The Devereux Student Strengths Assessment—*Second Step* Edition (DESSA-SSE)

A measure of social–emotional competencies of children in kindergarten through the 5th grade developed to accompany the *Second Step* Program.

Paul A. LeBuffe, Jack A. Naglieri, & Valerie B. Shapiro

Technical Manual, Support Portal Version
(Items & Norms Redacted)

Acknowledgments

The DESSA-SSE has been developed as a partnership between the Devereux Center for Resilient Children and the Committee for Children. The authors, on behalf of Devereux, would like to thank our partners at the Committee for Children. We are very appreciative of the many contributions of Sherry Catron Burke, Director of Programs, Partnerships, and Research, and Brian Smith, Ph.D., Research Scientist, who collaborated with us in the creation of the DESSA-SSE scales. Ms. Burke and Dr. Smith also played key roles in helping to design the DESSA-SSE materials and in reviewing this manual.

We would also like to acknowledge the leadership roles played by Richard Israel, Chief Operating Officer, and Sally Vilardi, Director of Marketing and Client Relations of the Committee for Children. As members of the project leadership team, they ensured that the DESSA-SSE met the high quality standards that characterize all of the products and services of the Committee for Children. Joan Duffel, the Executive Director of the Committee for Children, deserves special acknowledgement for her inspirational leadership and for being such an exemplary partner to the Devereux Center for Resilient Children.

At Devereux, we would like to thank Robert Kreider, President and Chief Executive Officer, and Martha Lindsay, Vice President for Product Development, for their unwavering commitment to the mission and staff of Devereux Center for Resilient Children. Their guidance and support have promoted the resilience of millions of children over the past decade. The authors also thank the Co-Director of the Devereux Center for Resilient Children, Linda Likins. It was Ms. Likins's vision and leadership that led to the partnership of Devereux and the Committee for Children, which resulted in the DESSA-SSE. We would also like to thank Jennifer Fleming, Research Associate at the Devereux Center for Resilient Children, for her assistance in preparing this manual.

We are also indebted to our publishing partners at the Kaplan Early Learning Company, especially Matthew Marceron and Clarissa Willis. We greatly appreciate their embrace of the partnership with the Committee for Children.

Many other individuals contributed to the development of the Devereux Student Strengths Assessment (DESSA), which is the basis for the DESSA-SSE. These individuals are acknowledged in the DESSA manual (LeBuffe, Shapiro, & Naglieri, 2009). Their efforts have made the DESSA-SSE possible.

About the Authors

Paul A. LeBuffe, M.A. is the Co-Director of the Devereux Center for Resilient Children in Villanova, Pennsylvania. He is a graduate of St. Mary's College of Maryland and received his Master's degree in experimental psychology from Bryn Mawr College. Paul is an author of many widely used assessments of behaviors related to children's social–emotional strengths and needs including the *Devereux Behavior Rating Scale—School Form,* the *Devereux Scales of Mental Disorders (DSMD),* the *Devereux Early Childhood Assessment (DECA),* the *Devereux Early Childhood Assessment Clinical Form (DECA-C),* the *Devereux Early Childhood Assessment—Infant/Toddler Form (DECA-I/T),* and the *Devereux Student Strengths Assessment (DESSA).* His major research interests are the measurement and promotion of social–emotional competence and resilience in children, psychometrics and test development, and suicide prevention education.

Jack A. Naglieri, Ph.D. is a senior research scientist at the Devereux Center for Resilient Children, research professor at the University of Virginia, and emeritus professor of psychology at George Mason University. Jack is a Fellow of APA Division 16, the recipient of the APA Division 16 2001 Senior Scientist Award, and the Italian American Psychological Assembly 2011 award for distinguished contributions to psychology. Dr. Naglieri obtained his B.A. in psychology from Long Island University, an M.S. from St. John's University, and his Ph.D. in educational psychology from the University of Georgia in 1979. The author of more than 250 scholarly papers, chapters, books, and tests, he has focused his efforts since the late 1970s on reconceptualizing intelligence. In addition to the several different Devereux scales, Dr. Naglieri is the author of the *Cognitive Assessment System* (Naglieri & Das, 1997), the *Wechsler Nonverbal Scale of Ability* (Wechsler & Naglieri, 2006), the *General Ability Measure for Adults* (Naglieri & Bardos, 1997), the *Naglieri Nonverbal Ability Test—Second Edition* (Naglieri, 2008), and the *Autism Spectrum Rating Scales* (Goldstein & Naglieri, 2008).

