IMPLICIT AND EXPLICIT ATTITUDES OF EDUCATORS TOWARD THE

EMOTIONAL DISTURBANCE LABEL

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF PHILOSOPHY

BY

JAMES PATRICK JONES

DISSERTATION ADVISOR: DR. BARBARA ROTHLISBERG

BALL STATE UNIVERSITY

MUNCIE, INDIANA

NOVEMBER 2009
ACKNOWLEDGEMENTS

Looking back at this labor of several years, there are many people to thank. I extend my gratitude to Dr. Barbara Rothlisberg, my committee chair, and in so many ways my prototype of the ideal professor-mentor: disciplined, exacting, and deeply knowledgeable, yet encouraging, invested, and accessible. Her guidance through this process served to bridge the many starts and stops - the potential momentum killers – and to reassure me that this goal was within reach.

I thank Dr. Sharon Paulson, whose constancy, resourcefulness, and supportiveness, were critical elements in accelerating my pace and helping me to get past bumps in the road. As a member of my committee, her strong commitment to this project was very much appreciated. To Dr. Anthony Mahon and Dr. Jerry Ulman, who completed my committee, I give thanks for your positive influences and constructive feedback. Your input helped me refine my research.

There are many family members who deserve credit for sacrificing time and energy on my behalf. Foremost, I am grateful for my wife. The importance of her willingness to allow me countless hours to work on my dissertation cannot be overstated. She cheered me on and never questioned whether this would happen, even when it was logical to ask, “What? There’s more?” I can say unequivocally that God blessed the broken road that led me to her.

My parents have been there when I’ve faced my greatest personal challenges. They have always painted a picture of how good life could be for me, and I will be forever grateful for their unwavering support and their many displays of love.
Two of my children preceded my dissertation efforts and one joined my family midstream. Each of them inspired me to carry on through challenges and setbacks. To leave this work incomplete was a legacy I could never bear to leave them.

The Ball State School Psychology program introduced me to many individuals who have been key to my progression. Thanks to Dr. Frank Sparzo, Dr. Andrew Davis, and Dr. Betty Gridley for being among my most instrumental influences in the program. I thank Dr. Cheryl Peresie for being a great supervisor and for showing strong interest in my development. The administrative staff for our department, Linda Jones and Penny Nowaczyk, are exemplary individuals who have assisted me whenever I had a need.

I am grateful for a predoctoral internship that set a very high standard. Dr. Eric Schwartz, Dr. Katherine Wardi, and Dr. Kim Quirk were superb supervisors whose encouragement and kindness helped to spur me on. My co-interns, Jeremy Vose and Nora Geraghty, made that year one of the most meaningful and enriching of my life.

One of the greatest obstacles of this project was the recruitment of participants. I thank my director, Lynn Gosser, for enabling me to conduct this research in her schools. My gratitude goes out to all the directors, psychologists, and staff within our school cooperative who expressed interest in my progress, and in so doing, helped me maintain my focus. To all of the teachers who gave of their time to take part in my study, I am deeply indebted to you.

In subtle but significant ways, my friends and extended family contributed to this endeavor as well. Thanks to you all.
# TABLE OF CONTENTS

Acknowledgments............................................................................................................... ii

Table of Contents............................................................................................................... iv

List of Tables ..................................................................................................................... vi

List of Figures ................................................................................................................... vii

Abstract.............................................................................................................................. ix

Introduction..........................................................................................................................1

Review of Literature ............................................................................................................16

  History of Measuring Attitudes toward Groups with Perceived Differences........ 17

  History of the Emotional Disturbance Label.......................................................... 36

  Attitudinal Research toward the Emotional Disturbance Label.......................... 52

  Theories of Labeling and Expectancy Effects.......................................................61

Method...............................................................................................................................72

  Participants ............................................................................................................... 72

  Measures ................................................................................................................ 73

  Procedure ............................................................................................................... 83

Results...............................................................................................................................85

  Question 1a ........................................................................................................... 85

  Question 1b ........................................................................................................... 91

  Question 1c ........................................................................................................... 91

  Question 2 ............................................................................................................ 92

  Question 3 ............................................................................................................ 94
Discussion ..........................................................................................................................95

Findings..................................................................................................................................95

Limitations..............................................................................................................................100

Further Research....................................................................................................................102

References..............................................................................................................................105

Appendix A: Research Protocol..............................................................................................126

Appendix B: Standard Recruitment Email/Letter.................................................................142

Appendix C: Consent Form ....................................................................................................143
LIST OF TABLES

Table 1. *Frequencies of Individual Characteristics within the Sample* ..................72

Table 2. *Mean Scores of Educators on Explicit and Implicit Attitude Measures* ........86

Table 3. *Summary of Intercorrelations for Scores on the SADP, SDS ED, SDS LD, and IAT for Entire Sample* ........................................................................................................87

Table 4. *IAT Difference Scores by Order of Presentation* .....................................89

Table 5. *Summary of Intercorrelations, Means, and Standard Deviations for Scores on the SADP, SDS ED, SDS LD, and IAT as a Function of Occupation* .................................93
LIST OF FIGURES

Figure 1. Demographics Page.................................................................126

Figure 2. Implicit Association Test – Instruction Screen .........................127

Figure 3. Target-Concept Discrimination – Sample Screen .....................127

Figure 4. Attribute Instructions .................................................................128

Figure 5. Associated Attribute Discrimination – Sample Screen ............128

Figure 6. Combined Task Instructions ......................................................129

Figure 7. Initial Combined Task – Sample Screen ................................129

Figure 8. Combined Task Instruction Screen – Testing Condition ..........130

Figure 9. Combined Task – Sample Screen ..............................................130

Figure 10. Reversed Target Concept - Instruction Screen ......................131

Figure 11. Reversed Target Concept Discrimination – Sample Screen ....131

Figure 12. Reversed Combined Task – Instruction Screen ......................132

Figure 13. Reversed Combined Task Practice – Sample Screen ..........132

Figure 14. Reversed Combined Task Test Condition – Instruction Screen .133

Figure 15. Reversed Combined Task Test Condition – Sample Screen ....133

Figure 16. IAT Completion Screen ..........................................................134

Figure 17. SADP Screen 1........................................................................135

Figure 18. SADP Screen 2........................................................................136

Figure 19. SADP Screen 3........................................................................137
Figure 20. SADP Screen 4.................................................................138
Figure 21. SADP Screen 5.................................................................139
Figure 22. Social Distance Scale......................................................140
Figure 23. Sample Drop-Down Menu.............................................140
Figure 24. Final Screen.................................................................141
Figure 25. Screen Lock.................................................................141
ABSTRACT

DISSEPTION: Implicit and Explicit Attitudes of Educators toward the Emotional Disturbance Label

STUDENT: James Patrick Jones

DEGREE: Doctorate of Philosophy

COLLEGE: Teachers College

DATE: November 2009

PAGES: 153

This study examined implicit and explicit attitudes of teachers toward the Emotional Disturbance (ED) label, the strength of association between implicit and explicit ratings, and the variance in attitudes between different types of teachers or among teachers in different settings. Ninety-eight teachers (52 regular education and 46 special education teachers), from a mix of urban and rural school districts in central Indiana, completed three computer-based attitudinal measures: an implicit association test (IAT), a Social Distance Scale (SDS), and the Scale of Attitudes toward Disabled Persons (SADP). Results indicated that teachers possess a significant preference for the category of Learning Disability (LD) compared to ED, and the correspondence between implicit and explicit attitudes toward those two labels was strong. This data supports the notion that the ED label is perceived far too negatively by educators to serve as an effective category of services for children with serious mental health issues.
Implicit and Explicit Attitudes of Educators toward the Emotional Disturbance Label

Of the 13 categories of educational disability presently identified in special education law, the emotional disturbance (ED) category has long been a subject of criticism due in part to definitional inconsistencies and significantly lower than projected rates of identification. The disparity of ED classification rates among the 50 states demonstrates inconsistencies in qualifying students for the category or perhaps reluctance to apply the ED label to students. Rates of identification range from a low of 0.11% in Arkansas to 1.5% in Vermont (U.S. Department of Education, 2005). Boundaries between ED and other categories, such as learning disability that has similarly uneven prevalence rates, have been thought to be obscured by comorbid features and a lack of distinctive criteria (Hallahan & Kauffman, 1977). Studies suggest that characteristics of an emotional disturbance can be observed within multiple special education categories (Cullinan, 2004), resulting in inappropriate placements for these children in programs for children with learning disabilities (Lopez, Forness, MacMillan, Bocian, & Gresham, 1996).

Although the percentage of students classified as ED in U.S. schools was estimated to average 0.73% in 2003 (U.S. Department of Education, 2005), experts have indicated that the actual prevalence of students meeting the criteria for ED is likely much
higher, perhaps in the range of 3 to 6 percent of school-age children (Friedman, Kutash, & Duchnowski, 1996; Kauffman, 2001). If one considers the projected numbers of students with conduct disorders and/or those in need of comprehensive mental health services, the number swells to approximately 20% of the school-age population (Hoagwood & Erwin, 1996). Walker, Nishioka, Zeller, Severson, and Feil (2000) contended that schools greatly underserve students with significant behavioral and emotional problems.

Some fluctuation of prevalence rates might be expected from state to state. One sees, however, that rates of identification are fairly consistent in categories that display more overt and readily diagnosable features of disability such as Hearing Impairments and Visual Impairments. This lends support to the notion that there are factors other than naturally-occurring variations among students that account for discrepancies in the ED prevalence rates. Efforts to uncover the sources of variation among these rates have found that roughly one-third are due to differences in states’ definitions, leaving two-thirds unexplained (Wright, Pillard, & Cleven, 1990). These findings reinforce the need to examine other potential sources of disagreement to see what factors could be standing in the way of identifying students.

In addition to concerns about inconsistent identification of ED, there are issues about the timing of identification. Students who are classified ED are identified much later on average than for any other category of special education services, with referrals peaking at the ages of 14- to 15-years (Walker et al., 2000). This delay in identification and servicing is all the more troublesome when considering the perspective some hold that conduct disorders and antisocial behaviors, which are sometimes characteristic of
students with ED, may become relatively fixed, chronic conditions after the age of eight (Kazdin, 1987).

Duncan, Forness, and Hartsough (1995) indicated that most students who later qualified for ED services were known to have serious problems years prior to identification. The research on the early identification of students who are classified as emotionally disturbed suggests that the sooner intervention occurs, the greater the likelihood that long-term negative consequences can be ameliorated (Conroy, Hendrickson, & Hester, 2004). If children with severe behavioral and emotional problems are not assessed and subsequently identified as such, then they are much more likely to be strung along in the status quo of general education and receive more generalized, normative attempts at intervention rather than idiographic interventions made possible in a more specialized setting. There is some evidence that severe behavioral problems can be dealt with better within self-contained classrooms (Schneider & Leroux, 1994). If early identification of ED is seen as desirable, as most would contend, then it may be helpful to obtain a better understanding of what factors play a role in the under identification and/or delayed identification of these youth.

Negative outcomes for students with ED are well-established. Wagner (1991) indicated that students receiving ED services drop out of school at significantly higher rates compared to students of other disability groups. Survey data shows that the prevalence of ED as a primary disability is eight times higher in detention and correctional facilities than in general school-age populations (Quinn, Rutherford, Leone, Osher, & Poirier, 2005). Considering the potential for individual loss and societal harm
brought on by antisocial behavior, there appears to be a strong case for further ascertaining what factors account for difficulties with service provision.

A logical concern that follows the assumption of under identification of students with severe emotional and behavioral problems is that there is a need to account for the shortfall. Kauffman, Brigham, and Mock (2004) suggested that there are three primary reasons why school personnel fail to identify students with emotional/behavioral disorders. These reasons include (a) personal philosophy, (b) definitional imprecision, and (c) pragmatic concerns. It is fair to consider that these reasons may not be exclusive of one another, that is, any combination of the three may contribute to decisions about whether to classify a student as ED.

Personal philosophies that run counter to the identification of students with all manner of disabilities are seen in movements such as the Regular Education Initiative, that has called for the elimination of special education labels altogether (Wang & Walberg, 1988). Philosophical debates about the merits of full inclusion have been prominent in the literature with strong voices in support of such an educational approach (Stainback, Stainback, & Ayres, 1996) and others promoting a continuum of placements (Kauffman, McGee, & Brigham, 2005). The Full Inclusion Movement advocates for the abandonment of categorical labels, with the aim of integrating students with all disabilities into classrooms with students who are non-disabled. Concerns about the appropriateness of identifying students are not specific to the ED label; however, ED as a category is beset by definitional problems, making it perhaps easier to deny services to a student on philosophical grounds, by using the justification that the nebulous criteria for ED have not been met.
Definitional imprecision has plagued the ED label since its introduction as the serious emotional disturbance (SED) category with P.L. 94-142 in 1975. Forness and Kavale (1997) stated, “In special education, the field of emotional or behavioral disorders seems to have considerably more problems in definition than other major categories…” (p. 45). Owing in part to imprecise definitions and a lack of distinction among the ED, LD, and mental retardation (MR) categories, Hallahan and Kauffman (1977) argued for the abandonment of a categorical framework and the adoption of a behavioral one. Unfortunately, the definition for ED has remained relatively unchanged since its introduction, having been taken from statements made by Eli Bower (Bower, 1960). Forness and Knitzer (1992) acknowledged the value of Bower’s work to the early work-up of the definition but cited the importance of modernizing the federal definition.

The current definition for emotional disturbance reads as follows:

"(i) The term means a condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree that adversely affects a child’s educational performance: (A) An inability to learn that cannot be explained by intellectual, sensory, or health factors; (B) An inability to build or maintain satisfactory interpersonal relationships with peers and teachers; (C) Inappropriate types of behavior or feelings under normal circumstances; (D) A general pervasive mood of unhappiness or depression; (E) A tendency to develop physical symptoms or fears associated with personal or school problems."
(ii) The term includes schizophrenia. The term does not apply to children who are socially maladjusted, unless it is determined that they have an emotional disturbance. (U.S. Department of Education, 1998, p. II-46).

Cullinan (2004) cited problems with several non-operationalized statements within the definition, including the qualifiers “marked degree” and “extended period of time.” Neither of those phrases allow for a strict delineation of severity and duration of students’ behaviors. Kidder-Ashley, Deni, Azar, and Anderton (2000) reported that few states’ guidelines specify what those terms mean, which makes it understandably difficult to assess uniformly whether a student qualifies for ED services. Perhaps the most perplexing and derided portion of the federal definition is the phrase, “does not include children who are socially maladjusted, unless it is determined that they have an emotional disturbance.” The inclusion of the social maladjustment clause has rendered the ED classification troublesome since its inception (Kauffman, 2001). It provides grounds to overlook students who present with externalizing problems. This amounts essentially to splitting hairs over two very interrelated terms (emotional disturbance and social maladjustment) whose boundaries are difficult to differentiate. Merrell and Walker (2004) observed that social maladjustment has never been defined in the federal definition, nor has there been any prescription given as to how to discriminate between emotional disturbance and social maladjustment. To the question of whether indeed the two conditions should be differentiated in assessment for special education qualification, they responded “no,” or at best “almost never” (p. 902). In spite of confusion surrounding the social maladjustment clause, Forness (1992) noted that survey data has shown that
40% of states do not exclude children with social maladjustment from the ED category. Therefore, a student who presents with traditional behavioral problems appears to receive more or less consideration in classification decisions, depending on the state in which the student lives.

Furthermore, there appears to be disagreement about the adequacy of the term emotional disturbance as a disability category. In a study that surveyed 41 state education agencies, Kidder-Ashley et al. (2000) found that state guidelines are at odds as to whether the name of the ED category should imply exclusively emotional, exclusively behavioral, or a combination of behavioral and emotional problems. Of the 41 states that responded to the researchers’ survey, 61% used a term that suggests only an emotional aspect of the disability (e.g., emotional disturbance); 15% used a term that is primarily behavioral in nature (e.g., behavioral disorders); and the remaining 24% referred to both emotional and behavioral problems (e.g., emotional and behavioral disorder or EBD) in their terminology. [Note: EBD is the term used by the Council for Children with Behavior Disorders (CCBD), a division of the Council for Exceptional Children (CEC).] This multiplicity of labels has resulted in confusion both in assessing ED and in determining placement for the category.

Significant efforts have been made to shore up extant problems with the definition of ED, including proposing a new definition as well as modifying the category name to reflect externalizing problems. Cullinan (2004) described how the National Mental Health and Special Education Coalition created an alternative ED definition; however, the definition remains unadopted for federal legislation. A long-standing issue, related to the social maladjustment clause of the federal definition of ED, is that children with
conduct disorders and severe behavioral problems are unjustly excluded from receiving ED services. Boucher (1999) discussed a movement to expand the terminology from ED to Emotional or Behavioral Disorders (EBD). EBD reflects an understanding that emotional and behavioral problems are often comorbid, and that a focus on one, at the exclusion of the other, may be unfair and limit services for students.

In contrast to the other sources of under identification that have been discussed and debated extensively of late, including the value of inclusive education and the need to refine the definition for emotional disturbance, the belief that ED is an overly-stigmatizing label has been a relatively unchallenged notion for many years. Kauffman and Landrum (2006) stated, “One sure way to thwart prevention is to deny the student’s deviance” (p. 126). Fear of labeling a student as ED may be one major reason that preventive efforts are not conducted satisfactorily on behalf of students who present with severe emotional and behavioral problems. Kauffman (2003) confirmed, “There is hesitancy to anticipate disabilities, particularly cognitive or behavioral disabilities, partly due to the superstition that expectations cause disabilities” (p. 196-197). This “superstition” that Kauffman described may play a part in the reluctance to identify students as ED. The power of teacher expectations to shape students’ performance has a lengthy, albeit contested history in the literature (Rosenthal & Jacobson, 1966), which has been one reason for persistent views that stigma should be avoided at all costs.

Kauffman et al. (2004) indicated, “Some professionals fail to identify students with EBD because of a personal distaste for labeling” (p. 18). A tendency of some educators to avoid the ED label simply on the grounds that it might be stigmatizing begs the question of which is more important, the need for educational services that are
appropriate for a student’s needs or the need to shelter a student from perceived labeling effects. Within that question lie two concerns of practical interest. First, does the ED label carry with it an inherently negative association relative to other disability categories (e.g., with LD)? Second, if there is an inherently negative view of the ED label, how does this affect students who qualify for ED services? This study seeks to examine the first question and determine whether the ED label carries with it a significant stigma among educators.

Evidence of negative perceptions toward students with ED traces back to research conducted in the 1970s and early 1980s. Wood, Bloomquist, and Chalmers (1992) surmised that studies using attitude surveys and experimental simulations showed negative valences toward students labeled “emotionally disturbed” or “behaviorally disordered” (Algozzine, Mercer, & Countermine, 1977; Boucher & Deno, 1979; Carroll & Repucci, 1978; Foster, Algozzine, & Ysseldyke, 1980; Stevens-Long, 1973; Ysseldyke & Foster, 1978). Hypotheses concerning negative attitudes toward students with ED pointed primarily to the power of the label itself to mediate educators’ tolerance of certain behaviors.

Interestingly, Ysseldyke and Foster (1978) found that teacher perceptions of behaviors of “normal” students, who were labeled LD and ED for purposes of the study, were the same. On the other hand, Algozzine et al. (1977) observed that teachers found disruptive, aggressive behavior less acceptable when attributed to a student labeled LD as compared to a student labeled ED. This provides evidence that, stigmatizing or not, the ED label may serve to moderate teachers’ attitudes toward externalizing behaviors at least relative to students labeled LD. Therefore, if a placement decision is going to be
made, there may be some justification for classifying a student as ED because problematic behaviors are more likely to be tolerated or understood as being a characteristic of the disability.

Additional justifications for performing a contemporary analysis of the ED stigma among educators include social changes over time and the introduction of more refined methods of assessing attitudes. In the intervening years since the early attitudinal surveys and experiments, perspectives on disability have changed as evidenced by movements to include children with disabilities with their nondisabled peers. What is known about disabilities generally, and emotional and behavioral disorders specifically, has expanded through the accumulated body of research on the subject. In discussing the history of scientific study of Emotional and Behavioral Disorders, Kauffman et al. (2004) noted, “the gradual accretion of scientific understanding through research has led us from nearly total ignorance to the beginnings of understanding what behavioral disorders are, why they occur, and how we can best treat them” (p. 16).

