magnitude, and the presence of an overall factor was theoretically based, the hierarchical three-factor model was accepted. These results differ from theoretical expectations in that only one overarching factor, which was named Barriers, was predicted. Factor loadings linking each item to its respective factor were large (>0.90).

Table 5: Measures of fit for part 4: Barriers to addressing stigma.

Table with 4 columns: Fit Statistics, One-Factor, EFA Suggested Three-Factor, Hierarchical Three-Factor. Rows include Chi-squared, df, RMSEA, SRMR, CFI, and TLI.

Note: *Indicates chi-squared are statistically significant at p<0.001.

Limitations

There may be some limitations in this study such as the sample, data collection method, and factors included in the survey. Regarding the sample, this was gathered from an individual school district with demographics similar to those of the teaching workforce across the United States. However, the limited sample may impact the generalizability of the survey measure and study results to school districts in other states or countries. Additionally, the response rate may hinder the generalizability of the results. The low response rate could have stemmed from the dissemination of the survey through email. The use of survey measures to collect information on this topic may be a limitation in itself. Specifically, teachers may have under or over reported use of practices they had used to address stigma, as they were surveyed at only one point in time. The results may have been more factual if the data were collected either through observation or by asking teachers to maintain a running longitudinal record of their use of practices. The use of surveys may have also resulted in a social desirability response bias in that teachers may have over reported their use of practices to address stigma.

There are additional factors that would have been useful to explore through this survey. Previous research, for instance, has suggested that teachers' experience of in-service training on related topics may influence their attitudes and behavior [8,19,46]. Therefore, information on in-service trainings may have been helpful to collect. It may have also been beneficial to investigate school climate in the survey, as literature has suggested it may impact the academic achievement, behavior, and wellbeing of students in schools [47-50]. There may also be other factors that could be predictive of teachers' use of empirically-based practices to address stigma in the classroom that would be helpful to include in the survey measures.

Conclusion and Implications

The measures of teachers' personal attitudes, perceptions of public stigma, employment of empirically based practices to address stigma, and barriers to implementing these practices comprised in the DSPP that were developed and refined through this study carry implications for future research in the field and practices at school or district levels. Future researchers are both ensured of the existence of these constructs and provided with methods of measurement to sanction

an indeterminate quantity of further research in this area in order to more broadly increase acceptance of individuals with disabilities. At the school or district level, the constructs of personal attitudes, perceptions of public stigma, employment of empirically based practices to address stigma, and barriers to implementing these practices could be measured through the DSPP in order to identify teachers who may require support in addressing disability stigma in their classrooms. The DSPP may be easily disseminated to teachers either in print form or electronically and scored at individual or grade, school, or district levels. With these results, further implications may be drawn by administration, school psychologists, or other experts in the field at various levels, in terms of how support may be provided to teachers through methods such as professional development, consultation, or provision of resources. This would allow teachers to receive effective support to increase students' peer acceptance of individuals with disabilities through education, social contact, or advocacy based methods in order to affect long-term benefits for all students' attitudes and outcomes.

Appendix

Disability Stigma Perspectives and Practice (DSPP)

Part 1: Personal attitudes.

Table with 2 columns: Item number and Item description. Items 1-16 describe personal attitudes towards students with disabilities.

*These items were removed following the first phase.

Part 2: View of public perceptions.

Table with 2 columns: Item number and Item description. Items 1-12 describe public perceptions of children with disabilities.

Note: All items were retained following the first phase.



Part 3: Attempts to address stigma.

Table with 2 columns: Item number and description of attempt to address stigma.

Part 5: Demographics.

Table with 2 columns: Item number and demographic question.

Note: For each item, the first item number represents the answer to the question of "Used in past year?" (items 1 to 24), the second "Do you think this is/would be effective?" (items 25 to 48) and the third "Would use if more time, training, resources, etc.?" (items 49 to 72).

*These items were removed from analysis with the scale following the first phase.

Part 4: Barriers to addressing stigma.

Table with 2 columns: Item number and description of barrier to addressing stigma.

Note: *These items were removed following the first phase.

References

List of 12 references related to disability stigma and education.