Valerie B. Shapiro, Ph.D. is an assistant professor in the School of Social Welfare at the University of California, Berkeley and a research analyst for the Devereux Center for Resilient Children. She earned her B.A. in psychology from Colgate University, her M.S.S. from the Bryn Mawr Graduate School of Social Work & Social Research, and her Ph.D. from the University of Washington School of Social Work (2012). A licensed social worker and Department of Education-certified school social worker, Dr. Shapiro has worked in diverse settings providing school-based prevention programming, community-based outpatient psychological services, and residential clinical interventions. Her current research examines the relationship between risk and protective factors, preventive interventions, and youth outcomes; the design and evaluation of tools to facilitate the use of protective factors in strength-based and evidence-informed practice; and the development of community-based infrastructures to promote and sustain the use of such tools in practice environments.

Table of Contents

Tables

Figures

Foreword

In 1996, after nearly 75 years of providing behavioral health, rehabilitation and special education services to individuals with special needs, Devereux embarked on a new mission: primary prevention. Responding to the burgeoning numbers of children and adolescents with social and emotional problems, the president and CEO of Devereux began an initiative that has evolved to become the Devereux Center for Resilient Children.

The Devereux Center for Resilient Children (DCRC) began by focusing on enhancing the social–emotional competence and resilience of preschool children through the Devereux Early Childhood Initiative (DECI). The mission of the DECI is to promote young children's social and emotional development, foster resilience, and build the skills for school and life success. The DECI staff developed the first nationally standardized assessment of behaviors related to resilience in preschool children, the Devereux Early Childhood Assessment or DECA (LeBuffe & Naglieri, 1999). The DECA and other resources developed by the DECI staff are now used in thousands of preschool and infant-toddler programs throughout the United States and Canada. More information on the DECI can be found at its website, www.devereuxearlychildhood.org.

In 2009, with the completion of the Devereux Student Strengths Assessment (LeBuffe, Shapiro, & Naglieri, 2009; DESSA), the child-centered, social–emotional competence-enhancing approach originating with the DECA was extended upward to children in Grades K–8. This expansion of the age range served was accompanied by the creation of the Devereux Center for Resilient Children (DCRC), which encompasses both the DECI and the DESSA. More information on the DCRC can be found at its website, www. centerforresilientchildren.org.

The Devereux Student Strengths Assessment—*Second Step* Edition (LeBuffe, Naglieri, & Shapiro, 2011; DESSA-SSE) is a result of a partnership between the DCRC and the Committee for Children. Both organizations have a shared mission to promote the social and emotional competence of all children. Through this joint venture, we are combining the assessment development expertise of the DCRC with the program development expertise of the Committee for Children to better serve the professional community and children.

The development of the DESSA-SSE was guided by the following underlying principles:

- ***Strength-Based***—All of the assessments, programmatic resources, and services provided by the DCRC are focused on building the social–emotional strengths of children. Although we appreciate the importance of addressing the behavioral concerns and problematic behaviors of children already experiencing significant social–emotional problems, we also stress the importance of promoting children's social and emotional competencies. We have described the many advantages of a strength-based approach elsewhere (LeBuffe & Shapiro, 2004), but perhaps the most important benefit is that these social–emotional competencies contribute to a child's resilience in the face of adversity.
- ***Excellence in Assessment***—Instruments used to guide practices that influence children's lives need to be constructed using the highest professional standards. This includes using well-established test development methods to ensure the highest psychometric qualities, especially a nationally representative sample for creation of norms and specific guidelines for use and interpretation of the scores that the rating scale provides. The DESSA-SSE has been developed to meet or exceed the standards for assessment instruments established by the American Educational Research Association and the American Psychological Association

(APA, 1999). Information on the development and psychometric properties of the DESSA-SSE is presented in Chapters 2 and 3.