Despite the fact that people may claim that students with ED are stigmatized disproportionately, the reality may be that there has been some normalization of attitudes vis-à-vis the term emotional disturbance in much the same way that conditions such as depression have become more socially accepted. It would be helpful for educators to know whether such a normalization has occurred, so that arguments involving stigma could play more or less of a role in discussions of appropriate services for students. Whether discussions of labeling effects are appropriate in determining special education eligibility decisions is a legalistic subject beyond the scope of the current inquiry.
Attitudinal research now offers means beyond explicit self-report measures (e.g., questionnaires and surveys) for evaluating positive or negative perceptions. Historically, measurement of attitudes has been subject to the biases inherent in self-reporting. When individuals are given time to think about a response, they may tend to consider or be swayed by how the response may be perceived by others (Fazio, Snbonmatsu, Powell, & Kardes, 1986). This is a known issue observed not only in attitudinal research but also in the use of self-rating scales for assessment of children and adults. Elliott and Busse (2004) acknowledged that the validity of such scales can be affected by “response bias factors,” sometimes manifested by acquiescence, faking, or seeking social desirability. Therefore, there may be uncertainty as to the reliability of attitudinal data that relies on explicit measures. One must wonder whether such explicit measures truly capture data that approximates the desired construct or whether they merely reflect what a respondent wants others to think. Also, responses to an explicit measure may reflect an individual’s knowledge of a stereotype as opposed to a privately held opinion.

More sophisticated attempts have been made to glean attitudinal data, such as comparing teacher and teacher-in-training ratings of the behaviors of students who were labeled randomly as either normal or emotionally disturbed. Foster et al. (1980) performed such a study and found that teachers were more likely to be influenced by the biasing label than teachers-in-training. Does this imply that greater familiarity with a label leads to stronger negative perceptions? Or does it suggest that those who possess greater expertise may be compelled to over interpret behaviors based on a past history with the label? This is not clear but illustrates the difficulty in assuming that negative
perceptions estimated by explicit attitudinal measures are the driving force behind labeling bias.

Skepticism about the utility of explicit measures of attitude gave rise to interest in discovering ways to capture implicit attitudes. Greenwald and Banaji (1995) described implicit attitudes as “introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects” (p. 8). In the social psychological research, a tool referred to as an Implicit Association Test (IAT) has been developed for the purpose of capturing automatized attitudinal feedback (Greenwald, McGhee, & Schwartz, 1998). An IAT is a mechanism through which the latency of individuals’ button-pressing responses is recorded when they are asked to pair (associate) a word or picture with either a negative or positive term (e.g., “good” and “bad”). For example, Greenwald and Nosek (2001) illustrated how male-math and female-arts might be paired, and then counterstereotypic pairings of female-math and male-arts could be requested. The assumption is that individuals who possess a negative bias toward a pairing will take longer to associate terms that run counter to their thinking. The initial research on the IAT (Greenwald et al., 1998) found that it could reliably detect differences in concept associations that are almost universal in the population, such as a preference for flowers over insects. The application of the IAT has been extended logically to examine implicit attitudes about weightier social issues including racial prejudice (Wittenbrink, Judd, & Park, 1997) and smoking (Swanson, Rudman, & Greenwald, 2001).

This study will mark the first time that efforts have been made to examine teachers’ implicit attitudes toward special education labels. Collecting data with explicit
as well as implicit attitudinal measures is a standard feature of studies involving an IAT, because comparing results of the two measures can point to potential differences in attitudes that individuals are willing to share overtly versus what appear to be privately held beliefs. Also, given that explicit attitudes have been the exclusive subject of past research in stigma and bias in special education labeling, it becomes helpful, if not essential, to use an explicit measure in the current study in order to draw inferences about changes in attitudes from past research and to compare with attitudes toward the LD classification.

Teachers, both general educators and special educators, are the targeted participants for this study. This participant selection is due to practical as well as historical considerations. Practically speaking, educators are the ones who are familiar with special education terminology and inform parents of what terms such as emotional disturbance mean. Additionally, teachers are the ones who would most likely be affected by the stigma of a special education label because the specific classification for a student’s services typically is not made known to students’ peers. Also, because of the teacher’s central role in managing the quality of educational services for students, there has been traditionally little exploration into the effects of categorical labels on parents’ and children’s attitudes (Palmer, 1983).

Previous research demonstrates a clear need to intervene earlier and provide greater assistance to students who meet the criteria for ED. Barriers to identification of ED include philosophical differences about the value of special education services and disagreements about the federal definition. Those barriers have been, and continue to be, vigorously discussed and debated. However, a third obstacle, concerns about an unduly
stigmatizing effect of the ED label, provides a viable and important subject for inquiry. Explicit attitudes toward the ED label have been studied in the distant past, but the advent of implicit measures of attitude provides potential for even greater understanding of how school personnel feel about this category of special education services. This study will explore the veracity of long-held beliefs about the stigma of ED. An emotional disturbance is a term defined by special education law and has no social meaning or value for services rendered outside of schools; thus, school personnel who have an impact on educational services for students will be the participants of interest.

This study will examine four research questions: Do educators hold negative perceptions toward the ED label compared to the LD label?, Do general attitudes of educators toward individuals with disabilities correspond to their attitudes toward individuals with the ED label? Are there significant differences between explicit and implicit attitudes of regular education versus special education teachers regarding the ED label?, and Are there significant differences between explicit and implicit attitudes of urban compared to rural teachers regarding the ED label? By exploring these questions, there will be the potential to strengthen or weaken the argument that there is an inherent stigma among educators toward the ED label. Examining biases toward the ED label may serve to lend greater clarity to school decisions about whether to risk deleterious effects of labeling a child as opposed to intervening with appropriate services for a student who meets the criteria for ED. Additionally, this study may bolster the efforts of advocates who have unsuccessfully petitioned legislators for several decades to modify the eligibility criteria for, and category name of, emotional disturbance.
A between group comparison of attitudes among special educators and general educators may offer important insights about where negative attitudes may be concentrated. Whether or not the notion of a harmful stigma is confirmed, this research has the potential either to identify the need to manage stigma more effectively or to stop using it as a reason for refusing to refer or identify students who would likely benefit from ED services. If evidence is found that the ED label carries no more negative association than LD, an oft-used proxy for ED (Gresham, 1985; Lopez et al., 1996), these findings could remove a significant barrier to the identification of some students who might otherwise qualify for ED services.

By looking at implicit versus explicit attitudes, inferences can be made ostensibly about how attitudes toward ED have changed over time and whether overt measures of attitude are closely aligned with more deep-seated beliefs. If educators are found to possess significantly negative perceptions of the ED classification relative to LD, it may then prove useful to provide more information to them of the potential benefits of providing specialized services to students who meet the criteria for ED. Additionally, this research might lend support to the importance of follow up studies that identify why the ED classification is stigmatizing and what might be done to eliminate negative perceptions that may lack an evidentiary basis.
CHAPTER 2

Review of Literature

This review will examine the history of research on attitudes toward stigmatized individuals generally and the Emotional Disturbance (ED) label specifically. The ED label (formerly, serious emotional disturbance) came into being with the enactment of the Education of the Handicapped Act of 1975 (Public Law 94-142). Therefore, any compilation of the history of the labeling of children with symptoms and behaviors characteristic of ED prior to 1975 requires some knowledge of those euphemisms that were formerly used in the literature to describe such children as well as the behavioral descriptions contained within the federal definition of emotional disturbance. The current federal definition of emotional disturbance reads as follows:

"(i) The term means a condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree that adversely affects a child’s educational performance: (A) An inability to learn that cannot be explained by intellectual, sensory, or health factors; (B) An inability to build or maintain satisfactory interpersonal relationships with peers and teachers; (C) Inappropriate types of behavior or feelings under normal circumstances; (D) A general pervasive mood of unhappiness or depression; (E) A tendency to develop physical symptoms or fears associated with personal or school problems."
(ii) The term includes schizophrenia. The term does not apply to children who are socially maladjusted, unless it is determined that they have an emotional disturbance. (U.S. Department of Education, 1998, p. II-46).

When considering the literature that predates PL 94-142, students who might be classified as ED in the current educational context were likely to be lumped under broader categories often pertaining to mental defectiveness, mental illness, or juvenile delinquency. Two distinct branches of research, deviance literature and mental health literature, looked at issues associated with the labeling of individuals most likely to be classified as ED. Much of the research on deviance originated in the field of social psychology with a strong focus on prejudice, attitudes, and theories of labeling. Kauffman and Landrum (2006) affirmed that the historical foundations of special education for children with EBD (emotional and behavioral disorders) are found in the fields of psychology and psychiatry. EBD is an umbrella term used by some researchers and educational organizations to represent the array of disorders that are commonly associated with children who are qualified for ED services. In particular, the Council for Children with Behavior Disorders (CCBD), a division of the Council for Exceptional Children (CEC), has brought the term EBD into the common vernacular of special education.

**History of Measuring Attitudes toward Groups with Perceived Differences**

Perhaps the most significant early attempt to measure attitudes toward disparate groups was conducted by Bogardus (1925) who generated a “social distance” scale to assess prejudice. Bogardus structured the scale to include statements implying more or
less social contact with a targeted group or individual. Thus, the underlying assumption of the scale was that greater prejudice would be manifest by individuals agreeing with statements that implied more social distance from the group in question.

Thurstone’s (1928) declaration that “attitudes can be measured” marked the onset of attitude research as a serious course of study, rendering the determinations of what constitutes an attitude, and how best to measure such a concept, as major aims of social psychology. With the belief that cognitions and affective states (presumably linked to a behavior or class of behaviors) could be measured, the attitudinal survey (Likert, 1932) became a ubiquitous tool with which researchers would attempt to estimate group perspectives about political, sociological, and psychological phenomenon. The vast majority of attitudinal research toward persons with disabilities has been conducted using such explicit, self-report measures; however, some researchers have used indirect ways to measure attitude-behavior consistency (LaPiere, 1934; Wax, 1948), and more recently, implicit measures of attitude have become a prominent alternative for attitude measurement (Greenwald, McGhee, & Schwartz, 1998).

Definitions of attitude abound and vary primarily according to whether the author includes an affective, cognitive, and/or behavioral component. One of the first and most cited definitions comes from Allport (1935), who described attitude as “the degree of affect for or against an object or a value” (p. 10). Contemporary researchers commonly define attitudes with three attitudinal components: behavior, cognition, and affect. The prevailing understanding of attitudes ties the concept to a belief (cognition), passion to act (emotion), and likelihood to act in a manner consistent with the belief (behavioral
intent). For example, Triandis (1971) defines attitude as “an idea charged with emotion which predisposes a class of actions to a particular class of social situations” (p. 2).

A renowned early attitudinal study examined the strength of association between attitude and behavior (LaPiere, 1934). The LaPiere study presented data on the acceptance rate of 250 establishments, both restaurants and hotels, for accommodating a Chinese couple who was traveling through a portion of the United States with the researcher. LaPiere found that the couple was refused service only one time. Several months after the trip, he wrote letters to the same establishments to make reservations for the Chinese pair. Roughly half of the businesses provided written responses, of which 92% were refusals to service the guests. He surmised from the results that the attitude-behavior correlation was a weak one. It appeared that the hoteliers and restaurateurs possessed attitudes that were not on display when a situation arose in which they physically encountered members of a stigmatized group.

In commenting on the methodological weaknesses of the LaPiere study, Triandis, Adamopoulos, & Brinberg (1984) noted there were no indications that the persons responding to the letters were the same ones who had admitted the Chinese couple previously, and the “attitude objects” varied in the two circumstances with the letter characterizing the visitors only as “Orientals” whereas the actual visits had included a presumably erudite white professor in an expensive car. Nonetheless, the LaPiere study served to illustrate some of the difficulties inherent in mapping a particular attitude to a specific behavior.

Wax (1948), in a well-known early study of prejudice, used an indirect method to examine attitudes of hoteliers in the province of Ontario. Wax contacted a variety of
hotels by letter to inquire about getting a reservation on a particular date. One letter
identified the petitioner as having the surname “Greenberg,” and the other letter used the
surname “Lockwood.” The study found that the response rate to the Lockwood letter was
substantially higher than to the Greenberg letter, as was the number of properties offering
accommodations. The clear implication was that particular attitudes or prejudices are tied
to surnames, and those attitudes in the Wax study appeared to be manifested in a
measurably discriminatory way. Such research marked the genesis of looking at how
names or labels could serve to tie a person to perceived undesirable qualities, which
could then lead to potential acting out of prejudice (i.e., discrimination).

A major focus in social psychology that seemed to be inspired by the uncovering
of the horrors inflicted in the concentration camps, was how prejudice develops and
manifests itself. Allport (1954) defined prejudice as “an aversive or hostile attitude
toward a person who belongs to a group, simply because he belongs to that group and is
therefore presumed to have the objectional qualities ascribed to the group” (p. 7).

Allport (1954) examined the subject of prejudice primarily along ethnic and racial
lines. However, he drew an important distinction between what he referred to as an “in-
group” and an “out-group” – a distinction that would carry over to the mental and
behavioral health literature about prejudice and stigma in later years. Allport described
members of an in-group as those who “use the term we with the same essential
significance” (p. 31). This idea holds that individuals organize themselves into distinct
groups and categorize themselves according to similar features. Those who do not
possess some particular property or feature of that category are then considered part of
the “out-group.” Foreshadowing the emerging concern about labels, Allport described the
need to consider the emotional effects that certain labels can evoke in others when he stated “…any program for the reduction of prejudice must include a large measure of semantic therapy” (p. 187). It was his view that labels often are employed by a speaker “not only to characterize a person’s membership but also to disparage and reject him” (p. 181).

Allport’s (1954) work provided not only a framework for how prejudice forms and manifests itself but also how prejudice can be reduced. He suggested that under certain conditions of interpersonal contact, disparate groups would observe a lessening of prejudice. The optimal circumstances for contact with a member of a stigmatized group would be characterized by four conditions: common goals, cooperative interactions, equality of status of the two groups, and institutional (i.e., societal, legal) support. Allport’s hypothesis would grow into “contact theory,” which would have a significant impact on philosophies that would later guide educational policy regarding the integration of students of varied ability levels and disabilities in the same classrooms. The notion of contact held that allowing children with various exceptionalities to interact with their non-disabled peers would bring about attitudinal change. Presumably, students with false impressions of, or pre-conceived ideas about, children with disabilities would benefit from having contact with those children. In such manner, prejudice would be decreased by allowing the non-disabled to form more realistic perspectives based on genuine interactions as opposed to stereotypes or cultural myths.

A line of generalized research toward the disabled opened up with Yuker, Block, and Campbell (1960). They developed what would become the most commonly-utilized instrument for the assessment of attitudes toward the disabled, the *Attitudes Toward*
Disabled Persons Scale (ATDP). Antonak and Livneh (1988) described the ATDP as “the best known and most widely used of the scales purporting to measure attitudes toward disabled people in general” (p. 134). The ATDP consisted originally of 20 statements describing persons with disabilities, their care, or those associated with them. Respondents evaluated the statements on a 6-point scale, expressing strong agreement at one pole and strong disagreement at the other. Over the next 25 years, the ATDP would be used extensively in research to look at how individuals viewed persons with disabilities compared to “normal” persons.

As early as 1963, researchers began to find evidence that attitudes toward the disabled were more multidimensional than unidimensional in nature. The multidimensional perspective would consider the influence of a variety of factors on attitude formation and the shifting of attitudes toward persons with disabilities based on the nature of the disability and the environmental conditions surrounding their contact with “normal” individuals. Gergen and Jones (1963) conducted a study that looked at several situational variables which could moderate views toward the mentally ill. They grounded their study in the Freudian concept of ambivalence (Freud, 1923/1961), maintaining that individuals begin with mixed feelings toward the disabled, initially torn between sympathy for the hardships associated with the disabling condition and revulsion at dissimilar features from themselves. Depending on how one’s encounter with a person with a disability unfolds, the ambivalence is thought to swing either in a distinctly positive or negative direction thus producing “response amplification” (Gibbons, Stephan, Stephenson, & Petty, 1980).
Gergen and Jones (1963) had 66 white, non-psychiatric VA patients rate their pre and post-impressions of a “stimulus person” (SP) labeled either as normal or mentally ill. First, the participants read a case history about the SP, with half of the participants getting a history that was “normal” or relatively free of notably distressing events and the other half receiving a much more trauma-filled history culminating with psychiatric treatment. The researchers provided an apparatus to allow for participants to work in one room on a matching activity with the unseen SP in an adjoining room. The participants were asked to choose cards from various card pairings that they thought the SP would prefer. The matching activity was staged under a consistent/inconsistent performance condition and a low/high consequence condition, with the experimenter controlling the feedback to the participants under the guise that it was originating from the SP. For the consistent/inconsistent condition, the SP showed more or less agreement and consistency with the participant on the matching activity. In the low consequence condition, a mismatch would result in a light flashing on the participant’s instrument. By contrast, the high consequence condition triggered a “noxious buzzer.” After the exercise, participants had the opportunity to evaluate the SP based on the matching activity. The results suggested that participants were much more likely to evaluate the mentally ill SP in a highly negative way following an interaction that was characterized by inconsistent behavior and the annoying buzzer. Thus, the researchers concluded that attitudes toward the mentally ill could be tilted quite negatively if the person with presumed psychiatric difficulties produced inconsistent behavior coupled with an aversive consequence to the normal individual. Such an interaction was not observed with SPs described as normal, which implies that the participants perhaps were more likely to consider the possibility of
their own performance as being problematic rather than their counterpart when dealing with a supposedly normal person.

One of the first examples of multi-dimensional scaling (MDS) of attitudes toward disability labels was performed by Guskin (1963). MDS is a means of data analysis whereby a researcher assesses what factors might contribute to variance of attitudes toward individuals based on the type of disabling condition they have. Guskin chose 10 “deviancy types” and had participants match behavioral descriptions with the perceived type. He found support for an overarching factor of normal vs. deviant but also observed key differences in perceptions between disabilities. For example, he concluded that children with behavioral deviance were seen as much more deviant than those with limited physical disabilities.

Jones (1974) found support for a general factor of attitudes toward the disabled in his factor analytic study. He examined attitudes of college students toward 13 categories of exceptionalities and discovered that three specific factors emerged from the general factor, including attitudes toward the psychologically disabled, physically disabled, and mildly retarded/non-exceptional. Jones’ clustering of conditions that loaded under the psychologically disabled dimension included delinquent, emotionally disturbed, and severely retarded. Data from these studies suggests that people do not view the construct of disability as a single amorphous entity, but rather, tend to differentiate disabilities according to a variety of factors. The idea of differentiated views toward particular disabilities and their labels would drive research aimed at establishing whether particular special education or diagnostic labels were more or less stigmatizing than others and should be applied with greater caution.
In a longitudinal study completed by Harasymiw, Horne, and Lewis (1976), they found in a sample of over 4000 people, which included teachers, students, and members of the general population, that preferences for interacting with individuals with disabilities could be ordered as follows (from most to least): physically disabled (e.g., diabetic, orthopedic), sensory impaired (e.g., blind, deaf), psychological disabilities (e.g., cognitive impairment, ED), and lastly, those with disorders stemming from social conditions (e.g., delinquent, drug addict). They concluded, based on the relative stability of the rankings across time and across disparate samples, that attitudes toward the disabled may extend along a continuum of normalcy in terms of social value, with perceptions being most negative toward those individuals least equipped to contribute to society in meaningful ways (i.e., low work ethic or productivity).

Abroms and Kodera (1979) surveyed university students in education classes and found three major factors around which the participants ranked disabilities, including organic or metabolic impairment, sensorimotor impairment, and functional impairment (e.g., learning disability, mental illness). They characterized their results as showing greater acceptability of those disorders most amenable to treatment. Disorders with a stronger history of responsiveness to medical interventions were viewed more favorably than those believed to be more resistant to such interventions. For example, disorders of organic impairment (e.g., ulcers, arthritis) were viewed most favorably, followed by sensorimotor impairments (e.g., speech or sight impairment), and lastly, psychoeducational/functional impairments. Emotional and behavior disorders, which would be included under functional impairments, often form a specific factor among attitudes toward the disabled. Such findings imply that there are perhaps distinct views of
the ED label that would cause such children to stand out in the minds of others in some meaningful and potentially stigmatizing ways. If, for example, educators viewed emotional and behavioral impairments as being relatively unalterable (i.e., unresponsive to treatments), then it would follow that there might be a reluctance to expend effort to interact with the child. Teachers desiring a sense of self-efficacy could be repelled by children who are characterized by fairly fixed states of problematic conduct and emotional symptomology.