13. Ludwikowski WM, Vogel D, Armstrong PI (2009) Attitudes toward career counseling: The role of public and self-stigma. *Journal of Counseling Psychology* 56(3): 408-416.
14. Campbell JM, Ferguson JE, Herzinger CV, Jackson JN, Marino CA (2004) Combined descriptive and explanatory information improves peers' perceptions of autism. *Research in Developmental Disabilities* 25(4): 321-339.
15. Ohan JL, Visser TA, Moss RG, Allen NB (2013) Parents' stigmatizing attitudes toward psychiatric labels for ADHD and depression. *Psychiatric Services* 64(12): 1270-1273.
16. Rosenbaum PL, Armstrong RW, King SM (1987) Parental attitudes toward children with handicaps: New perspectives with a new measure. *Developmental and Behavioral Pediatrics* 8(6): 327-334.
17. Vogel DL, Bitman RL, Hammer JH, Wade NG (2013) Is stigma internalized? The longitudinal impact of public stigma on self-stigma. *Journal of Counseling Psychology* 60(2): 311-316.
18. Sirey JA, Bruce ML, Alexopoulos GS, Perlick DA, Raue P, et al. (2001) Perceived stigma as a predictor of treatment discontinuation in young and older outpatients with depression. *American Journal of Psychiatry* 158(3): 479-481.
19. Bell L, Long S, Garvan C, Bussing R (2011) The impact of teacher credentials on ADHD stigma perceptions. *Psychology in the Schools* 48(2): 184-197.
20. Kellison I, Bussing R, Bell L, Garvan C (2010) Assessment of stigma associated with attention deficit hyperactivity disorder: Psychometric evaluation of the ADHD stigma questionnaire. *Psychiatry Research* 178(2): 363-369.
21. McAuliffe MD, Hubbard JA, Romano LJ (2009) The role of teacher cognition and behavior in children's peer relations. *Journal of Abnormal Child Psychology* 37(5): 665-677.
22. McGoey KE, Rispoli KM, Venesky LG, Schaffner KF, McGuirk L, et al. (2014) A preliminary investigation into teacher perceptions of the barriers to behavior intervention implementation. *Journal of Applied School Psychology* 30(4): 375-390.
23. Cohen J (1988) *Statistical power analysis for the behavioral sciences*. In: 2nd (Edn.), Hillsdale, Lawrence Erlbaum Associates, Inc., USA, pp. 1-579.
24. Faul F, Erdfelder E, Buchner A, Lang AG (2009) *Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses*. *Behavior Research Methods* 41: 1149-1160.
25. National Center for Education Statistics (2016) *Digest of Education Statistics, 2015*. USA, pp. 1-970.
26. George D, Mallery P (2005) *SPSS for windows step by step: A simple guide and reference*. Social Sciences, 11th (Edn.), Pearson Publishers, England, pp. 1-408.
27. Bond TG, Fox CM (2015) *Applying the rasch model: Fundamental measurement in the human sciences*. 3rd (Edn.), Lawrence Erlbaum Associates, USA.
28. Gorsuch RL (2003) *Factor analysis*. In: Shinka JA, Velicer F (Eds.), *Handbook of psychology: Research methods in psychology*. John Wiley, USA, 2: 143-164.
29. MacCullum R, Roznowski M, Necowitz LB (1992) Model modifications in covariance structure analysis: The problem of capitalizations on chance. *Psychological Bulletin* 111(3): 490-504.
30. Bryne BM (2012) *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. 1st (Edn.), Routledge Publishers, USA, pp. 1-432.
31. Schmitt TA (2011) Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment* 29(4): 304-321.
32. Jöreskog KG (1977) *Factor analysis by least-squares and maximum likelihood methods*. In: Enslein K, Ralston A, Willf HF (Eds.), *Statistical methods for digital computers*. John Wiley & Sons, USA, pp. 125-153.
33. Tatroyn DJ, Wood JM, Gorsuch RL (1999) Setting the value of k in promax: A monte carlo study. *Educational and Psychological Measurement* 59(3): 384-391.
34. Kaiser HF (1960) The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20: 141-151.
35. Cattell RB (1966) The scree test for the number of factors. *Multivariate Behavioral Research* 1(2): 245-276.
36. Glorfeld LW (1995) An improvement on horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement* 55(3): 377-393.
37. Horn J (1965) A rational and test for the number of factors in factor analysis. *Psychometrika* 30: 179-185.
38. Buja A, Eyuboglu N (1992) Remarks on parallel analysis. *Multivariate Behavioral Research* 27(4): 509-540.
39. Velicer WF (1976) Determining the number of components from the matrix of partial correlations. *Psychometrika* 41: 321-327.
40. Browne M, Cudeck R (1993) Alternative ways of assessing model fit. In: Bollen K, Long J (Eds.), *Testing structural equation models*. Sage Publishers, USA, pp. 136-162.
41. Hu L, Bentler PM (1995) Evaluating model fit. In: Hoyle RH (Ed.), *Structural equation modeling: Concepts, issues, and applications*. Sage Publishers, USA, pp. 76-99.
42. Tucker LR, Lewis C (1973) A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 38: 1-10.
43. Kline RB (2010) *Principles and practice of structural equation modeling*. 3rd (Edn.), Guilford Press, USA.
44. Yu CY (2002) Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. pp. 1-183.
45. Thurstone LL (1947) *Multiple-factor analysis*. Chicago University Press, UK.
46. Park M, Chitiyo M (2011) An examination of teacher attitudes towards children with autism. *Journal of Research in Special Educational Needs* 11(1): 70-78.
47. Bear GG (2010) *School discipline and self-discipline: A practical guide to promoting prosocial student behavior*. Guilford Press, USA, pp. 1-256.
48. Bear GG, Witcomb SA, Elias MJ, Blank JC (2015) SEL and schoolwide positive behavioral interventions and supports. In: Durlak J, Gullotta T, Domitrovich C, Goren P, Weissberg R (Eds.), *Handbook of Social and Emotional Learning*. Guilford Press, USA, pp. 453-467.
49. Emmer ET, Sabornie EJ (Eds.), (2015) *Handbook of classroom management*. 2nd (Edn.), Routledge Publishers, USA, pp. 1-616.
50. Harrison P, Thomas A (2014) *Best practices in school psychology*. National Association of School Psychologists, Bethesda, USA, 4.