Numerous studies (Panda & Bartel, 1972; Shears & Jensema, 1969; Tringo, 1970) have shown that individuals tend to rate those disabilities less favorably which appear self-imposed and related to personal conduct (e.g., alcoholic, ex-convict) as compared to those that originate outside the individual’s locus of control, such as a physical disability (e.g., amputee). Attitudes have been found to vary across categories of disabilities, which supports the multi-dimensional view of disability attitudes. Researchers in the 1980s (Antonak & Livneh, 1988; Livneh, 1982) called into question the assumption of unidimensionality present in the most prominent disability attitudes scale (i.e, the ATDP) and the need to revise the instrument to reflect more contemporary social views about disabilities. The assumption of unidimensionality of attitudes toward disabilities dominated much of the early research, suggesting that only one major factor appeared in attitudes toward disabled persons - a dimension of normalcy versus abnormalcy.

Additional research (Schmelkin, 1985) looked at other dimensions along which disabilities may be viewed, including dimensions of physicality, behavioral/emotional vs. cognitive, specific vs. diffuse, and visibility. The physical dimension includes perceptions of disorders that are clearly physical in origin (e.g., cancer, heart disease) as compared to
those conditions for which the etiology is unclear (e.g., autism, mental illness). On the
dimension of behavioral/emotional versus cognitive, Schmelkin found that individuals
sorted some disabilities along the lines of being more purely psychiatric (e.g.,
schizophrenia, ED) as compared to cognitive (e.g., learning disability, mental
impairment) in nature. The specific versus diffuse dimension isolated perceptions of
disorders along a continuum of a specific impairment (e.g., stuttering, deafness)
compared to diffuse or more pervasive impairments (e.g., autism, mental retardation).
The final dimension formed from the analysis was that of visibility, as participants ranked
disabilities according to their obviousness to the eye. Thus, individuals in wheel chairs or
with other assistive devices would be considered more visible in their disabilities than
others who might be able to mask their disabilities more effectively.

Schmelkin’s (1985) study of the multidimensionality of disability labels included
a sample composed of teachers as well as non-educators, and she observed that
perspectives toward the labels were more differentiated among teachers than non-
teachers. Greater differentiation of attitudes or perceptions of labels reflects teachers’
tendencies perhaps to have greater exposure to individuals with a variety of disabilities
and to sort or categorize them along more dimensions. Non-teachers, who presumably
had less exposure to individuals with disabilities, did not evidence as clear separations of
the disabilities on the various dimensions. Therefore, it appeared that the non-teachers
might be less familiar with diagnostic differences between disability labels or perhaps to
approximate more closely the unidimensional perspective of normal versus abnormal in
regards to disabilities. Schmelkin further characterized the perceptions of special
educators as the “most complex” with the implication that perspectives toward
disabilities showed greater “perceptual space” (i.e., became more contrasting and nuanced) based on greater familiarity with individuals with disabilities. Although it might appear on the surface that more knowledge of group differences rather than less knowledge would be a desirable condition, in relating these findings to the work of educators with children with emotional and behavioral disorders, Schmelkin’s findings do not cast any light on whether greater insight into EBD would produce more or less positive attitudes toward those children.

Guided by a theoretical perspective, Jones et al. (1984) highlighted six key dimensions or factors in the formation of stigma, including concealability, course, disruptiveness, aesthetic qualities, origin, and peril. Concealability pertains to how effectively individuals with a disability can keep characteristics of their disability hidden from others. This is what Goffman (1963) described as “passing” – a process whereby individuals are able to pass themselves off as “normals” despite possessing some feature or behavior that society looks down upon or views disapprovingly. The course of a disability (referred to as a “mark” by the authors) is the pattern of presentation, whether it be an enduring condition over the lifespan or something that abates with time. They note that disruptiveness is not as conceptually clear as the other factors that form stigma; it is the property of a disability that disturbs the quality of interpersonal relationships. Aesthetics of disability involve the degree to which an individual will be perceived as being physically attractive versus disfigured. When considering a person’s own culpability, or lack thereof, in producing a disability, the concern about origin is a consideration in the formation of stigma. Lastly, stigma may be heightened or lessened by the perceived peril that an individual with a disability poses to others. The idea of
peril is especially relevant to populations of mental patients and those who have performed violent acts.

When one applies the six factors from Jones et al. (1984) to the criteria that typically exemplify children with behavioral disorders, it would appear that children with EBD would perhaps particularly be suited for stigmatization given their outward manifestation (i.e., limited concealability), somewhat pervasive and enduring course during the lifespan, high level of disruptiveness in peer and other social relationships, and potential for injury or harm to others (especially with extremely antisocial behavior). In terms of the origin of disorders for those who are termed behaviorally-disordered, Page (1984) noted, “Those with conduct stigmas are generally considered to be personally responsible for their failings. It is commonly believed that such individuals have deliberately chosen to behave in socially unacceptable ways” (p. 6). Furthermore, Palmer (1983) suggested that teachers or peers who like a child with a disability will attribute negative social or academic behaviors to external factors, whereas children who are disliked will have such behaviors attributed to internal factors. Given that deficits in interpersonal functioning are often characteristic of children with ED (Ialongo, Vaden-Kiernan, & Kellam, 1998), they are a population that would seem particularly susceptible to the negative attributions of others.

In a review of literature about teacher attitudes toward children with disabilities, Hannah and Pliner (1983) examined the relationship between teachers’ contact with students with disabilities and the effect of that contact on their attitudes. They concluded that positive contact has been shown to have a favorable impact on teachers’ attitudes.
Consistent with the idea of response amplification, those teachers who possessed negative attitudes were likely to have had negative interactions with students with disabilities.

Yuker (1988) stipulated a number of variables involved with the contact between a person with a disability and a non-disabled person that can determine whether the outcome is positive or negative. Those factors include competence, social skills, misbehavior, communication skills, athletic skills, age, education, personality, and credibility. Within the educational context, *competence* refers to the student’s level of academic functioning. When Yuker cites *age*, he explains the effect of individuals having more positive views toward younger children with disabilities compared to older children, due to being considered cuter. A curious effect is observed in terms of educational level. Yuker indicated that the timing is important in the contact; individuals with disabilities in college are more likely to evoke positive responses from others, whereas high school-age children with disabilities produce mixed effects. Among those *personality* characteristics more likely to produce positive contact with an individual with a disability, he points to manageability, likability, and treatability. Finally, it appears that *credibility*, or the consistency of attitudes and behaviors of individuals with disabilities, plays an important role in establishing the favorability of the impressions of their non-disabled counterparts.

Corrigan and Penn (1999) identified contact, protest, and education as the “stigma-reduction armamentarium” for combating mental health-related stigmas. As an example of effective use of protest, they cite advocacy groups voicing opposition to unfair and errant portrayals of individuals with mental illness in the media. However, Corrigan and Penn noted that protest is a reactive method, thereby doing little to provide
others with alternative, reality-based conceptions of the group in question. They emphasized the importance of a proactive strategy, such as education, which plays an important role in countering prejudicial assumptions that others might have about those with mental disabilities. Undergirding the assumptions about combating stigmas, similar to notions about fighting blind stereotypes, is that the stigmas are incorrect worldviews of a particular disorder or individuals with the disorder. If indeed a serious stigma toward the ED label exists among educators, then these prescriptions toward mental illness suggest that teachers need to have more contact with EBD students, more education dispelling myths about perceived negative attributes, and more opportunities to hear alternative opinions that protest unfair characterizations of such students.

Although there is strong evidence that general attitudes only weakly predict specific behaviors or actions (Ajzen, 1982) and that traditional attitudinal measures do not relate consistently to actual behavior (Fishbein, 1967), evidence is much stronger that general attitudes correlate with, or are predictive of, general patterns of behavior (Fazio & Roskos-Ewoldsen, 1994). Also, there has been shown to be a fairly strong relation between attitude and behavior when multiple measures of behavior are obtained to correspond to a general measure of attitude, when measures of attitude have been administered properly, and when the qualities of the attitude object closely resemble those of the measured behavior (Fishbein & Ajzen, 1975).

A major question of concern in attitudinal research has been and remains, what is the relation between attitude and behavior? Indeed, if attitude had little or no bearing on one’s actual behavior, then it would seem to be hardly an important subject of study. The value of examining attitudes has always resided in the assumption that the behavioral
intent (attitude) to some extent would serve as a predictor of one’s behavior (Ajzen & Fishbein, 1980).

Self-monitoring has been cited as an important variable in moderating individuals’ propensity to behave publicly in accordance with private attitudes (Becherer & Richard, 1978; Lutsky, Woodworth, & Clayton, 1980; Zann, Olson, & Fazio, 1980; Zuckerman & Reis, 1978). Snyder (1982) contends that low self-monitoring individuals, or those who are less susceptible to conforming to behavioral patterns of a specific social context, will demonstrate stronger consistency between their attitudes and covert actions. The assumption that underlies much of the literature that will be cited here is that individuals possessing negative attitudes will be more likely to respond in unfavorable or stigmatizing ways toward an attitude object (i.e., student). In the discussion of attitudes toward the ED label, self-monitoring research would portend that educators who are less conforming are more likely to have stronger correspondence between their self-reported attitudes and their behaviors. A disproportionality of conformists within the ranks of the teaching profession perhaps might yield greater discrepancies between stated attitudes and actual behavior.

Researchers discovered a new tool for examining attitudes in a less transparent way with the development of the Implicit Association Test (IAT) (Greenwald et al., 1998). Typically staged as a computer-based measure, an IAT presents respondents with attitudinal objects (i.e., words or pictures) that they must use keystrokes to associate with pairs of overarching categories. For example, a picture of a daisy might appear on the screen, and the participant would have to choose to associate the picture with the words “good” and “flower” in one corner of the screen and “bad” and “insect” in another corner.
of the screen. The latency of tapping the key on the keyboard then provides an indicator of the strength of the association. Conditions are reversed several times, thus rendering the opportunity to make a presumably more difficult association of “insect” and “good” when a picture of a spider appears on the screen. As Nosek, Greenwald, and Banaji (2005) described it, “…the IAT rests on the assumption that it ought to be easier to make the same behavioral response (a key press) to concepts that are strongly associated than to concepts that are weakly associated.” An IAT, as a measure of implicit cognition, provides a means to examine mental associations in ways that are thought to differ from the more deliberate, introspective results of self-report measures in meaningful ways.

By developing a means of assessing attitudes toward stimuli by looking at implicit cognitive associations, researchers now have a means to avoid some of the shortcomings connected to self-report measures. Some of the recognized shortcomings include respondent sensitization, response styles, and reactivity (Antonak & Livneh, 1995).

*Respondent sensitization* occurs when respondents to an attitude measure have perhaps little or no connection to what is being measured (i.e., the attitude referent). The researcher then may be susceptible to overinterpreting the stability or generalizability of the attitudinal data, which in truth reflects an ephemeral phenomenon. When personality attributes affect participants’ pattern of responding (for example, answering most items in a positive or negative direction) and do not reflect their genuine attitudes, they are said to have produced *response styles*. The third area of vulnerability with explicit attitudinal measures, *reactivity*, refers to participants’ awareness of being assessed or involved in a research venture and their corresponding efforts to respond in a fashion to impress or assist the researcher or to meet their own hidden needs in some way. Reactivity can take
a variety of forms and may be a consequence of the participant: (a) seeking to confirm the researcher’s hypotheses (experimenter demand effect), (b) impressing others with their level of sophistication (evaluation apprehension effect), (c) erring on the side of caution in their treatment of the attitudinal subject (generosity effect), (d) embracing only that content that is perceived to be socially and/or culturally approved (social desirability bias), (e) undermining the aims of a study by taking a purposefully opposite point of view (sabotage effect), (f) bringing little or no insight and introspection to the task at hand due to disinterest or low motivation (thoroughness effect), and (g) refusing to share information due to a guardedness of one’s own private world and reservations about disclosing potentially atypical attitudes (refusal bias).

In the relatively brief existence of the IAT, a wealth of studies has been done using the method. The IAT garnered early attention for indicating implicit racial bias in favor of white versus black (Greenwald et al., 1998). Hofmann, Gawronski, Gschwendner, Le, and Schmitt (2007) performed a meta-analysis involving 126 studies, comparing the correlations between the IAT and explicit attitudinal measures. Although Hofmann et al. observed that implicit and explicit measures have traditionally produced relatively low correlations, they concluded that the two are systematically related but may show weaker correlations under conditions in which the two tests lack conceptual correspondence and in which response to the explicit measure is less spontaneous. A lack of conceptual correspondence is observed when the explicit and implicit measures have not been developed to measure the same construct, thus producing understandably weaker correlations. Greater deliberation in responding to explicit measures has been linked to lower correlations with IAT measures due to the greater likelihood of reactivity
or other self-report effects holding greater sway on the respondent’s choices. Because the IAT is thought to elicit attitudes that are below the level of conscious thought, it stands to reason that more spontaneous (i.e., initial “gut feeling”) responses on explicit measures would be more likely to resemble the IAT results than those in which respondents second-guess and over think their answers.

In a study of the IAT that looked at the stigma of smoking (Swanson, Rudman, & Greenwald, 2001), the researchers found a much stronger correlation between the responses of non-smokers on explicit and implicit measures compared to the responses of smokers. Whereas the smoking participants indicated more favorable attitudes toward smoking on the self-report measure, their results on the IAT suggested a negative assessment of smoking. Swanson et al. speculated that the inconsistency of smokers’ responses on the implicit measure may have evidenced the wide-held understanding of smoking’s status as a stigmatized and addictive behavior. That would be a case, some would argue, where the IAT results may reflect knowledge of a stereotype as opposed to capturing one’s own deep-seated belief. However, one might also argue that smokers are more likely to hold negative implicit attitudes toward smoking of which they are not consciously aware, or are unwilling to convey, in a more introspective format such as a self-report measure.

White, Gordon, and Jackson (2006) used an IAT and various explicit scales to examine attitudes of undergraduate students toward athletes with disabilities. Their research found evidence of a strong negative bias toward the athletes on the administrations of the IAT; that is, the latency of responses on the implicit tests suggested that participants were more much comfortable associating positive terms with non-
disabled athletes than with athletes with disabilities. White et al. suggested their study was the first to employ the IAT in assessing attitudes toward individuals with disabilities. Their work points to the benefit of looking at potential bias in implicit attitudes of teachers toward disabilities. Given the increasing exposure that regular education teachers have to students with a wide range of disabilities (and corresponding special education labels), it may be helpful to identify whether certain groups of children are more likely to be viewed negatively than others and whether stigma-reduction services are needed.

**History of the Emotional Disturbance Label**

In the 1800s, the tendency of many individuals was to diminish or play down the existence of psychological and behavioral pathologies of youth and focus mainly on the presentation of adult syndromes. One of the few exceptions, as cited in Winzer (1993), was Henry Maudsley, who in 1867 affirmed the notion that adult psychoses could be observed in children as well. In stark contrast to contemporary models of servicing children with cognitive and emotional disabilities, early educational services were offered within segregated, institutional environments. This also was an age in which mental deficiency and moral deficiency were identified as separate afflictions of the mind and spirit, though both were often viewed as being brought on by an individual’s sinfulness.

Although children with severe emotional problems often were not differentiated from children with mental defectiveness, Samuel Gridley Howe, founder of one of the first schools for the mentally retarded in 1848, attempted to delineate between mental idiocy and what he referred to as “moral idiocy.” Gelb’s (1989) citation of Howe’s thinking on the matter is informative: “This [moral idiocy] was a condition “in which the
sentiments, the conscience, the religious feeling, the love of neighbor, the sense of beauty” were defective, although the intellectual faculties were normal” (p. 365). Howe maintained that many persons who would be characterized as “moral idiots” were key constituents in prison populations at that time. In many ways, his descriptions shared features with, or resembled characteristics of individuals who later would be termed severely behavior disordered.

Subsequent to Howe’s efforts, similar schools arose such as the Pennsylvania Training School for Idiotic and Feeble-Minded Children in 1852 and the renowned New Jersey Training School for Feeble-Minded Boys and Girls in Vineland, New Jersey. The proliferation of institutions for children with mental impairments spiked throughout the second half of the 19th century, growing from five institutions in North American in 1857 to an estimated 14 states having such institutions by 1890 (Davies, 1959).

Winzer (1993) acknowledged that those children who might be categorized as severely emotionally disturbed were the recipients of less educational concern prior to the 20th century. She notes, “But it was not until mental retardation was viewed as a distinct category of handicap, with its own teaching and institutional approaches, that severe emotional disturbance could in turn emerge as a discrete entity” (p. 115). Nonetheless, special segregated classes (sometimes referred to as “ungraded classes”) arose in the latter portion of the 1800s to serve children found to be incorrigible, low achieving, and/or frequently absent. Winzer surmised that those early special education classes functioned “as the transmission belt to move disabled children and youth and those displaying behavioral problems beyond schools” (p. 322). With those facts in mind, it can
be inferred that the groundwork for truly differentiated educational services for children with severe emotional disorders still needed to be laid at the turn of the 20th century.

For children who presented with greater behavioral dysfunction in the EBD pantheon, reformatory and correctional facilities arose to rehabilitate, or at the very least, contain such individuals. Specialized schools called “industrial schools” were formed in the latter half of the 19th century with the objective of assisting children who were not amenable to public schooling. Antisocial children were a key demographic who attended the industrial schools, and the emphasis of their educational programming was on learning moral behavior and adopting vocational skills. The United States was operating 135 industrial schools by 1919 (U.S. Bureau of Education, 1917-1918). The prevailing social mindset toward young people with severe behavior problems appeared to be that they were best served and rehabilitated within an institutional setting.

An important factor in the servicing of children with various exceptionalities, including behavioral and emotional problems, was the enactment of compulsory attendance laws in 1909. Such laws stiffened the requirement of schools to monitor and enforce students’ attendance, which elevated the need to consider truancy among those problem behaviors that warranted intervention. Whereas schools had grown accustomed to letting children with disabilities and/or conduct problems simply choose not to attend, decisions about dealing with such students more effectively would become increasingly more pressing.

Although labels such as ‘idiot,’ ‘feeble-minded,’ and ‘imbecile’ had been applied to children with exceptionalities in the 19th century, new classifications emerged in the early part of the 20th century that had the appearance of being more sensitive to these
youth. The new classifications contributed to greater diagnostic specificity and wider services for educational agencies that began to sort children in more varied ways and with an eye toward differentiated services. Two categories that splintered off from the mental retardation classification included emotional disturbance and learning disability.

Winzer (1993) indicated that the term *emotional disturbance* first came into use in 1910. Although the term would not be ratified as a formal federal special education label until almost 65 years later, Winzer noted, “The old, elastic category once labeled “incorrigible” now was characterized as mild emotional disturbance (behavior disorders) and it continued to be a concern of the educational and legal systems” (p. 339).

*Incorrigible*, along with the euphemisms *maladjusted* and *disturbed*, were some of the prevalent labels and descriptors in the early years of identifying children who later would be later referred to as emotionally disturbed (Kauffman & Landrum, 2006).

One of the earliest attempts to formally label students exhibiting behaviors of severe conduct problems and/or emotional disturbance was done by John Horn (1924). Horn created a wider taxonomy of ways to group children of various exceptionalities that included three broad categories and seven subgroupings. Kauffman and Landrum (2006) provided the following representation of Horn’s system:

I. Children who are exceptional for reasons primarily mental
   1. The most highly endowed group.
   2. The most poorly endowed (but not feeble) group.

II. Children who are exceptional for reasons primarily temperamental
   1. Incorrigibles and truants.
   2. Speech defectives.
III. Children who are exceptional for reasons primarily physical

1. The deaf.
2. The blind.
3. The crippled. (p. 49)

Horn’s broad category that corresponds to emotional disturbance was termed exceptionalities “for reasons primarily temperamental” and included the subgroup of children labeled “incorrigibles and truants.” Incorrigibility would remain in fashion for these students for the first half of the 20th century as the process of defining and categorizing their condition(s) continued to evolve.

For the first half of the 20th century, children with severe emotional and behavioral problems were very likely to receive educational services within segregated institutional settings. Kauffman and Landrum (2006) noted that at midcentury, there remained little differentiation between services for children with EBD and broader, more ubiquitous categories such as mental retardation. Significant steps toward educating children in non-institutionalized settings was seen with the advent of the “600” schools in New York City in 1946 (Berkowitz & Rothman, 1967). The 600 schools provided education for youth then termed “maladjusted” or “disturbed” in day schools housed within regular school settings, although schools in more restrictive settings such as residential treatment were common as well. Children were often channeled to such settings when they simply exceeded the capacities of public schools and/or parents to manage them. The focus of treatment was on rehabilitation as educators and therapists sought to remediate students’ undesirable behaviors. It is perhaps of historical interest to acknowledge that the nomenclature of the schools as a numeric quantity reflected efforts
of administrators and educators even at that time to be sensitive to the possible stigmatization that might follow if the schools were named according to the malady or disabling condition for which they operated. Thus, the generic number represented a purposeful ambiguity intended to offset one potential means of fostering prejudice. In spite of the presumed good intentions, the 600 branding was eventually dropped for the very reason that it became synonymous with emotional disorders.

Kauffman and Landrum (2006) noted that the special education of children with emotional and behavioral disorders really ascended by the end of the 1950s. Along with more systematic attempts by states and localities to provide educational services for youth with emotional and behavioral disorders as evidenced by the founding of specialized schools like the 600 schools, came the first instructional texts for educating students with those needs. Leonard Kornberg’s (1955) classic text, *A Class for Disturbed Children: A Case Study and Its Meaning for Education*, incorporated aspects of psychoanalysis and interpersonal therapeutic process, looking at the importance of the therapeutic contact between persons in a classroom in the transmission of meaning. A text from Berkowitz and Rothman (1960) reflected a treatment/educational approach grounded in Freudian thought, namely, that instruction should only proceed after underlying mental conflict for a student has been resolved.

Although the early texts in support of children with EBD reflected the strong Freudian training of psychologists at the time, psychoanalytic ideas and treatment methods would quickly be supplanted in popularity by behaviorism (Skinner, 1953). The rise of behaviorism marked a radical shift in the conceptualization of, and treatment of, behavioral problems. Skinner’s ideas of operant conditioning held that many behaviors
are shaped and maintained by environmental consequences. In sharp contrast to the psychoanalytic viewpoint of internal mechanisms accounting for neurotic behaviors, behaviorism theorized that the etiology of problematic behaviors as well as the internal world of children need not be speculated upon in order for successful modification of behavior (i.e., treatment) to occur. Perhaps more so than any other psychological model, behaviorist principles would have a significant influence on classroom management of students with EBD and would be inextricably intertwined with special education law at the end of the 20th Century.

In addition to the evolution of how emotional and behavioral disorders would be conceptualized and treated, the field would be greatly affected by several public laws. First, Public Law 88-164, signed into law in 1963, added “seriously emotionally disturbed” as a category of students for which colleges and universities would receive grants for the specific training of teachers. Increasingly more attention was being given to the provision of services for students with emotional disturbance during the 1960s, but it was not until PL 94-142 in 1975, the Education of the Handicapped Act, that a formal definition for the SED category would be introduced and that a mandate would be given to states to provide individualized educational service plans to children with disabilities and to educate them within the least restrictive environment. Whereas segregation of special education services had been the norm prior to PL 94-142, schools and districts had to begin rethinking the placement of their exceptional students and consider the possibility of supporting their needs within general education classrooms. Parental rights took a giant leap forward as well with schools being required to maintain confidentiality of special education records and to grant them the right of due process.
The basis for the five criteria used in the PL 94-142 definition of serious emotional disturbance came from the writings of Bower (1960). Remarkably, and in spite of growing controversy over the next several decades, the original definition would remain largely unchanged even until the most recent reauthorization of special education law, the Individuals with Disabilities in Education Improvement Act (IDEIA) in 2004.

A debate among educators and advocacy groups was being waged about the adequacy of the federal definition as soon as PL 94-142 was passed. They cited shortcomings of the definition and the need to clarify ambiguous terms that made it difficult to identify these children (Kauffman, 1976; Hallahan & Kauffman, 1977). Calls to revise the definition of ED have been made over many years on the basis of confusion about the term socially maladjusted (Costenbader & Buntaine, 1999; Merrell & Walker, 2004; Nelson & Rutherford, 1990; Skiba & Grizzle, 1991), the lack of clarity about inclusion of students with behavioral disorders (Forness & Kavale, 2000), and vague specifiers regarding duration and severity (Kauffman, 2001).

Some support was given in the early 1980s for the use of the DSM-III to resolve definitional problems, with Slenkovich (1983) contending that special education law intended for the SED category to reflect students with internalizing rather than externalizing problems. The dichotomy of internalizing versus externalizing has come to represent a common mode of distinguishing between types of behaviors. Gresham and Kern (2004) defined internalizing behaviors as “behaviors that are directed inwardly toward the individual and represent an overcontrolled and inner-directed pattern of behavior” (p. 262). Examples include social withdrawal, depressive features, and anxiousness. By contrast, externalizing behaviors are manifested as outwardly-directed
patterns of behavior that consist of acting out toward objects within the environment. Aggression, defiance, and hyperactivity are representative of externalizing behavior.

Criteria from the ED definition that have been criticized as ill-defined and vague include the specifiers “long period of time” and “marked degree” (Cullinan, 2004; Forness & Kavale, 1997). School districts have been left without any guidance as to what constitutes a lengthy duration of presentation of problem behavior. Similarly, no prescriptions have been provided in federal legislation to suggest what magnitude and severity of emotional/behavioral symptoms would be meant by the description “marked degree.”

The so-called “social maladjustment” clause within the federal definition has generated significantly more negative feedback than any other portion. The clause indicates that the ED “does not apply to children who are socially maladjusted, unless it is determined that they have an emotional disturbance” (U.S. Department of Education, 1998, p. II-46). Traditionally, this has been thought to be a means of excluding children with conduct disorders from eligibility for special education services under the ED label (Nelson, Leone & Rutherford, 2004). Students who are conduct-disordered, it has been frequently maintained, are more suited to rehabilitative services in the juvenile justice system. Some institutions have viewed emotional disturbance as a disabling condition that is not a product of willful behavior on the part of a student, whereas social maladjustment or conduct disorders are often seen as volitional. Such a conceptualization has been contested heavily by researchers who contend that there are no practical or meaningful guidelines for differentiating social maladjustment from emotional disturbance (Merrell & Walker, 2004), and as a consequence, many states do not exclude
children from being qualified under the ED category in spite of the definition (Forness, 1992).

The very name of the ED category came under fire from critics who indicated it was not suggestive of, or linked in any way to behavioral dysfunction (Forness & Knitzer, 1992). Kidder-Ashley et al. (2000) surveyed 41 state education agencies and found many questioned the adequacy of the phrase *emotional disturbance* as a disability category. In fact, of the 41 agencies, 15% eschewed the federal term in favor of a behavioral one (e.g., behavioral disorder) and 24% included both emotional and behavioral terms in their classification term (e.g., emotional and behavioral disorder or EBD). Thus, roughly 40% of states appeared to break from the federal terminology to employ descriptive terms of their own choosing.

In 1989, the Council for Children with Behavioral Disorders proposed that the term *serious emotional disturbance* be changed to *emotional or behavioral disorder* (EBD). However, it wasn’t until the reauthorization of the Individuals with Disabilities Education Act of 1997 that special education law would reconsider the label. As it turned out, the modifier of ‘serious’ was dropped from the name, but the emotional disturbance moniker remained in place.

In the early 1990s, the National Mental Health and Special Education Coalition attempted to remedy some of the confusion within the federal criteria for emotional disturbance, by recommending a new definition for the category. The NMHSEC definition eliminated the nebulous clause about social maladjustment and opened up the possibility of being more inclusionary of children with conduct disorders. As cited in Forness and Knitzer (1992), it reads:
(i) The term emotional or behavioral disorder means a disability characterized by behavioral or emotional responses in school so different from appropriate age cultural, or ethnic norms that they adversely affect educational performance. Educational performance includes academic, social, vocational, and personal skills. Such a disability

(A) is more than a temporary, expected response to stressful events in the environment.

(B) is consistently exhibited in two different settings, at least one of which is school-related; and

(C) is unresponsive to direct intervention in general education or the child's condition is such that general education interventions would be insufficient.

(ii) Emotional and behavioral disorders can co-exist with other disabilities.

(iii) This category may include children or youth with schizophrenic disorders affective disorders, anxiety disorders, or other sustained disorders of conduct or adjustment when they adversely affect educational performance in accordance with section (i) (p. 12).

Subsequent special education law has not adopted any of the proposed alternative language for the definition. With the most recent reauthorization of IDEA in 2004, the Federal Register (U.S. Department of Education, 2006) confirmed that the matters of dropping the social maladjustment clause and changing the category name to emotional
and behavioral disorders were subjects of discussion. Nonetheless, the record indicated, “Given the lack of consensus and the fact that Congress did not make any changes that required changing the definition, the Department recommended that the definition of emotional disturbance remain unchanged” (p. 46550). In making that determination, the 2004 law retained the definition and classification name virtually unchanged from its inception 30 years previously.

In addition to focusing on the many semantic concerns related to legal terminology about the ED label, the research literature in the 1990s began to examine more intently the problems that might account for under identification of children with emotional and behavioral disorders. Some of the under identification was hypothesized to relate, in fact, to the definitional imprecision of the category. Prevalence rates for ED placement varied considerably from one state to the next, and Wright, Pillard, and Cleven’s (1990) analysis found that one-third of the variance in the rates was due to differences in states’ definitions; they surmised that two-thirds of the variance was due to unexplained phenomena.

Another source of considerable alarm within special education generally, and within the community of behavioral disorder researchers in particular, has been, and is, the relatively late stage identification of children with EBD. On average, students who are classified ED are identified much later on average than for any other category of special education services, peaking at the ages of 14- to 15-years (Walker, Nishioka, Zeller, Severson, & Feil, 2000). The relatively late identification of these students is particularly troublesome in light of a multitude of negative academic, occupational, and career outcomes that have been well-documented over the years. Wagner (1991) found
that students receiving ED services were characterized by dropout rates significantly higher than those of students of any other disability group. Also, the prevalence of ED-labeled children among residents of correctional and detention facilities was found to be eight times higher than among the general school-age population (Quinn, Rutherford, Leone, Osher, & Poirier, 2005). The literature has shown that early intervention services, in addition to benefitting those children with significant academic delay, provide strong benefits for the prevention and remediation of children with significant behavioral and emotional problems (Conroy, Hendrickson, & Hester, 2004). Several writers and researchers have speculated on why it takes so long to identify EBD students for special education services. Kauffman, Brigham, and Mock (2004) identified three primary reasons that account for the delay in identifying these students, including (a) personal philosophy, (b) definitional imprecision, and (c) pragmatic concerns. *Personal philosophy* relates to the distaste for labeling children and especially for labeling children with what are perceived to be particularly stigmatizing conditions. *Definitional imprecision*, as discussed previously, is a factor in under identification simply because schools are confused about, or reluctant to use, the vague diagnostic criteria of the federal language in qualifying certain children for ED services. Chief among the *pragmatic concerns* that may drive schools to under identify these children might be fear about significant behavioral problems (e.g., aggression, fire starting, etc.) that would not be as easily disciplined for children bearing an ED labeled compared to children who are unlabeled.

When considering the eligibility of children for ED services, the assessment procedures traditionally have included behavioral ratings and observational data.
Behavioral rating scales are administered to parents, teachers, and in some cases, the children themselves. Such ratings are normed to show whether the rate of occurrence or perceived level of certain behaviors is at or above average for a certain age and/or gender. A school psychologist or other school staff would perform structured or unstructured observations, typically in the classroom, with the intent of discerning the function of problematic behaviors or to estimate the frequency, duration, or severity of target behavior(s). Polsgrove (2004) identified a third means of assessing students for EBD, namely, developing and monitoring academic and behavioral intervention plans. Closely aligned with a specific type of assessment referred to as a functional behavioral assessment (FBA), a behavior intervention plan takes the prescriptions of observational data from an FBA and puts into place environmental changes to increase desired behaviors and/or reduce problematic behaviors.

One of the first to recommend a shift away from the categorical identification of children with emotional disturbance, Gresham (1991) advanced the notion that behavior disorders should be established through a resistance to intervention criterion rather than through traditional psychometric measures. He recommended that practitioners and teachers follow a behavior-analytic view of student’s behavioral problems, thus considering what changes might be made to the student’s school environment in order to facilitate behavioral change rather than assuming that undesirable behaviors are a product of intra-individual pathology. Gresham identified a variety of factors that should be considered in determining resistance to behavioral intervention, including severity of behavior, chronicity of behavior, generalization of behavior change, tolerance of behavior, and treatment strength, acceptability, integrity, and effectiveness. According to
Implicit and Explicit Attitudes of Educators

his diagnostic standard, Gresham contended that behavior disorders could be determined if a student did not show adequate responsiveness to classroom interventions.

There have long been concerns about the effectiveness of special education. Worries about labeling were, in part, tied to the notion that the discrimination or stigmatization that might result from the labeling would not be offset by appreciable gains in academic, social, or behavioral functioning as a consequence of receiving special education services. Landrum, Tankersley, and Kauffman (2003) reviewed the literature on what educational practices had been used to good effect for students with characteristics of EBD. They highlighted a number of behavioral techniques, including reinforcement, precision requests, behavioral momentum, time-out, response cost, and group-oriented contingencies for the management of inappropriate behavior. For academic learning problems, they identified self-monitoring, continuous progress monitoring, Classwide Peer Tutoring, and direct instruction. In addressing unsatisfactory interpersonal relationships, the authors acknowledged the positive effects of facilitating practice in natural settings, modifying antecedents and consequences, and providing direct instruction of individually targeted behaviors. They acknowledge the relative resistance of students with EBD to long-term positive behavioral change and treatment effects due to being identified so late in their development as well as having disorders that are widely considered to be lifelong in spite of preventive efforts.

Echoing the frustration of many in special education about the years-long debate over the differentiation between social maladjustment (SM) and ED, Heathfield and Clark (2004) advocated for a move away from the traditional model of service delivery and the adoption of a “comprehensive school-based mental health model” in serving
children with emotional and behavioral problems. Their model incorporates proactive efforts such as school-wide screening of social-emotional skills as well as the implementation of primary, secondary, and tertiary prevention services.

With the implementation of IDEIA 2004, the language of special education law relative to emotional disturbance was not altered in any meaningful way. However, the door was opened for an evolution of the way in which students with ED are identified. For example, in much the same way that Gresham (1991) advocated for a method of looking at a child’s response to behavioral intervention for informing the ED eligibility decision, the new law recommended that states look at a process of using response to scientific, research-based interventions for determining the presence of a specific learning disability. Formerly, most states had employed an assessment model that used discrepancies between scores on standardized ability and academic achievement measures in making the determination of a learning disability. In some areas of the country, schools have already begun to apply a response to intervention (RtI) paradigm to students’ behavioral problems in much the same way that such a process has been and is increasingly being used to resolve academic problems. Through such an RtI process (Batsche et al., 2006), categorical yes/no considerations for considering an ED label take a back seat to the collection of data in the classroom to show whether a student responds to an intervention.

Emotional disturbance remains, however, a category of services with lingering, unresolved questions going forward as to the best methods for applying meaningful diagnostic criteria, identifying children early for preventive services, and intervening effectively on the behalf of students with EBD. Although conventional wisdom holds that
attitudes toward the ED label have remained negative since landmark studies in the late 1970s and follow up studies in the 1990s, legislators have done little to address changing the terminology of the potentially stigmatizing categorical name (emotional disturbance) or to modifying the confusing eligibility criteria. Therefore, it appears that there has not been sufficiently convincing evidence to suggest that the term may elicit particularly negative perceptions among educators. A timely examination of educators’ attitudes toward the label may serve to strengthen the case for clarifying the ED terminology and thus improve the likelihood that students receive such services.

**Attitudinal Research toward the Emotional Disturbance Label**

Because of the ongoing refinement of assessment and attempts to better understand the characteristics of students who fall under the category of emotional disturbance, there is no literature about teacher attitudes toward these children in the early part of the 20th century. This was due not only to the lack of a cohesive classification of children with severe emotional and behavioral difficulties, but also to the changing circumstances of their educational placements.

Csapo (1984) indicated that the educational establishment did not take ownership of providing services for the seriously emotionally disturbed until the 1960s. Thus, it follows that teachers’ perspectives toward such students were not a subject of formal study until that decade and beyond. Research about attitudes toward children with ED or toward the ED label most often took the form of comparative studies where participants rank order their preferences of one special education category relative to others. Commonly, researchers have compared the ED label to the learning disability (LD) and mental retardation (MR) labels.
In an early labeling study that compared teachers’ attitudes toward mental deficiency, psychopathy, schizophrenia, and cerebral palsy, Combs and Harper (1967) found evidence that attitudes of teachers and teachers-in-training toward retardation labels were more positive than those toward mental/psychological conditions. The researchers did not speculate or attempt to explain why this effect may have occurred but warned clinicians against using diagnostic labels with teachers, especially in the case of psychotic disorders, that “elicited significantly more negative attitudes than did labels for neurotic, retardant, or neurological disorders” (p. 403).

Hughes, Kauffman, and Wallace (1973) had a sample of teachers rate the following labels: learning disabled, educationally handicapped, maladjusted, problem child, emotionally disturbed, and behaviorally disturbed. They found teachers rated labels that indicated an academic deficit more positively than those that were psychological or behavioral in nature. Foster, Ysseldyke, and Reese (1975) observed that teachers-in-training assessed children labeled ED negatively, and those biases appeared to persist even after behaviors inconsistent with the label were presented.

MacDonald and Hall (1969) took a different approach in their study of graduate students’ perceptions of the effects of five types of disabilities, including internal disorders, sensory disorders, disfigurements, amputations, and emotional disorders, on the social functioning of individuals with the disabilities. They assessed the raters’ levels of locus of control (either external or internal) with Rotter’s Internal-External Locus of Control Scale (Rotter, 1966). The researchers defined locus of control as dealing with “the extent to which persons view themselves as having some control over their reinforcements” (p.655). The graduate students also completed a disability scale that
estimated their perceptions of how debilitating the five disability types were. Those participants with higher internal locus of control scores were found to evaluate emotional disorders as being much more debilitating than did those individuals who were identified as having higher external loci of control. Thus, there was support for the notion that individuals more inclined to believe that people control their own destinies (i.e., internally controlled individuals) viewed emotional disorders as being more debilitating than those who would place greater weight on the role of external forces in determining behavior. Of the dimensions of social functioning examined (vocational, marital, social, parental, familial, and personal), the emotional disorders category was rated to have the least impact on vocation but the most negative effect on the parental relationship. This research brought out questions of what characteristics of individuals who deal with persons with disabilities might contribute to greater stigmatization in their interactions. Also, the multi-categorical approach to assessing social functioning suggested that non-disabled individuals might have more nuanced views of how disabilities affect different relationships and/or present differently within a variety of settings.

Research about attitudes toward specific special education labels may have begun with Jones (1972), who suggested that no empirical study of labels and stigma on exceptional children in the public schools had been conducted prior to his study of perceptions of the labels *educable mentally retarded, culturally disadvantaged*, and *culturally deprived.* He set out to examine a hypothesis purportedly held by many special educators, namely, that deficit labels adversely affect the learning of labeled individuals. Among his findings, Jones concluded that teachers possessed lower expectations of children labeled culturally deprived and culturally disadvantaged, and teachers confirmed
the presence of stigma associated with special class placement for those children labeled educable mentally retarded.

Studies that have measured attitudes toward ED compared to the LD and MR labels have been complicated by the fact that there are diagnostically similar features and significant overlap found in the behavioral symptomology of the conditions. In particular, students with ED and LD often bear similar qualities of academic struggles that can make it difficult to differentiate students for services (Hallahan & Kauffman, 1977).

Shotel, Iano, and McGettigan (1972) examined regular education teachers’ attitudes toward the integration of three groups of children into their classrooms – those labeled educable retarded, emotionally disturbed, and learning disabled. The teachers were drawn from two groups: an experimental group that had the use of a resource room and a control group that had strictly self-contained classrooms. One noted difficulty with the study, observed similarly in other studies predating PL 94-142, was that many teachers appeared unfamiliar with the term “learning disabled” as their state (Pennsylvania) had used the term “minimally brain damaged” to identify the learning disabled population of children. Teachers appeared to equate learning disabled with culturally disadvantaged. Interestingly, teachers with the services of a resource room held significantly more favorable attitudes toward the inclusion of students labeled ED in their classrooms than did teachers in the control group. Only a slightly higher percentage of the experimental teachers indicated positive attitudes about the academic and social potentials of the children with ED. The researchers concluded that the teachers generally had more positive attitudes toward the children with learning disabilities compared to the emotionally disturbed and educable retarded.
Foster and Ysseldyke (1976) looked at teacher bias across four conditions – one each for hypothetical children labeled normal, mentally retarded, emotionally disturbed, and learning disabled. Having been given one of the four labels for a child, the teachers viewed a video of the same normal child and were asked to complete a referral form citing the behaviors of concern from the video. The researchers found that the deviancy labels had an effect of eliciting negative behavioral expectancies from the teachers, even when normal behavior was presented. The educable mentally retarded condition produced the strongest stigmatization effect; however, both the ED and LD labels produced bias effects as well. Highlighting perhaps a more compelling contrast between the ED and LD labels, one study found teachers much more likely to recall the ED label than the LD label when used in case reports for children (Boucher & Deno, 1979). Also, Levin, Arluke, and Smith (1982) found support for the notion that teacher expectations were much lower for the ED label than for the labels dyslexic (a specific variant of LD) and mentally retarded.

Labeling effects observed in research, however, have not always been shown to be negative for the ED label relative to other labels. In considering the possibility of differential treatment of teachers toward children according to special education label, Algozzine, Mercer, and Countermine (1977) found that teachers showed greater tolerance toward aggressive and disruptive behaviors that were tied to children labeled ED compared to children labeled LD. Thus, research has illustrated that teachers hold negative expectancies of the ED label but may be more willing to excuse or put up with severe behaviors from such labeled children. Also, greater specificity in labeling has been shown to perhaps minimize some of the perceived stigma of the ED label.
Tisdale and Fowler (1983) distinguished among teachers’ perceptions of more finely grained labels of behavioral disorders and emotional disorders and found when qualifiers of *mild* and *moderate* were employed, the labels did not appear to affect teachers’ perceptions of the prevalence of such children in their classrooms. They did, however, find teachers much more reticent to label students emotionally disordered as opposed to behaviorally disordered. In suggesting that the emotionally disordered or “emotionally handicapped” label might be anathema to many teachers, the researchers advocated that perhaps the label should be revised in the interest of serving more of those students who might be excluded for consideration of that category of eligibility merely on semantic grounds. Some of the negative attributions toward an emotional label as compared to a behavioral label may stem from the observability of behavioral problems, which renders them more readily assessed by educational professionals, whereas judgments about emotional difficulties are inherently subjective and traditionally assessed by mental health professionals. Thus, these differences in perception may stem in part from the lack of clear diagnostic criteria for the ED label (Cullinan, 2004; Kauffman, 2001).

One of the challenges in assessing attitudes toward the ED label is that the constellation of behaviors associated with the label are typically rated most disturbing by teachers. Students’ behavioral characteristics that teachers find most problematic often include acting-out (aggressive) behavior (Coleman & Gilliam, 1983). Hannah (1988) concluded that research on students with the ED label has shown views of these children as *unmotivated to learn*, *unfriendly*, *impolite*, *dishonest*, *unhappy*, *aggressive*, *unable to relate to others*, *in great need of professional help*, and as *manifesting hallucinations*,
tics, compulsions, and phobias. Therefore, in seeking out whether the label itself is stigmatizing, one must wonder whether attitude assessment is tapping into negative perceptions of a teacher’s interactions with a particular labeled student as opposed to attitudes toward the label generally.

As increasingly more research accumulated to suggest that attitudes toward the disabled are more multidimensional in nature, disparate findings regarding teacher attitudes toward students with ED could be viewed more clearly. For example, some research has suggested a relatively favorable view of the educability of students with ED compared to other disabling conditions (Kvaraceus, 1956; Morris & McCauley, 1977; Murphy, Dickstein, & Dripps, 1960; Warren & Turner, 1966). Those studies appeared to isolate the attitudes and perceptions of educators on a dimension of students’ capacity to reason and succeed academically as opposed to likability, behavioral compliance or noncompliance, and other qualities that would be more likely to trend negatively for the ED population. In such cases where a premium is placed on evaluating students’ potential for learning, it is understandable that teachers might tend to rate children with cognitive impairments as being somewhat intractable in that domain relative to children with emotional and behavioral disorders. By contrast, research that has examined attitudes of teachers and other professionals about personality characteristics of children with disabilities have consistently shown highly unfavorable attitudes toward interacting with those students with emotional and behavioral problems (Tringo, 1970) and toward the behaviors they often exhibit (Algozzine, 1977; Boucher & Deno, 1979; Coleman & Gilliam, 1983).
Graham and Leone (1987) looked at a group of 88 pre-service special education teachers’ evaluations of writing samples for four groups of children: normal, emotionally disturbed, behavior disordered, and conduct disordered. The samples were assigned randomly to one of the four expectancy conditions, and teachers were asked to score the written work after being informed of the developmental status of the supposed author. The researchers found that scores for the written products were not affected negatively by the child’s label, but rather, reflected the quality of the writing and the grader’s level of expertise.

As the 1980s drew to a close, research on the potentially stigmatizing effects of the ED label waned. The literature shifted away from looking at unfair treatment and negative attitudes toward students with EBD and began to explore the need to soften pejorative classification terms and consider alternative means of identifying and providing services for these children. In essence, it appears that educators and researchers tacitly accepted the earlier findings of the stigma of special education labels and began making efforts to ensure greater diagnostic accuracy in determining special education eligibility or to abandon the system of disability labeling altogether.

Although investigations into the potential stigmatization of the ED label trickled off dramatically since the 1970s, a few follow-up studies occurred in the ensuing years. Fox and Stinnett (1996) compared outlooks of a variety of professionals, including teachers, school psychologists, and psychologists-in-training, toward the future functioning of children with the labels of seriously emotionally disturbed (SED), conduct disordered, socially maladjusted, and no exceptionality. They found bias extended across all of the professional groupings, with respondents forecasting significantly poorer
outcomes, especially in interpersonal functioning, for those children labeled SED. Based on their findings, the authors advocated that professionals consider a label other than SED for children with behavioral problems in order to minimize labeling bias. However, the removal of the qualifier of ‘seriously’ from the ED label in the reauthorization of the Individuals with Disabilities in Education Act of 1997 (IDEA 1997) might render some of the Fox and Stinnett’s findings less relevant to current connotations of the label.

More recently, Thelen, Burns, and Christiansen (2003) investigated the effects of the LD, ED, and mild mental retardation (MMR) labels when applied to a behavioral vignette read by a sample of 409 teachers, college students, and high school students. They concluded that expectations were more negative for the ED label condition than for the other two in all three assessed domains of functioning: behavioral, academic, and interpersonal. Also of note was the finding that elementary teachers appeared to possess more favorable expectations of interpersonal functioning for the different labels than did high school teachers. The researchers suggested the difference between expectations for primary and secondary teachers might be due to the salience of social skills deficits in adolescence.

Providing a voice of opposition to those who have decried the negative effects of labeling children as ED, Kauffman (1999) has contended that fears of using the label are not justified and are not supported by research evidence. He questioned whether the antilabeling sentiment in fact acts to prevent children who are at the greatest risk of having emotional and behavioral disorders from obtaining services. Citing aggression, peer rejection, academic failure, and affiliation with deviant peers as the behaviors that signal the highest risk for development of EBD (Ialongo et al., 1998; Kazdin, 1985;
Walker, Colvin, & Ramsey, 1995), Kauffman maintained, “Eventually, many of those with emotional or behavioral disorders are formally and informally labeled unequivocally, sometimes brutally” (1999, p. 453). He implied that children who exhibit EBD, whether labeled by their school or not, will be labeled unceremoniously by critical peers, by frustrated educators, and/or by other social institutions (e.g., mental health agencies, juvenile detention facilities) who deal with the reality of antisocial or self-destructive personal behavior.

Minor et al. (2002) reported on the creation of the first scale created specifically for measuring teachers’ attitudes toward the ED (SED) label. Their stated purpose in creating the instrument was to have a tool that could be used as a criterion measure for documenting changes in attitudes toward students with ED over time. Their validation study of the *Attitudes Towards Children with Serious Emotional Disturbance Scale* (ATCSED) found that the 29-item (later pared to 23-item) instrument produced one first-order factor (a Total Score) and three second-order factors, ecology, family factors, and parental involvement. The authors indicated that the external validity of the scale remained to be established, and subsequent research using the instrument has not been found. It is curious that the authors retained the term ‘Serious’ in the labeling of the instrument, given that it had been eliminated from the 1997 reauthorization of the Individuals with Disabilities in Education Act.

**Theories of Labeling and Expectancy Effects**

As the mainstreaming of education for children with EBD was becoming a wider consideration in American schools in the 1950s, a movement in the sociological literature was beginning to question the value of labeling individuals with disabilities. Based on a
concept called *labeling theory* or *social reaction theory* (Lemert, 1951), social scientists in the 1950s and 1960s began to study the potentially aversive effects of labeling observed in society’s categorization of individuals according to their membership in various groups.

Labeling theory posits that deviance is a product of social reactions by agents of control, such as government bodies, through the process of labeling (Becker, 1963; Erikson, 1962). It essentially rejects the notion of mental illness and deviant behavior as being of intra-individual, pathological origin, but rather, considers such behaviors to be products of labeling and subsequent stigmatization (Scheff, 1999). The theory breaks from earlier views of deviance, which maintained that deviance was a property or act (Davis, 1972). Also referred to as *social behaviorism* (Quadagno & Antonio, 1975), labeling theory emphasizes social and environmental variables as the critical considerations in understanding deviance. Labeling theory has particular relevance to the ED label because the theory has traditionally targeted areas of functioning that are characteristic of students with ED, namely mental illness and conduct problems (i.e., delinquency).

Lemert (1972) contributed the ideas of *primary deviation* and *secondary deviation* to the sociological literature. Primary deviation is the initial act that draws attention to an individual from a social entity and results in a deviance label. For students labeled ED, primary deviance would be observed in their initial eligibility for special education services. Secondary deviation according to Lemert is “a special class of socially defined responses which people make to problems created by the societal reaction to their deviance” (p. 63). Thus, secondary deviance is the deviance of concern with labeling
theory, as it is contended that individuals once labeled will act in ways that meet the social expectancies tied to their deviance label and adopt a deviant identity and way of life (Warren, 1974). The concept of secondary deviance has provided the framework for criticizing the utility of diagnostic labels being applied by mental institutions and schools. It inspires the conventional wisdom that labeling children with ED (or any other special education category) might in fact have the effect of stigmatizing those children, result in others exercising prejudice according to the label as opposed to actual behavior, and reinforce the labeled children to act in ways that are consistent with expectations of those who possess the label.

Although labeling theory began as a theory to explain societal deviance in criminal justice populations, it was expanded to include related critiques on mental health/psychiatry (Goffman, 1968; Scheff, 1999; Szasz, 1961). Offered as an alternative to the mental health model of psychopathology as an intra-individual deficiency of organic origin, an “antipsychiatry” movement (Sedgwick, 1982) arose, that alleged that mental illness was largely a social construct perpetuated by profiteers within the mental health fields and control agents of government.

With the view that disabilities might in fact be socially constructed and therefore exist only because society wishes to isolate and label those who are aberrant from perceived norms, doubts began to arise as to whether therapeutic and educational services were truly benefitting individuals and might, in fact, be perpetuating problematic behaviors. The questioning of the efficacy of mental health services and the potential for such services to perhaps selectively reinforce deviant behaviors would distill into the school setting with concerns about youth labeled as delinquent and emotionally disturbed.
That is, the culture of schools appears to have demonstrated the same mistrust toward labeling of individuals with mental and behavioral problems that has been observed in other institutions and society at large. Many individuals adhere to the idea that identifying individuals for mental health and behavior management services will in actuality exacerbate a labeled individual’s problems by calling more attention to them and altering the manner in which others interact with them.

Fisher (1972) examined labeling theory in the school context in his study that compared groups of publically labeled “deviant” 7th, 8th, and 9th grade students with unlabeled students on measures of academic performance. Fisher theorized that the so-called deviant students, if labeling theory were to have an effect, would experience an “amplification” of deviant behavior, that is, the deviant label would generate significant negative consequences for the students labeled as such. Although the results indicated a positive association between the deviancy label and negative school performance, Fisher questioned whether it was the labeling per se that produced the negative reaction, and instead, noted that by accounting for the factor of ability, the effects of labeling were eliminated. Thus, the study provided some evidence that preexisting differences in ability level might account for diminished academic performance of delinquent students, calling into question labeling theory’s attribution of academic decline due to labeling effects.

A landmark study by Rosenhan (1973) was instrumental in perpetuating the notion that psychiatric labeling might be a specious thing. For Rosenhan’s study, 12 “normal” adults were admitted to psychiatric hospitals under the pretense that they were hearing voices and were displaying anxious behavior. On average, the individuals were retained in the facilities for 19 days, ranging from a minimum of 7 to a maximum of 59
days. One can conclude from the research that if individuals with merely the suggestion of mental problems could be falsely held for a lengthy period of time within an institutional setting, then there is a strong potential for clinicians to make false positive diagnoses and to over identify the population of mentally ill. It speaks to the precariousness of an inherent aspect of mental health treatment – a strong reliance on self-report or the report of others to establish the greater likelihood of meeting the criteria for mental disorders. Although the study served to stoke the flame of consciousness about the potential for misdiagnoses of mental illness under a flawed mental health system, a boom occurred in the 1970s with ever increasing numbers of individuals receiving psychiatric services.

Perhaps the most ardent critic of labeling theory, Gove (1970), suggested that the empirical evidence established that, “a person’s behavior determines the expectations of others to a much greater degree than the reverse” (p. 882). Although he acknowledged that labeling theory had played a “critical role” in the process of deinstitutionalization of patients from mental institutions starting in the mid-1950s, Gove (1982) argued that the psychiatric “revolution” served to undercut labeling theory, as individuals demonstrated responsiveness to the treatments and appeared to exhibit few signs of being perpetually cast as mentally ill following the amelioration of symptoms. Furthermore, criticisms aimed at the imprecision and vagueness of psychiatric diagnostics were addressed with the development of the *Diagnostic and Statistical Manual of Mental Disorders, Third Edition* (DSM-III) in 1980. Spitzer and Williams (1982) heralded the advancements in the DSM-III as including more specific diagnostic criteria and the provision of a multiaxial system for evaluation.
Smith, Osborne, Crim, and Rhu (1986) looked at labeling theory as it applied to the application of the LD label. They compared differences of perceptions among special educators, parents, and school officials about the meaning of the label and found that there was considerable variance among the different roles as to how they defined the label. For example, some individuals provided descriptors more appropriate for classifications of retardation, socio-emotional disorders, and hyperactivity. The researchers noted that school personnel often characterized the LD label as including behavioral and socio-emotional problems, whereas parents ranked physical disorders and attention-span problems as the top descriptors. Of interest to the discussion of who plays the strongest role in affixing a label to a child, it was found that school psychologists were mentioned most frequently and regular teachers a close second as to who had the most influence in the process of labeling.

Despite evidence that labeling theory was showing a decline as a comprehensive and predictive theory, it nonetheless had left its imprint on a variety of social institutions, with schools being the most relevant for this current discussion. A body of literature that emerged concurrently with that of the labeling theorists, and in some ways showed congruence with its concerns, was that of expectancy effects. Wodarski (1977) maintained, in fact, that expectation theory provided much of the empirical support for labeling theory. Expectancy effects within the context of research studies had been known and explored empirically for a long time, but the application of such effects to the school context became a serious interest in the 1950s and beyond. For example, as early as 1939, Roethlisberger and Dickson’s study of industrial behavior established an expectancy effect that would become known as the Hawthorne Effect. Based on a series
of experiments done with factory employees, the researchers found that changes made to the work environment (e.g., dimming lights, brightening lights) seemed to boost worker productivity. Thereafter, it was speculated that participants who are involved in an experimental process may show some positive reactivity simply because they are receiving attention and are perhaps recognizing an expectancy for improvement.

Where labeling theory was concerned with labeling as the “selective reinforcement of one’s deviant behaviors by significant others…resulting in an increased frequency of subsequent deviant behaviors” (Wodarski, p. 123), the research on expectancy effects looked at how and why others (e.g., caretakers, teachers, etc.) adopted false expectations about individuals’ capacities to perform various acts and thus set in motion behaviors that caused the expectations to be reinforced – so-called “self-fulfilling prophecy” effects (Merton, 1948). In a review of the research about self-fulfilling prophesies, Jussim (1986) summarized the concept as “situations in which one person’s expectations about a second person lead the second person to act in ways that confirm the first person’s expectation” (p. 429). Derived from the Thomas theorem, which holds that “if men define situations as real, they are real in their consequences” (Thomas & Thomas, 1928), the notion of a self-fulfilling prophesy became a descriptor for conditions in which individuals might be prone to perform adversely on a task if there is a negative expectancy for their performance, or conversely, perform to a higher standard under circumstances when there is an expectancy of stronger performance.

A convergence of labeling theory and expectancy theory was seen most notably in Pygmalion in the Classroom, the seminal work by Rosenthal and Jacobson (1968). Their study (Rosenthal & Jacobson, 1966) was conducted in 18 classrooms across 6 grades.
The researchers had students take a nonverbal intelligence test early in the year and then again eight months later. The test was administered under the guise of being a test that was predictive of academic “blooming,” and approximately 20% of the children were identified for teachers as being likely to make significant academic gains in their classrooms. The study concluded that the children labeled bloomers evidenced significantly higher gains on retesting of IQ than did the control children. Thus was born the notion that teacher expectancies alone could have quite dramatic effects on student behaviors, whether they be performance on IQ tests, achievement tests, or simply teacher-student interactions. In making a compelling case that merely labeling certain students as “bloomers” helped to facilitate interactions between teachers and students that would enhance intellectual growth, the expectancy literature branched out from Rosenthal and Jacobson’s work to look at how expectancies of students with disabilities might be established unfairly through the process of labeling.

Although the Rosenthal and Jacobson study was highly influential, it had its share of critics on the grounds of methodological concerns (Barber & Silver, 1968; Elashoff & Snow, 1971; Thorndike, 1968). Elashoff and Snow (1971) illustrated that follow-up studies using IQ as a dependent variable did not replicate the findings of Rosenthal & Jacobson. However, the central idea of the research would be carried out in a multitude of other studies, typically using achievement scores (Brophy & Good, 1970; Meichenbaum, Bowers, & Ross, 1969; Palardy, 1969; Seaver, 1971) and teacher behaviors (Good, 1970; Simon, 1969) as dependent variables. In summarizing the overall findings of 20 years of research on teacher expectancy effects following the Rosenthal and Jacobson study, Good (1987) stated that there existed “a consensus that teachers’ expectations can and
sometimes do affect teacher-student interaction and student outcomes, along with the recognition that the processes involved are much more complex than originally believed” (p. 33). Good conceded that low teacher expectations might reflect inappropriate knowledge of how to work with students who have learning difficulties as opposed to being a product of simply labeling students.

As the prevailing wisdom suggested that labels hurt and stigmatize individuals, forces for reform within the educational environment began to advocate for the abolition of the system of labeling and segregated special education services altogether. One of the most well-known efforts to fully integrate regular education with special education was the Regular Education Initiative (Reynolds, Wang, & Walberg, 1987). Others (Stainback, Stainback, & Ayres, 1996) have argued similarly for a unitary construction of educational services and placements under the name of the Full Inclusion Movement. Under a system of full inclusion, categorical labels for students would be eliminated, and students of all levels of ability and functioning would be placed in the same classrooms.

While labeling theory was being scrutinized and exposed for it inadequacies as a comprehensive theory of deviance, a related ideology, postmodernism, was rising in prominence in sociological and educational fields. Postmodernism has been described in the following manner:

The ideology of postmodernism is difficult to describe clearly…but it is generally the notion that objectivity is unbelievable and impossible, that truly scientific investigation is therefore impossible or irrelevant, that the scientific notions of universality and predictability have been discredited, and that subjective or
“situated” knowledge should be considered superior to all other “ways of knowing.” (Kauffman, Brigham, & Mock, 2004, p. 25)

Postmodern ideology would feed into, and add to, perceptions that labeling of any kind is unacceptable on the grounds that attempts to categorize others are the manifestations of an objective view of the world, of hard singular truth. The postmodern perspective dictates a relativistic, subjective notion of truth, calling for the revoking of not only categories of disabilities but of broader social identifiers such as race and gender as well. Carrier (1983) applied postmodernist thinking to the LD label, contending that the view of the label as descriptive of a real condition was mistaken and was better explained by social forces.

A number of voices have arisen in the special education and behavioral disorder literature advocating from a postmodern perspective (Danforth & Rhodes, 1997; Elkind, 1998; Gallagher, 2004). Skrtic (2005) stated his objective as a “critical theorist” as follows: “My goal is to disrupt this power relation, to deconstruct it and the social categories it creates by exposing inconsistencies, contradictions, and silences in special education's functionalist knowledge tradition, and by disseminating alternative interpretations of special education and student disability” (p. 149). In this manner, the debate about the stigmatization of special education labels has begun to move away from arguments based on the social and academic fallout or effects of labels on the children and has encompassed an alternative worldview that calls into question whether disability is a real or useful construct.

Voices of opposition in special education and psychology have warned of the dangerous consequences of adhering to multirealist (postmodern) philosophy (Krueger,
2002; Locke, 2002; Mostert, Kauffman, & Kavale, 2003; Sasso, 2001). Kauffman (1999) has faulted the ideology foremost on the basis that it serves to hinder efforts to prevent the development of emotional and behavioral disorders. He notes, “Prevention thus becomes at once a primary object and impossible; we cannot prevent what we are unwilling to categorize as different from the typical or normative, nor can we practice anything other than primary prevention if we are unwilling to categorize interventions as special (not normative)” (p. 455). Thus, the climate of concern about identifying children with ED persists with questions about stigmatizing effects but also has grown larger to question the validity of the label and its usefulness for serving any children at all.
CHAPTER 3

Method

Participants

For this study a stratified sample of near equivalent numbers of regular education and special education teachers was sought out (see Table 1, for characteristics of sample). Additionally, efforts were made to survey both rural and urban educators.

Table 1

Frequencies of Individual Characteristics within the Sample

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Regular Educators (n = 52)</th>
<th>Special Educators (n = 46)</th>
<th>Total Sample (n = 98)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>42</td>
<td>36</td>
<td>78</td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Setting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>25</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>Urban</td>
<td>27</td>
<td>21</td>
<td>48</td>
</tr>
<tr>
<td>Education level completed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelors degree</td>
<td>13</td>
<td>13</td>
<td>26</td>
</tr>
<tr>
<td>Some graduate work</td>
<td>10</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>Masters degree</td>
<td>28</td>
<td>28</td>
<td>56</td>
</tr>
<tr>
<td>Other advanced degree</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Teaching experience (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1-5</td>
<td>11</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>6-10</td>
<td>12</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>11-15</td>
<td>6</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>16-20</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>20+</td>
<td>17</td>
<td>14</td>
<td>31</td>
</tr>
<tr>
<td>School level of instruction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary</td>
<td>36</td>
<td>21</td>
<td>57</td>
</tr>
<tr>
<td>Middle</td>
<td>9</td>
<td>13</td>
<td>22</td>
</tr>
<tr>
<td>High</td>
<td>7</td>
<td>12</td>
<td>19</td>
</tr>
</tbody>
</table>
The sample consisted of 98 teachers (78 women, 20 men, $M_{age} = 43.3$ years, age range: 22-65 years) who were recruited through email or letters placed in their school mailboxes. All participants were drawn from a consortium of school districts in central Indiana; one urban district and two rural districts were included.

**Measures**

**Explicit measure of attitudes toward ED and LD labels.** A primary factor of interest was explicit attitudes of educators toward the ED label relative to the LD label. The instrument for assessing attitudinal differences between those two labels was a modified version of Tringo’s (1970) *Disability Social Distance Scale* (DSDS) and Crandall’s (1991) *Social Distance Questionnaire* (see Appendix A). The DSDS was selected for several reasons, including its relatively parsimonious manner of estimating differences in ED/LD attitudes, the flexibility of administering the measure to groups or individuals, the relevance of social distance as a manifestation of stigmatizing attitudes toward labeled children, and the opaqueness of the measure in allowing respondents to rate numerous disability categories as opposed to singling out LD and ED specifically.

The format of the DSDS is practical in answering the overarching research question of whether there is greater stigma attached to ED versus LD, a question which would be difficult to answer with measures aimed at attitudes toward one particular disability and providing no referent disability. Furthermore, the adoption of the DSDS follows Yuker’s prescription (Antonak, 1988, p. vi) that when possible it is preferred to improve upon old scales rather than develop new ones.

Antonak (1988) provided a critique of the DSDS in which he called for future researchers to make several changes to the instrument, including updating the list of
disabilities to reflect more contemporary terms and eliminating the confusion of mixing terms that represent names of disabling conditions (e.g., epilepsy) with identifiers of disabled individuals (e.g., paraplegic). Accordingly, the list of disabilities was changed to match several categories of special education classification included in the most recent federal legislation (IDEIA 2004). Care was taken to ensure that the terms reflected the disability label (e.g., learning disability) as opposed to a descriptor of an individual (e.g., learning disabled). The finalized list of disabilities for this measure included autism spectrum disorder, blind or low vision, deaf or hard of hearing, emotional disturbance, language or speech impairment, learning disability, orthopedic impairment, and traumatic brain injury. Categories perceived to be inherently vague, broadly-defined, or serving as catch-all categories, such as multiple disabilities and other health impairment, were excluded from this measurement.

One element of the DSDS that appeared somewhat anachronistic and unsuitable for the purposes of the present study was the wording of the item ratings. The measure was set up for respondents to grade particular disabilities on a nine point scale of degrees of social proximity. On the extremely negative or more distancing end of the scale, the option read, “Would put to death.” It was determined that such a response reflected outmoded notions about the treatment of individuals with disabilities and would be unlikely to discriminate respondents in any meaningful way. Turning toward a more recent revision of the Social Distance Scale, Crandall’s (1991) examination of stigma and attitudes toward homosexuals and drug users featured more contemporary language in its scale items. Because Crandall’s *Social Distance Questionnaire* is structured to accommodate or follow up on respondent’s sentiments toward behavioral vignettes, the
scale items were modified to be more appropriate for descriptions of special education labels rather than being applied to specific cases and individuals. The seven scale options included (a) spouse, (b) close personal friend, (c) co-worker, (d) neighbor, (e) I would tend to avoid, (f) I would prefer to be segregated from other people, and (g) I believe should be kept in an institution.

Scoring for the modified Social Distance Scale was done on a seven-point scale, with seven points given for ‘spouse,’ six for ‘close personal friend,’ continuing down to one point for ‘I believe should be kept in an institution.’ Thus a higher score on the SDS conveyed a stronger willingness to affiliate with, or have a more intimate relationship with, a person with a particular disability as compared to lower scores suggesting social aversion to such individuals. Because the two scale items of interest for this study were the impressions of social distance toward the Emotional Disturbance and Learning Disability labels, the items rating social distance toward individuals with ED and LD provided the units of comparison both within this scale and relative to the implicit attitude scores.

Explicit measure of general attitudes toward disabilities. In order to allow comparison between general attitudes toward disabilities with specific attitudes toward ED and LD, the Scale of Attitudes Toward Disabled Persons, Form R (SADP-R) was chosen (Antonak & Livneh, 1988). Whereas the explicit ED/LD attitudinal measure assessed the construct of social distance, the SADP provides more varied data to investigate attitudes toward persons with disabilities as a group. The SADP was constructed based on interviews with “experts in the fields of special education and rehabilitation” (Antonak, 1988, p. 139), thus rendering item content particularly well-
suited for administration to teachers. The SADP (see Appendix A) consists of 24 statements to which respondents may rate their agreement on a six point scale from -3 (“I disagree very much.”) to +3 (“I agree very much.”).

Antonak (1985) structured the SADP-Form R such that for half of the items a positive response (e.g., +3, +2, or +1) would convey a positive attitude whereas for the other half, a negative response (e.g., -1, -2, or -3) would be suggestive of a positive attitude. The scoring algorithm for the SADP-Form R involves several steps, including first reversing the sign of the negatively-worded items either from positive to negative or from negative to positive, then summing the values for the adjusted scores, and finally, adding a constant of 72 to the total to avoid producing a summary score with a negative value. The range of possible total scores on the SADP-Form R is 0 – 144 with higher point totals being indicative of more positive attitudes.

The SADP-Form R affords an additional layer of analysis with the identification of three subscales: Optimism-Human Rights, Behavioral Misconceptions, and Pessimism-Hopelessness. These subscales are calculated by reversing the sign of the negatively-worded item responses and summing the signs of the responses. For the first subscale, Optimism-Human Rights, the possible range of initial scores for the 11 items is -33 to +33. Subscale II, Behavioral Misconceptions, consists of 7 items; thus, scores may range from -21 to +21. Lastly, the Pessimism-Hopelessness subscale with its 6-item total may range from -18 to +18. Consistent with the procedure used to derive the overall SADP score, constants are added to eliminate negative scores on the subscales; therefore, final scores are transformed by adding either a constant of 33, 21, or 18 for the respective
The subscale. After adjusting the scores, they may range from 0 to 66 for subscale I, 0 to 42 for subscale II, and 0 to 36 for subscale III.

Antonak (1982) reported on analyses of construct validity and external validity in the creation of the SADP. Construct validity was looked at in the scale’s capacity to discriminate between populations that were theorized to have variance in their attitudes. The researcher found no significant difference between scores of males and females; however, data showed that SADP scores correlated significantly with the variables of educational level and age. External validity was examined by comparing items from the SADP with psychometric data reported by Yuker, Block, and Younng (1966) for the Attitudes Toward Disabled Persons – Form O (ATDP-O). Antonak discovered in the item analysis of the two scales that the ability of the SADP items to discriminate between low and high scores on the ATDP-O was low to moderate. The researcher observed that one quarter of the items on the ATDP failed to significantly discriminate overall scores on the ATDP itself. A correlational value of .54 was found between the SADP and ATDP, suggesting that the scales assessed a common construct. Scale reliability (.81) and internal consistency (.88) were described as acceptable for the SADP.

Implicit attitudes toward Emotional Disturbance label. A second primary factor for this research is implicit attitudes toward ED. This factor was measured with an Implicit Association Test (IAT), which was constructed to calculate the latency of participants’ responses when pressing a key on the keyboard to associate stimulus words on the computer screen.

Following a prescribed structure and sequence outlined by Lane et al. (2007), the administration of the implicit measure consisted of seven stages of responding. The
participant reviewed a screen with task instructions for completing the IAT, which included references to which two keystrokes to use that corresponded to a word or pair of words in the upper left or right side of the screen (see Appendix A).

The first stage in the 7-block sequence was an initial target-concept discrimination stage. This provided the participant with training to discriminate between terms descriptive of emotional disturbance versus terms descriptive of learning disability. The orientation of the terms emotional disturbance and learning disability was alternated according to the participant id#, whether even or odd. For odd participants, the words Emotional disturbance appeared in the upper left corner of the screen and the words Learning disability appeared in the upper right corner of the screen for the initial target-concept discrimination task, whereas their positions were flipped for even-numbered participants.

In order to assure that the number of descriptive words (i.e., stimulus items) used to represent ED and LD would meet the minimum standard for establishing the desired effect magnitude and reliability, five terms were identified for each target concept based on the federal definitions as well as research about teachers’ perceptions of those characteristics which are most descriptive of students labeled ED or LD (Boucher & Deno, 1979). For emotional disturbance, the stimulus words selected include depression, peer problems, disruptive behavior, aggression, and self-harm. For learning disability, the terms include dyslexia, brain dysfunction, language disorder, perceptual problems, and letter reversals. Nosek, Greenwald, and Banaji (2005) found that IAT effects emerged with as few as two stimulus items per target concept, with the effect being “only slightly less robust” than the use of eight items per concept. Greenwald et al. (1998) had
observed on the higher end that effect magnitudes were essentially the same when using five stimulus items as compared to 25. Following the prescription of Nosek et al. (2005), fewer items of stronger representation were chosen for the target categories as opposed to numerous items of weaker quality, so as to lend improved construct validity to this measure. Furthermore, attribute stimuli were constructed exclusively as nouns in accordance with findings that evaluative nouns produced significantly higher correlations between IAT and self-report data compared to evaluative adjectives and thematic words (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2004).

As the participant was exposed to the target-concept discrimination phase, words associated with either ED or LD appeared in the center of the screen one at a time. For example, if the word *aggression* flashed in the middle of the screen, the participant was expected to press a key to the left which corresponded to the presence of the words *Emotional disturbance* in the upper left corner of the screen. Conversely, if the term *dyslexia* appeared in the middle of the screen, the participant was expected to press the key to the right, corresponding with the *Learning disability* marker in the upper right corner. The program provided corrective feedback to the participant in those instances in which the stimulus word in the middle of the screen was paired incorrectly with a concept. Error feedback was indicated with a red X in the middle of the screen when an error occurred. Overriding the error involved having the participant select the other key to make the correct association between stimulus word and concept.

In the second stage, what is referred to as an *associated attribute discrimination* procedure was administered. This involved the presentation of terms related generally to perceptions of good versus bad. The pleasant and unpleasant stimulus words were taken
directly from previous studies of the IAT (Swanson et al., 2001) and included the following “pleasant” terms: cuddle, happy, smile, joy, warmth, peace, paradise, and love, and the following “unpleasant” terms: pain, awful, disaster, grief, agony, brutal, tragedy, and bad. These words appeared one at a time in randomized order in the middle of the screen, and the respondent pressed a key to the left or the right to associate the stimulus word with either the word pleasant in the upper left corner or unpleasant in the upper right corner.

The third stage was titled an initial combined task. This task presented two words in both the upper left and right corners. For example, the participant could see Emotional disturbance and pleasant paired in the upper left corner and Learning disability and unpleasant paired in the upper right corner of the screen. The IAT was designed to alternate the presentation of the concept words for this stage, depending on the participant’s id number. For odd-numbered participants, Emotional disturbance appeared in the left corner. For even numbered participants, Emotional disturbance appeared in the right corner. Stimulus words related to ED, LD, pleasant, or unpleasant were displayed one-by-one in the center of the screen. The individual was required to press a key to associate the words again with either the left or right side. The word pleasant or unpleasant appeared beneath either Emotional disturbance or Learning disability in the upper corners of the screen.

The fourth stage replicated the third, providing an additional trial of the same task. Thus, where the third stage provided a practice set to familiarize the participant with the combined task, the fourth stage followed up to obtain a more extensive collection of data for comparative purposes.
The fifth stage provided a reversal of the first stage, namely a reversed target-concept discrimination task. The screen reverted to having only one word in the upper left and one word in the upper right corner, with the positions of the terms Emotional disturbance and Learning disability flipped to opposite sides of the screen from their prior presentation. This procedure helped to control for potential differences in responding with one hand versus the other. As with the earlier stage, words populated the center of the screen one at a time, and the participant pressed either the “E” key to associate the word with the concept in upper left corner or the “I” key to associate the word with the concept in upper right corner.

The last two stages, six and seven, featured reversed combined tasks. As with stages three and four, two words appeared in the upper left and right corners. This time, the concepts of Emotional disturbance and Learning disability were paired with the opposite attribute from the earlier pairing. In other words, if Emotional disturbance was paired with pleasant in the 3rd and 4th stages, then it was paired with unpleasant in the 6th and 7th. Upon completion of the 7th stage, the participant was finished with the IAT. On average it was estimated that the entire IAT took approximately 10 minutes for participants to complete.

Participants’ latencies of response to all stimulus words as well as their errors were saved to a log file on the host computer for each session and with a unique identifier to map the IAT data to a participant’s demographic information. From the latency data, the researcher derived a D-score for each participant according to conventions established by Greenwald, Nosek and Banaji (2003). Those researchers recommended the IAT be scored in the following sequence:
1. Delete trials greater than 10,000 milliseconds
2. Delete subjects for whom more than 10% of trials have latency less than 300 milliseconds
3. Compute the “inclusive” standard deviation for all trials in Stages 3 and 6 and likewise for all trials in Stages 4 and 7.
4. Compute the mean latency for responses for each of Stages 3, 4, 6, and 7.
5. Compute the two mean differences \((\text{Mean}_{\text{Stage6}} - \text{Mean}_{\text{Stage3}})\) and \((\text{Mean}_{\text{Stage7}} - \text{Mean}_{\text{Stage4}})\).
6. Divide each difference score by its associated “inclusive” standard deviation
7. \(D = \) the average of the two resulting scores derived in step 6.

The value for a \(D\)-score is expressed in positive or negative terms, depending on whether the implicit association is favorable or unfavorable for one concept compared against another. The \(D\)-score is most often interpreted like an effect size of implicit preference. Per Greenwald, Nosek, and Banaji (2003) a common convention used with web-administered IATs has been to share with participants a descriptor of the magnitude of implicit preference (e.g., slight, moderate, large) at the conclusion of an IAT. For purposes only of informing test-takers of potential implicit bias, scores between 0.15 and 0.35 have been described as representing a “slight” preference for one concept versus another; scores in the range between 0.35 and 0.65 reflect a “moderate” preference; and scores of 0.65 and above are indicative of a “strong” preference. In this study, the preference of interest was learning disability compared to emotional disturbance; thus, positive scores in the aforementioned ranges were suggestive of levels of implicit preference for the concept of learning disability over emotional disturbance.
Demographics. Other variables of interest that were recorded included the participant’s occupation (regular education or special education teacher), school setting (i.e., rural or urban), number of years spent working/teaching in a school environment, level of school currently taught (i.e., elementary, middle or high school), amount of education, and gender.

Procedure

After IRB approval was obtained for this study, permission to solicit involvement in the study was sought from either building principals or special education directors within the three districts surveyed. A standard recruitment email (see Appendix B) was sent out to all teachers listed with an email address on the school website. For buildings that showed limited response to the email solicitation within approximately a week’s time, a letter was placed in teachers’ mailboxes at the school.

For teachers who volunteered to participate, times for administration of the survey were scheduled either before or after school in their classrooms. The researcher reported to the office of the school and joined the teacher in every instance in a classroom setting. While the laptop computer was powered on and the Inquisit software loaded, teachers were given a consent form to read and sign (see Appendix C).

The demographic survey, IAT, and explicit measures (in that order) were administered via an Inquisit script launched on the laptop computer. Refer to Appendix A for the sequence of screens encountered by participants. Although some research has shown that the order of IAT and explicit measures does not have any significant effect on correlations of the two measures (Hofmann et al. 2004), other researchers have recommended administering the implicit measure first so as to avoid priming the
participant’s responses (White, Jackson, & Gordon, 2006). Nosek, Greenwald, and Banaji (2005) confirmed that administering the IAT first “does not appear to induce reactance or assimilation tendencies in subsequent self-report” (p. 176). Therefore, this study adhered to the convention of administering the IAT prior to the explicit measures.

After completing each of the measures on the host computer, participant data were saved automatically to Inquisit log files. Periodically, the data were backed up to a secure university server. As an inducement for taking part in the study, teachers were informed that they would have an approximate 1 in 25 chance of receiving 100 dollars if they completed the 15-20 minute administration of a demographic profile and three attitudinal measures. The names of one special educator and one regular educator from both the rural and urban groups were drawn at random at the conclusion of the study and awarded the incentive.

In order to compile the data, all of the information from the Inquisit files was imported into a spreadsheet. Although each of the computer-based attitudinal measures produced a distinct file, unique user ids connected data for each of the participants across each of the files. Once the data were aligned in the spreadsheet according to participant id, the researcher examined the data for anomalies or missing information. After copying the spreadsheet data into SPSS, the statistical analyses were performed.
CHAPTER 4

Results

After collecting data for implicit attitudes of teachers toward the ED label relative to the LD label, general attitudes toward individuals with disabilities, and social distance ratings of individuals with ED and LD, data were analyzed to see if meaningful differences were present across the measures and across key demographic areas. These analyses of implicit and explicit views looked at descriptive information for the sample of teachers as a whole as well as across subgroups. The comparisons of interest included special education vs. regular education teachers and urban vs. rural teachers. The following results are discussed according to the guiding questions in conducting the study.

Question 1a: Do educators hold more negative perceptions of the ED label compared to the LD label? Several instruments were used to estimate the perceptions of educators toward the ED and LD special education categories. First, the explicit attitudes toward ED and LD were elicited with ratings of social distance or the closest relationship participants would have with a person with Emotional Disturbance or with a Learning Disability. As Table 2 shows, the mean social distance rating given for persons with Emotional Disturbance ($M = 4.86$) was lower than the mean rating for persons with Learning Disability ($M = 6.52$). The median social distance score for ED was a 5, which equated to “co-worker,” whereas the median social distance score for LD was 7
Thus, the median score suggests that the closest relationship many teachers would maintain with an individual with ED would be as a co-worker. In looking at the range of responses, the most socially distant rating given for a person with ED was a 2 (“I would prefer to be segregated from other people”) as compared to the most extreme rating for a person with LD, which was a 4 (“neighbor”). A paired samples \( t \)-test was performed that indicated that educators’ responses to the social distance measures were meaningfully different when rating ED compared to LD. Results suggest a greater social aversion to ED relative to LD (\( t = 12.41, p < .001 \)).

Table 2

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Range</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAT (D-score)</td>
<td>98</td>
<td>2.04</td>
<td>-0.64</td>
<td>1.39</td>
<td>0.57</td>
<td>0.40</td>
</tr>
<tr>
<td>SADP</td>
<td>98</td>
<td>76.00</td>
<td>60.00</td>
<td>136.00</td>
<td>106.78</td>
<td>14.28</td>
</tr>
<tr>
<td>SDS ED</td>
<td>98</td>
<td>5.00</td>
<td>2.00</td>
<td>7.00</td>
<td>4.86</td>
<td>1.31</td>
</tr>
<tr>
<td>SDS LD</td>
<td>98</td>
<td>3.00</td>
<td>4.00</td>
<td>7.00</td>
<td>6.58</td>
<td>0.64</td>
</tr>
</tbody>
</table>

*Note.* IAT = Implicit Association Test; SADP = Scale of Attitudes Toward Disabled Persons; SDS ED = Social Distance Scale, Emotional Disturbance item; SDS LD = Social Distance Scale, Learning Disability item.

One finding of note can be seen in the correlational data between mean social distance scores for ED and mean social distance scores for LD (see Table 3). A coefficient of 0.14 suggests no trend between participants’ responses on one item relative to the other. Thus, the differences between the mean social distance responses to ED and LD may be significantly different, but respondents were perhaps as likely to select a relationship of similar social distance on both items as they were to select a more distant relationship for one disability compared to the other.
### Table 3

**Summary of Intercorrelations for Scores on the SADP, SDS ED, SDS LD, and IAT for Entire Sample.**

<table>
<thead>
<tr>
<th>Measure</th>
<th>SADP</th>
<th>SDS ED</th>
<th>SDS LD</th>
<th>IAT (D-score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SADP</td>
<td>1.00</td>
<td>0.24*</td>
<td>0.21*</td>
<td>0.18</td>
</tr>
<tr>
<td>2. SDS ED</td>
<td>0.24*</td>
<td>1.00</td>
<td>0.14</td>
<td>-0.06</td>
</tr>
<tr>
<td>3. SDS LD</td>
<td>0.21*</td>
<td>0.14</td>
<td>1.00</td>
<td>0.16</td>
</tr>
<tr>
<td>4. IAT (D-score)</td>
<td>0.18</td>
<td>-0.06</td>
<td>0.16</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Note.* For these scales, higher values represent stronger agreement between measures in the direction of the assessed construct (i.e., attitude or preference toward disability). SADP = Scale of Attitudes Toward Disabled Persons; SDS ED = Social Distance Scale, Emotional Disturbance item; SDS LD = Social Distance Scale, Learning Disability item; IAT = Implicit Association Test.

* *p* < .05 (2-tailed).

Participants’ latency responses in milliseconds were recorded for all trials of an implicit attitudinal measure which looked at preference for the Learning Disability label compared to the Emotional Disturbance label. According to the scoring procedure prescribed by Greenwald, Nosek, and Banji (2003), all trials greater than 10,000 milliseconds were first deleted and then data were reviewed to see if any participants had more than 10% of trials with latencies below 300 milliseconds. The researcher would have been required to omit all such participant data which met the threshold of 10% or more trials below 300 milliseconds; however, in this study, none of the 98 respondents to the implicit measure reached that threshold. These scoring procedures allowed for the removal of data that outwardly bore the appearance of being compromised due to excessively rapid or slow participant responding. The lack of such indicators of compromised data in this study’s sample suggests that all participants performed the
implicit attitudinal tasks within the desired range of responding one would wish to see for validity purposes.

A one-sample $t$-test of the implicit measure $D$-scores was conducted to determine whether the mean implicit score was significantly different than a zero value. The result suggested that the mean $D$-score was significantly different from zero ($t = 14.08, p < .001$). The difference score ($D$-score, see Table 2) was calculated based on mean latency differences between conditions suggestive of a negative association or positive association with LD as compared to negative or positive association with ED. The mean $D$-score for the sample, 0.57, can be interpreted as representing a moderate preference for LD over ED. Such a score falls in a range that Nosek, Greenwald, and Banaji (2005, p. 168) have reported in large sample studies for well-established implicit preferences for white over black ($D$ mean effect = 0.52) and for young over old ($D$ mean effect = 0.62).

The administration of the implicit attitudinal measure was varied according to participants’ id numbers. For odd-numbered participants, an “unfavorable” stage of stimuli presentation was administered prior to a “favorable” stage. The unfavorable stage required participants to match a stimulus word in the middle of a computer screen with either the left side, which featured the words Emotional Disturbance paired with the word pleasant. This was considered the unfavorable stage, because it is hypothesized that it would be more implicitly difficult, and thus more time consuming, for respondents to make positive associations with ED. The favorable stage, on the other hand, presented the words Emotional Disturbance with the word unpleasant, and thus, would be hypothesized to be an association that would be easier to make and thus occur more rapidly. For even-numbered participants, the favorable stage preceded the unfavorable
stage. By alternating those stages among even- and odd-numbered participants, an examination of potential order effects was made possible.

Looking at Table 4, the mean latency differences between unfavorable and favorable stages is consistent with the hypothesis of greater latency in making positive associations with the ED label as compared to the LD label. The mean latency difference score for odd-numbered participants is positive, indicating a much higher number ($M = 236.12$) –that is, a longer latency of responding–for trials in the unfavorable condition as compared to trials in the favorable condition. By contrast, the mean latency difference score is negative ($M = -452.77$) for the even-numbered participants who completed the favorable trials before the unfavorable trials, suggestive of quicker reactions (i.e., lower latency of response) when making negative associations with ED as compared to positive associations.

Table 4

*IAT Difference Scores by Order of Presentation*

<table>
<thead>
<tr>
<th>IAT Latency Differences (milliseconds)</th>
<th>Mean</th>
<th>SD</th>
<th>$D^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional Disturbance + Pleasant (L) or Learning Disability + Unpleasant (R) before Learning Disability + Pleasant (L) or Emotional Disturbance + Unpleasant (R) [Odd-numbered participants]</td>
<td>236.12</td>
<td>238.92</td>
<td>-0.55</td>
</tr>
<tr>
<td>Learning Disability + Pleasant (L) or Emotional Disturbance + Unpleasant (R) before Emotional Disturbance + Pleasant (L) or Learning Disability + Unpleasant (R) [Even-numbered participants]</td>
<td>-452.77</td>
<td>360.64</td>
<td>0.59</td>
</tr>
</tbody>
</table>
Note. Means are for differences in average latency scores for combined and reversed combined IAT tasks. Positive scores suggest that the respondent responded more slowly on average for the first of the two stages shown. A higher value (i.e., slower response time) is indicative of a more negative implicit association for the initial pairings as compared to the second pairings. A negative score suggests respondents on average demonstrated a faster reaction time for the first group of pairings and thus presented with an implicit preference for making those associations as opposed to the second grouping.

The $D$-score represents the mean effect for participants who were administered the IAT in the described order.

In order to determine whether the order of administration of the favorable or unfavorable conditions was significant, two paired sample $t$-tests were performed to compare participants’ mean latency scores across the favorable and unfavorable stages of responding. For the favorable condition, results were insignificant ($t = -.30, p = .76$), which implies that the Learning Disability+pleasant/Emotional Disturbance+unpleasant stage was responded to in similar fashion by both odd-numbered and even-numbered participants. Results for the unfavorable condition, however, established that there was a significant difference between the mean latency scores of the two groups ($t = -.34, p < .01$). A review of the data showed that the group that was administered the favorable stage before the unfavorable stage took significantly longer on the unfavorable stage than did the group that was administered the unfavorable stage first. In summary, there appears no evidence that the ordering of the favorable stage created any meaningfully different outcome in response times for participants, but the ordering of the unfavorable stage first may have contributed to longer average latency of responding for participants in that group. In terms of the overarching research question, however, the results are indicative of significantly higher implicit negativity toward the ED label relative to LD.
Question 1b: Do educators’ general attitudes toward individuals with disabilities correspond with their attitudes toward individuals with the ED label?

By looking at the correlation between scores on the general measure of attitudes toward disabilities (SADP) and the scores on the explicit attitudinal measure of ED, there is evidence of a mild but significant correlation in responses to those measures. As Table 3 shows, the comparison between SDS ED and SADP scores yielded a coefficient of 0.24, which is significant at the .05 level. Therefore, one can conclude that individuals who possessed more positive attitudes toward individuals with disabilities generally tended to express a willingness to maintain a more favorable social distance with individuals with ED. Conversely, those who evidenced a more negative overall disposition toward individuals with disabilities were more likely to rate a stronger social aversion toward individuals with ED. This effect of corresponding general and specific attitudes was observed in the comparison of general attitudes and explicit attitudes toward LD as well ($r = .21, p < .05$).

Question 1c: Is there agreement between educators’ implicit and explicit attitudes toward the ED label? Analyses for comparing the degree of concordance between implicit and explicit attitudes required some manipulation of data in order to place data from the SDS and the IAT measures on a similar, standardized scale. Given that scores for both measures could be represented with magnitudes within a circumscribed range, difference scores for the SDS (subtracting the lowest score from the highest) were used to create a scale from the lowest SDS rating to the highest, with the mean difference score representing the middle of the distribution. Similarly, a scale was created for the IAT scores, using the low and high scores for the poles of the distribution.
and the mean score for the center. With ratios of proportionality in place, the values for
the SDS scale were then transformed into the approximate equivalents on the IAT scale,
thus allowing for a comparison of like means. A t-test was performed comparing the
mean D-score \((M = .57)\) to the transformed mean SDS difference score \((M = .48)\). The
result \((t = 1.68, p = .10)\) suggests there is no significant difference between the explicit
and implicit attitudes of these teachers, which is consistent with the individual analysis of
both measures that showed moderate preference for the Learning Disability label over the
Emotional Disturbance label.

**Question 2: Are there significant differences between explicit and implicit
attitudes of regular education versus special education teachers regarding the ED
label?** In order to address this question, a couple of analyses were performed looking at
differences between variables according to occupation (i.e., regular educator or special
educator). First, independent sample \(t\)-tests were conducted to look at differences in mean
scores for each of the measures (see Table 5) according to status of participants as special
educator or regular educator. For the measure of attitudes toward persons with disabilities
generally, the SADP, results between scores for regular and special education teachers
were found to be significant \((t = -2.39, p < .05)\), with special educators obtaining
significantly higher scores on average than regular educators. An examination of the
implicit attitudinal results showed a significant difference \((t = -2.66, p < .01)\) between \(D\)-
scores in those two groups. Mean scores for special educators were on average
significantly higher on the implicit measure than were those for regular educators, which
implies that special educators show a much stronger preference for the LD label
compared to the ED label. The \(t\)-tests for both explicit attitudes of ED \((t = .22, p = .83)\)
and LD ($t = 1.35, p = .18$) yielded insignificant results. Therefore, mixed results were observed across the measures; attitudinal variation between special and regular educators occurred to a significant degree with general and implicit attitudes, whereas results showed no significance for the two groups with regard to explicit attitudes toward ED and LD.

Table 5

Summary of Intercorrelations, Means, and Standard Deviations for Scores on the SADP, SDS ED, SDS LD, and IAT as a Function of Occupation.

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SADP</td>
<td>–</td>
<td>0.18</td>
<td>0.09</td>
<td>0.12</td>
<td>103.62</td>
<td>14.61</td>
</tr>
<tr>
<td>2. SDS ED</td>
<td>0.35</td>
<td>–</td>
<td>0.15</td>
<td>-0.07</td>
<td>4.88</td>
<td>1.32</td>
</tr>
<tr>
<td>3. SDS LD</td>
<td>0.34</td>
<td>0.13</td>
<td>–</td>
<td>0.19</td>
<td>6.50</td>
<td>0.70</td>
</tr>
<tr>
<td>4. IAT (D-score)</td>
<td>0.14</td>
<td>-0.03</td>
<td>0.03</td>
<td>–</td>
<td>0.47</td>
<td>0.40</td>
</tr>
<tr>
<td>M</td>
<td>110.35</td>
<td>4.83</td>
<td>6.67</td>
<td>0.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>13.16</td>
<td>1.30</td>
<td>0.56</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Intercorrelations for regular education participants ($n = 52$) are displayed above the diagonal; intercorrelations for special education participants ($n = 46$) are displayed below the diagonal. Means and standard deviations for regular education participants are displayed in the vertical columns, and means and standard deviations for special education participants are displayed in the horizontal rows. For these scales, higher values represent stronger agreement between measures in the direction of the assessed construct (i.e., attitude or preference toward disability). SADP = Scale of Attitudes Toward Disabled Persons; SDS ED = Social Distance Scale, Emotional Disturbance item; SDS LD = Social Distance Scale, Learning Disability item; IAT = Implicit Association Test.

A two-way ANOVA was performed to look at the main effects by category (i.e., SDS ED or SDS LD) and occupation (i.e., regular educator or special educator) and to determine whether there was an interaction between the two on explicit attitudes. The results confirmed the significance of the main effects of category ($F = 1.54, p < .001$) and
occupation \( (F = 154.21, \ p < .001) \); however, the interaction was insignificant \( (F = .70, \ p = .41) \).

**Question 3: Are there significant differences between explicit and implicit attitudes of urban compared to rural teachers regarding the ED label?** Much like the analysis used for looking at differences in explicit attitudes among special and regular educators, a repeated measures analysis of variance was performed to see if there were attitudinal differences among rural and urban educators. Using the categories of SDS ED and SDS LD as within subject factors and setting (i.e., rural or urban) as a between-subject factor, results confirmed significance for the main effect of the category \( (F = 1.54, \ p < .001) \), whereas the interaction between setting and the categories of SDS ED and SDS LD was not significant \( (F = 1.67, \ p = .20) \).

Additionally, independent sample t-tests were run to determine whether the sample for urban teachers varied significantly from the sample for rural teachers on the various measures. Results were insignificant on the SDS LD rating \( (t = 1.29, \ p = .20) \); however, significant differences in the two samples were identified on the SDS ED rating \( (t = 2.02, \ p < .05) \). The data revealed that the rural teachers on average possessed greater social aversion to individuals with the ED label as compared to the urban teachers.

For the measure of general attitudes toward persons with disabilities (SADP), results between scores for urban and rural teachers were found to be insignificant \( (t = .99, \ p = .32) \). Likewise, the comparison of implicit attitudes for both urban and rural teachers yielded insignificant results \( (t = -.26, \ p = .80) \).
CHAPTER 5

Discussion

Findings

Some of the findings of this study on teachers’ explicit and implicit attitudes towards the ED label are consistent with anticipated outcomes, whereas others are at variance with, or break from, established hypotheses. The primary consideration of whether educators hold more implicitly negative attitudes toward ED as compared to LD was substantiated persuasively. The finding of a significant preference for LD over ED has been demonstrated in previous explicit attitudinal studies (Levin, Arluke, & Smith, 1982; Thelen, Burns, & Christiansen, 2003); however, such results may have evidenced participants’ knowledge of the ubiquitous stigma of ED as opposed to deeply-held personal beliefs. In other words, one could question whether overtly negative evaluations of the ED label in previous research reflected a bias for confirming a widely-propagated stereotype as opposed to tapping into the actual experiences or perceptions of the raters. The attitudinal data of the current study confirms that indeed there is a strong correspondence between what teachers maintain outwardly about the ED label and their perceptions of the label that are ostensibly below the threshold of conscious thought.

In looking at differences in attitudes between special educators and regular educators, conventional wisdom dictated that teachers with specialized training in working with students with various exceptionalities might possess more moderate views
of ED than those who were not so trained. The data, in fact, portray a different picture inasmuch as the implicit attitude scores show a greater preference for LD over ED among special educators. By contrast, the general attitudes toward disabilities were found to be significantly higher (i.e., more positive) on average for special educators than those for regular educators. These findings are in many ways consistent with those of Schmelkin (1985) who identified greater complexity and wider differentiation in the views of disabilities among special educators compared to other teachers. Thus, one might see that special educators, who service children with a variety of disabilities, are able to draw upon a more positive exemplar when rating individuals with disabilities generally. Blanket evaluative statements about disabilities in general might be more likely to be in opposition to the thoughts and experiences of educators who have observed greater numbers of students from a wider spectrum of disabilities. However, when evaluating the more specific area of ED, special educators may have more negative exemplars to draw upon in making judgments about a population that on balance is characterized by highly negative patterns of behavior.

The findings of this study may call into question whether one of the principal aims of inclusion has been met, namely, to normalize perspectives of non-disabled individuals toward individuals with disabilities by allowing greater levels of exposure to individuals with disabilities. The results tend to support the notion espoused by some, that students with significant emotional or behavioral problems will facilitate negative reactions from those who work with them (Kauffman, 1999). Certainly, there have been changes in the provision of special education services in recent decades, which, being
once very exclusionary, now expose many general education teachers to students who have been identified with ED.

Whereas previous generations of regular education teachers may have been isolated from interactions with students of various disabilities, there is now an expectation of more inclusive environments in the schools that can in part demystify students with the ED label or educate teachers in seeing, and perhaps appreciating, individual differences among labeled students. It would seem likely, however, that the sizable gap in implicit preferences for LD over ED among the special compared to regular educators would be less a reflection of a failure of inclusionary practices and more confirmatory of the practice of continuing to isolate children with the most extreme behaviors in secluded special education settings.

Although the implicit measure established significant differences in magnitude of negative perceptions held by special educators compared to regular educators toward ED, the fact remains that both groups demonstrated much higher negative attitudes both implicitly and explicitly toward ED than toward LD. Such apparent cohesiveness in attitudes might reflect a shared disdain for the types of behaviors commonly associated with the ED label. The IAT was structured to present common behavioral terms (e.g., disruptive behavior, aggression, peer problems, etc.) that historically have been identified as some of the most problematic for classroom management specifically and aggravating to teachers generally (Algozzine, 1977; Boucher & Deno, 1979; Coleman & Gilliam, 1983). The definitional imprecision of the term emotional disturbance renders it possible that respondents to an implicit test for the ED label were uncertain about what the label entailed but were appropriately reticent in making positive associations with behaviors
already viewed as being among the most troubling for them to cope with in an educational setting.

In addition to the finding of particularly negative perceptions of the ED label, mean scores for the generalized measure, the SADP, cast some doubt on whether attitudes of educators toward students with disabilities have progressed in a favorable direction in the intervening years of inclusionary school services. When comparing the mean SADP score for this sample \((M = 106.78)\) with the mean score for the standardization sample \((M = 121.51)\) indicated in the original SADP study by Antonak (1982), it would appear at the very least that attitudes have not grown more favorable toward persons with disabilities. Different characteristics of the two samples could of course explain some of that variance, with Antonak’s sample including 225 undergraduate and graduate students compared to the 98 full-time educators included in the present study. Nonetheless, given that higher values on the SADP are suggestive of more favorable attitudes, it is curious to see how a group of educational professionals would score lower on an attitudinal measure nearly three decades after the norming of the instrument than a group of non-professionals did at its inception. Such an effect appears to be similar to earlier research showing more positive attitudes for teachers-in-training than for established teachers regarding the disabled (Foster et al., 1980) and perhaps reflects on the uncomfortable possibility that inclusionary practices may not serve to moderate negative beliefs toward individuals with disabilities as was originally hoped.

Given the finding that rural teachers report being more averse to interacting with individuals with ED than do urban teachers, there is support for the hypothesis that school setting, or perhaps the behavioral expectancies for students at a given school, may
play a role in influencing perceptions of teachers toward the ED label. Plausible explanations for why rural teachers on average rated greater social distance from individuals with ED included the following: rural teachers have less exposure to individuals identified with ED; rural teachers may be more likely to define abnormal behavior much more broadly than urban teachers do; or fewer quality services and programs exist for rural communities to manage severe problem behaviors.

A major implication of this research is that it reinforces the viewpoint that the ED label may be perceived too negatively by educators to effectively serve as a special education category for children with severe mental health needs. High estimates of school children who have significant mental health problems suggest that there is great need to provide more school-based supports for those children. However, given the categorical nature of special education, it would seem that the options for serving those children would require either a shift to a less stigmatizing classification or a transformation to a non-categorical system of providing interventions of varying intensity (Gresham, 1991). Quibbles about modifying the name, definition, and eligibility criteria for the ED category reflect in part the unfortunate reality that funding for a widening of school-based mental health services may not be a sustainable possibility.

Essentially, there is practical value for the schools in maintaining an antiquated, stigma-laden service category such as ED because of a lack of qualified personnel to provide services to larger numbers of students with emotional difficulties and the need for administrators to keep expenses down while unshackling schools from responsibilities for seemingly intransigent students (Forness & Knitzer, 1992). When considering the evident stigma of the ED label, a chicken-egg sort of argument arises. Do educators have
a negative bias toward the label because it has traditionally been synonymous with the most behaviorally maladaptive children? Or do only the most maladaptive children qualify for ED services because educators hold negative perceptions of the label? Unlike other disability categories in special education, there does not seem to be a continuum of placements afforded to the child who is emotionally and behaviorally disturbed. Rather, the category is reserved for the most extreme cases, thus perpetuating the already negative views educators have.

**Limitations**

Among the limitations in this study, the use of convenience sampling for recruitment of participants yielded what was likely an unrepresentative sample of educators. For example, a disproportionate number of the teachers who responded to the survey were from the elementary level, where traditionally, only a small percentage of the population of students identified as ED is qualified. Thus, it is possible that some portion of the sample lacked meaningful exposure to children known to have the ED label. However, in some cases, the grade level where some teachers reported their present level of instruction was different than where they had taught in the past, due to reconfigurations or changes in the district due to staffing realignments. Thus mobility of educational personnel across grade levels of instruction may have rendered the concern about familiarity with ED less relevant.

Numerous participants commented on taking part in the study perhaps out of sympathy for the researcher’s task of finding volunteers, which may have skewed the sample in the direction of including individuals of a more conscientious and empathetic nature. Nonetheless, if the sample did consist of individuals who were more predisposed
to care about the plight of others (and perhaps students with disabilities as well), one might anticipate that a sample that included a greater mix of less empathic individuals would provide results with even higher levels of negativity toward the ED label and confirm the presence of implicitly negative perceptions among educators to an even greater degree than does the current study.

A further characteristic of this sample that weakens the study’s external validity is the geographically narrow band of school districts from which the sample was taken. When one considers the varied terms that have been used throughout the country to label students with significant emotional and behavioral problems (Kidder-Ashley et al., 2000), it stands to reason that an analysis of implicit preferences of the ED label across areas where different terms have been commonly used might provide insights into the favorability or negativity that accompanies semantic differences in the labels or the differing composition of students receiving ED services in those areas.

In spite of the clear finding of implicit preference for the LD over the ED label, the construction of the implicit attitudinal measure could be considered a limitation of the study. For many studies involving the capture of implicit attitudes via an IAT, the two concepts being compared might be viewed as polar opposites (e.g., young vs. old, smoker vs. non-smoker). Therefore, the resulting attitudinal data would more likely reflect a measurement of preference extending from a positive concept across a wider distance to a negative concept. In the case of the current study, using two disability labels, one would not anticipate participants to view either of the concepts as positive; they are unfavorable to more or less of a degree. This pairing of the LD and ED concepts was chosen in order to answer the questions posed for the study; however, it would remain of interest to
quantify, for example, the magnitude of implicit preference for a concept expressing an absence of disability (e.g., non-disabled) as compared to ED. Such a comparison would be more in line with the traditional usage of an IAT. There would be little justification to verify the presence of an implicit preference for non-disabled over ED as that can be inferred reasonably from the current results, but the size of that preference could be useful as a metric for comparisons with similar IATs for other disabilities.

**Further Research**

For future research, it may be worthwhile to look more closely at the construct of EBD and tease apart which aspects or characteristics are most repellant to teachers. Because students are made eligible for ED services with an amalgam of presenting problems, including both conduct and emotional difficulties (e.g., depression), there may be value in assessing whether more narrow service categories might be useful that provide mental health supports under a less-stigmatizing label. The current study affirms that behaviors commonly associated with the ED label are perceived negatively by educators; however, not all students who qualify as ED exhibit such behaviors, and others who might be eligible could be excluded from emotional support services in order to avoid being stigmatized. Perhaps the time has come to define more clearly an area of services for students with various mood disorders (e.g., anxiety and depression) or other psychological disorders that do not fit cleanly in the ED category, which has become a sort of catch-all for behavioral maladies. The debate about whether to exclude the “socially maladjusted” or conduct-disordered from the ED category in favor of students who present with truly organic emotional disturbances appears to be an unproductive one (Merrell & Walker, 2004), because data clearly suggest that educators hold especially
negative views toward the category regardless of attempts to sanitize it from including delinquent youth.

Concerns about limitations of funding for special education services are justified. A major shift in the definition or eligibility criteria for ED could open floodgates of identification of students as school administrators fear. However, there seems to be a middle ground needed. Data from this research substantiate the commonly-held perspective that the ED label bears negative associations in the minds of educators, but the category remains a toothless, unremarkable option for servicing children when schools are too concerned about its stigmatizing effects and restricting the eligibility only to those children for whom treatment responsiveness is perhaps the most improbable. If there is widespread recognition that the numbers of children in need of significant mental health services is rising, then it stands to reason that the special education category aimed toward redressing psychological dysfunction should see some increase in students identified.

As long as numbers are kept artificially low for the ED category of services, a concerning cycle will persist. Based on low numbers of identification, correspondingly low funding is extended for services to those students. Lower budgets and numbers of students then dictate lesser concern for the category from legislators and administrators, and the need to increase the quality of educational programming and interventions is perhaps overlooked as more attention is given to higher incidence categories of service. One need only look at what has occurred in recent years with autism spectrum disorders to see what happens both from a funding, research, and perhaps quality of services perspective when there is an increase in public awareness about a particular disability and
an increased focus on identification of students who would benefit from services for that condition. In the state of Indiana, there has been a dramatic rise in the identification of students with autism since its official inclusion as a special education category in the early 1990s (Pratt, 2009).

Whereas the concern of stigma has been a limiting factor in the early identification of, and intervention for, students with emotional disorders, there does not appear to have been a similar such stigma to slow the rate of identification of children with autism. Future research might look at the phenomenon of autism and how the concern of stigma has been effectively eliminated as a major deterrent to the identification of those children. Perhaps there can be lessons learned from the rise of eligibility for autism to help curb the use of stigma in denying services for students with ED. Both the categories of ED and autism are characterized by very strong cases for the need for, and effectiveness of, early intervention to help offset some of the severe and pervasive behavioral challenges that become intractable as students get older. Also, the high costs to society for both populations of students, when such preventative measures have not been taken, are well established as well.
References


R.B. Rutherford (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 199-215). New York, NY: Guilford Press.


Appendix A: Research Protocol

All measures were contained within a master file that ran in the Inquisit program on the researcher’s laptop computer. The following information details the sequence of screens that participants may have seen.

1. The first screen is the demographics page (see Figure 1).

![Figure 1 Demographics Page](image)
2. After completing all the fields on the Demographics page and clicking the *Finish* button, an implicit association test will initiate (Figure 2). Participants will receive instructions relative to associating terms that appear in the center of the screen with concepts in either the upper left or upper right corners of the screen via the key strokes E and I.

![Image](image1)

---

**Learning disability**  **Emotional disturbance**

Put your middle or index fingers on the E and I keys of your keyboard. Words representing the categories at the top will appear one-by-one in the middle of the screen. When the item belongs to a category on the left, press the E key; when the item belongs to a category on the right, press the I key. Items belong to only one category. If you make an error, an X will appear - fix the error by hitting the other key.

This is a timed sorting task. **GO AS FAST AS YOU CAN** while making as few mistakes as possible. Going too slow or making too many errors will result in an uninterpretable score. This task will take about 5 minutes to complete.

Press the SPACE BAR to begin.

---

**Figure 2 Implicit Association Test – Instruction Screen**

3. After pressing the SPACE BAR, a series of concept discrimination pages will be presented similar to Figure 3, with terms presented randomly in the center of the screen. Participants strike either the E or I to associate the term with one of the two concepts (disability categories) above.

![Image](image2)

---

**Learning disability**  **Emotional disturbance**  **Self-harm**

---

**Figure 3 Target-Concept Discrimination – Sample Screen**
4. After target words are randomly generated and responded to by the participant, a new screen (see Figure 4) will appear that notifies the participant of a new task that involves associating stimulus words in the center of the screen with attributes in the upper left and right corners.

![Figure 4 Attribute Instructions](image)

See above, the categories have changed. The items for sorting have changed as well. The rules, however, are the same.

When the item belongs to a category on the left, press the E key; when the item belongs to a category on the right, press the I key. Items belong to only one category. An X appears after an error - fix the error by hitting the other key. GO AS FAST AS YOU CAN.

Press the SPACE BAR to begin.

![Figure 5 Associated Attribute Discrimination – Sample Screen](image)

5. After pressing the SPACE BAR, a series of randomly presented associated attribute discrimination screens will appear (see Figure 5).
6. A screen appears (see Figure 6) that introduces a new task, which pairs the presentation of concepts and attributes into a combined task. Once again, participants will see a screen that notifies the participant of a new task that involves associating stimulus words in the center of the screen with attributes in the upper left and right corners. *initial combined task.*

<table>
<thead>
<tr>
<th>Learning disability</th>
<th>Emotional disturbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>or Pleasant</td>
<td>or Unpleasant</td>
</tr>
</tbody>
</table>

See above, the four categories you saw separately now appear together. Remember, each item belongs to only one group. For example, if the categories flower and good appeared on the separate sides above - words meaning flower would go in the flower category, not the good category.

The blue and white labels and items may help to identify the appropriate category. Use the E and I keys to categorize items into four groups left and right, and correct errors by hitting the other key.

Press the SPACE BAR to begin.

![Figure 6 Combined Task Instructions](image)

7. After pressing the SPACE BAR, a series of randomly presented screens will appear similar to Figure 7. Once again, the participant must associate the stimulus word in the center with either a word in the upper left or right of the screen.

<table>
<thead>
<tr>
<th>Learning disability</th>
<th>Emotional disturbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>or Pleasant</td>
<td>or Unpleasant</td>
</tr>
</tbody>
</table>

Love

![Figure 7 Initial Combined Task – Sample Screen](image)
8. A screen appears (see Figure 8) as before with a combined task; however, this time, the participant is encouraged to work quickly. *Note:* This is a testing condition for which the latency of responding was recorded, whereas the previous combined task was a practice condition for the participant.

### Figure 8  Combined Task Instruction Screen - Testing Condition

<table>
<thead>
<tr>
<th>Learning disability</th>
<th>Emotional disturbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>or</td>
<td>or</td>
</tr>
<tr>
<td>Pleasant</td>
<td>Unpleasant</td>
</tr>
</tbody>
</table>

Sort the same four categories again. Remember to go as fast as you can while making as few mistakes as possible.

The blue and white labels and items may help to identify the appropriate category. Use the E and I keys to categorize items into the four groups left and right, and correct errors by hitting the other key.

Press the SPACE BAR to begin.

### Figure 9  Combined Task - Sample Screen

9. After pressing the SPACE BAR, a series of randomly presented screens appear similar to Figure 7. Once again, the participant associated the stimulus word in the center with either the left or right.

### Figure 9  Combined Task - Sample Screen

<table>
<thead>
<tr>
<th>Learning disability</th>
<th>Emotional disturbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>or</td>
<td>or</td>
</tr>
<tr>
<td>Pleasant</td>
<td>Unpleasant</td>
</tr>
</tbody>
</table>

Bad
10. A screen appears (see Figure 10) that explains how the concepts at the top of the screen have been switched. Now the terms on the left and right have been reversed.

<table>
<thead>
<tr>
<th>Emotional disturbance</th>
<th>Learning disability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notice above, there are only two categories and they have switched positions. The concept that was previously on the left is now on the right, and the concept that was on the right is now on the left. Practice this new configuration.

Use the E and I keys to categorize items left and right, and correct errors by hitting the other key.

Press the SPACE BAR to begin.

Figure 10  Reversed Target Concept - Instruction Screen

11. After pressing the SPACE BAR, a series of randomly presented screens will appear similar to Figure 11. Once again, the participant must associate the stimulus word in the center with either the left or right.

<table>
<thead>
<tr>
<th>Emotional disturbance</th>
<th>Learning disability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Brain dysfunction

Figure 11  Reversed Target Concept Discrimination - Sample Screen
12. A screen appears (see Figure 12) that explains how the concepts at the top of the screen are in a new configuration now, paired with the opposite attributes.

<table>
<thead>
<tr>
<th>Emotional disturbance</th>
<th>Learning disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>or</td>
<td>or</td>
</tr>
<tr>
<td>Pleasant</td>
<td>Unpleasant</td>
</tr>
</tbody>
</table>

See above, the four categories now appear together in a new configuration. Remember, each item belongs to only one group.

The blue and white labels and items may help to identify the appropriate category. Use the E and I keys to categorize items into the four groups left and right, and correct errors by hitting the other key.

Press the SPACE BAR to begin.

**Figure 12** Reversed Combined Task - Instruction Screen

13. After pressing the SPACE BAR, a series of randomly presented screens will appear similar to Figure 13. Once again, the participant must associate the stimulus word in the center with either the left or right.

<table>
<thead>
<tr>
<th>Emotional disturbance</th>
<th>Learning disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>or</td>
<td>or</td>
</tr>
<tr>
<td>Pleasant</td>
<td>Unpleasant</td>
</tr>
</tbody>
</table>

Bad

**Figure 13** Reversed Combined Task Practice - Sample Screen
14. A screen appears (see Figure 14) that explains that the participant should now attempt to work fast on the task [test condition] in the same condition as before.

<table>
<thead>
<tr>
<th>Emotional disturbance</th>
<th>Learning disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>or</td>
<td>or</td>
</tr>
<tr>
<td>Pleasant</td>
<td>Unpleasant</td>
</tr>
</tbody>
</table>

Sort the same four categories again. Remember to go as fast as you can while making as few mistakes as possible.

The blue and white labels and items may help to identify the appropriate category. Use the E and I keys to categorize items into the four groups left and right, and correct errors by hitting the other key.

Press the SPACE BAR to begin.

**Figure 14** Reversed Combined Task Test Condition - Instruction Screen

15. After pressing the SPACE BAR, a series of randomly presented screens will appear similar to Figure 15. Once again, the participant must associate the stimulus word in the center with either the left or right.

<table>
<thead>
<tr>
<th>Emotional disturbance</th>
<th>Learning disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>or</td>
<td>or</td>
</tr>
<tr>
<td>Pleasant</td>
<td>Unpleasant</td>
</tr>
</tbody>
</table>

Perceptual problems

**Figure 15** Reversed Combined Task Test Condition - Sample Screen
16. A completion screen for the IAT appears (see Figure 16), instructing the participant that the implicit association test is finished.

You have completed the first portion of the research study.
Press the space bar to proceed with the concluding items.

Figure 16  IAT Completion Screen
17. After pressing the SPACE BAR, the first screen of a measure of general attitudes toward the disabled, the *Scale of Attitudes toward Disabled Persons* (SADP) displays (see Figure 17).

Figure 17  SADP - Screen 1

<table>
<thead>
<tr>
<th>Statement</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children who are disabled should not be provided with a free public education.</td>
<td>-3, -2, -1, 0, +1, +2, +3</td>
</tr>
<tr>
<td>Persons who are disabled are not more accident prone than other people.</td>
<td>-3, -2, -1, 0, +1, +2, +3</td>
</tr>
<tr>
<td>Individuals who are disabled are not capable of making moral decisions.</td>
<td>-3, -2, -1, 0, +1, +2, +3</td>
</tr>
<tr>
<td>Persons who are disabled should be prevented from having children.</td>
<td>-3, -2, -1, 0, +1, +2, +3</td>
</tr>
</tbody>
</table>
18. After clicking the **Next** button, the second screen of the SADP displays (see Figure 18).

![Figure 18 SADP - Screen 2](image-url)
19. After clicking the **Next** button, the third screen of the SADP displays (see Figure 19).

![Screen 3 of SADP](image)

**Figure 19  SADP - Screen 3**
20. After clicking the **Next** button, the fourth screen of the SADP displays (see Figure 20).

Figure 20  SADP - Screen 4
21. After clicking the **Next** button, the final screen of the SADP displays (see Figure 21).

![Figure 21 SADP - Screen 5](image)
22. After clicking the **Finish** button, a Social Distance Scale screen appears *(see Figure 22).* Participants then choose their responses to each of the eight items from drop down menus *(see Figure 23).*

![Figure 22 Social Distance Scale](image1)

![Figure 23 Sample Drop-Down Menu](image2)
23. After clicking the **Next** button, a final screen displays *(see Figure 24).*

![Final Screen](image)

*Figure 24  Final Screen*

24. If the participant clicks the **Finish** button *(see Figure 24)*, an end screen appears with the word **Done** in the middle *(see Figure 25).* This screen cannot be minimized or closed unless overridden with a combination of keystrokes known only to the researcher; thus, the participant is prohibited from accessing any data or programs on the computer.

![Screen Lock](image)

*Figure 25  Screen Lock*
Appendix B: Standard Recruitment Email/Letter

Dear Educator,

My name is Patrick Jones, a school psychologist working in the X schools. I am conducting a research study through my graduate program at Ball State University that involves gathering teachers’ attitudes toward students’ disabilities. The purpose of this research is to look at perspectives of educators toward certain special education categories. The results may help inform the need to modify criteria or terminology that is currently in our legislative language for special education classifications.

Who is needed: I am hoping to find 50 participants from this school district. Participants must be full-time general or special education teachers.

Incentive: I am offering participants a chance to receive $100 for taking part in this study. The time is estimated at 15-20 minutes to complete the survey/measures, and I will come to your school at a time convenient for you to participate. Two names from the 50 participants will be drawn at random at the conclusion of the study and given the monetary gift. Your odds of receiving the gift are approximately 1:25.

For more information or to enroll as a participant in this research study, send an email to pat_jones@xxxxxx.k12.in.us or call 765-xxx-xxxx.
Appendix C: Consent Form

**Study Title**  Teachers’ Perceptions of Special Education Categories

**Study Purpose and Rationale**
The purpose of this research project is to examine how educators feel about various special education classifications that are commonly used for students. Findings from this research may help schools and advocacy groups better understand some of the factors that influence the qualification of students for certain services and the lower than expected rates of identification for certain areas.

**Inclusion/Exclusion Criteria**
To be eligible to participate in this study, you must be a full-time special education or general education teacher at your school. You should have attained a Bachelor’s Degree as a minimum in academic training.

**Participation Procedures and Duration**
This study involves the completion of a series of computer-based questionnaires and measures, including a brief demographic survey, a word association activity, a general measure of attitudes toward disabilities, and an assessment of social desirability of disability categories. The entire process should take approximately 15-20 minutes for you to complete.

**Data Confidentiality or Anonymity**
All data will be maintained anonymously and no identifying information such as names will appear in any publication or presentation of the data.

**Data Storage**
Because all data will be collected electronically, there will be no paper record of your responses to any questionnaire items. The data will be stored on the researcher’s password-protected laptop computer and backed up to a secure network directory on a university website. This information will be maintained for no more than five years and then destroyed. No one except for the researcher will have access to the data.

**Risks or Discomforts**
The only foreseeable risk from participating in this study is that you may feel uncomfortable responding to some of the questions. You do not have to answer any question that distresses you.

**Benefits**
One benefit from the study may be the sense of contributing to the literature in the area of special education and classification of students.

**Voluntary Participation**
Participation in this study is voluntary and you may withdraw at anytime for any reason without penalty or prejudice from the researcher. Please feel free to ask any questions before signing this form and at any time during the study.
IRB Contact Information
For your rights as a research participant, you may contact the following: Research Compliance, Sponsored Programs Office, Ball State University, Muncie, IN 47306, (765) 285-5070, irb@bsu.edu.

Study Title  Teachers’ Perceptions of Special Education Categories

**********

Consent

I, __________________, agree to participate in this research project entitled, “Teachers’ Perceptions of Special Education Categories.” I have had the study explained to me and my questions have been answered to my satisfaction. I have read the description of this project and give my consent to participate. I understand that I will receive a copy of this informed consent form to keep for future reference.

To the best of my knowledge, I meet the inclusion/exclusion criteria for participation (described on the previous page) in this study.

_____________________________  ____________________
Participant’s Signature                  Date

Researcher Contact Information

Principal Researcher:  Faculty Supervisor:

J. Patrick Jones, Graduate Student  Dr. Barbara Rothlisberg
School Psychology  Professor of Educational Psychology
Ball State University  Ball State University
Muncie, IN 47306  Muncie, IN 47306
Telephone: (765) 639-1707  Telephone: (765) 285-8520
Email: jones@bsugmail.net  Email: brothlisberg@bsu.edu