- ***Parent–Professional Collaboration***—The probability of a successful outcome for a child is enhanced when parents and professionals such as teachers, after-school staff, and mental health and child welfare professionals work in concert to support the child, ensuring a consistent approach across environments. The DESSA-SSE has been designed to support this collaboration.
- ***Support Effective Practice***—We believe that assessments must support parents, teachers, and other professionals in delivering effective services that will lead to improved outcomes for the child. The DESSA-SSE provides two key elements in this process. First, it provides valuable information about the child's social–emotional strengths and needs. Second, the DESSA-SSE can be used to evaluate the effectiveness of interventions and thereby support data-driven practice, provide practice-based evidence of program efficacy, and promote professional accountability.
- ***Promote Communities of Practice***—Since the publication of the DECA in 1999, there has been increasing interest in empirically supported approaches to enhancing social–emotional competence and resilience in at-risk children. The authors of the DESSA-SSE hope that the publication of this assessment will support and extend the current efforts of communities to recognize the importance of and promote these competencies. The authors of the DESSA-SSE as well as the staff of the Devereux Center for Resilient Children welcome the opportunities to collaborate with colleagues, students, and organizations who share these goals. We can be reached through the DCRC website: www.centerforresilientchildren.org.

CHAPTER 1: Introduction

The Devereux Student Strengths Assessment–*Second Step* Edition (DESSA-SSE) is a standardized, norm-referenced behavior rating scale that assesses social–emotional competencies of children in kindergarten through 5th grade. Developed to assist school, after-school, social service, and mental health professionals who utilize the *Second Step* program (Committee for Children, 2011), the DESSA-SSE provides a systematic, empirically based approach to assessing social–emotional competencies and evaluating the impact of the *Second Step* program.

The DESSA-SSE measures social–emotional competence in children and youth in kindergarten through 5th grade. *Social–emotional competence* means that children have the skills to successfully manage their emotions and behavior, cooperate with others, form positive relationships, and make responsible decisions. We view the DESSA-SSE as measuring "How well the child has met, and continues to meet, the expectations explicitly or implicitly set in the society for children as they grow up" (Wright & Masten, 2005, p. 21). The goal should be to assist the child in developing strengths in these areas and not to merely support the attainment of minimally adequate social–emotional skills.

The social–emotional competencies measured by the DESSA-SSE are critically important personal attributes that benefit all children. In a recent meta-analysis of 213 studies involving more than 270,000 students, Durlak and colleagues (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011) documented that high-quality social and emotional learning programs like the *Second Step* program, when implemented with fidelity, result in significant improvements in five areas:

1. decreased disruptive behavior
2. decreased anxiety
3. increased positive attitudes toward school, peers, and self
4. increased social and emotional competence
5. improved academic performance.

In short, the skills taught by the *Second Step* program and measured by the DESSA-SSE are essential for children's success in school and life.

Description of the DESSA-SSE

The DESSA-SSE is a 36-item standardized, norm-referenced behavior rating scale that assesses the social–emotional competencies taught by the *Second Step* program. The DESSA-SSE can be completed by parents, teachers, or staff at schools and child-serving agencies, including after-school, social service, and mental health programs. The assessment is entirely strength based, meaning that the items query positive behaviors (e.g., get along with others) rather than maladaptive ones (e.g., annoy others).

The DESSA-SSE is organized into conceptually derived scales that provide information about four critically important social–emotional competencies. Standard scores can be used to calibrate each child's competency in each of the four dimensions. For each question, the rater is asked to indicate on a five-point scale how often the student engaged in each behavior over the past 4 weeks. The scale names, scale definitions, and sample scale items are as follows:

- ***Skills for Learning*** (nine items): A child's ability to use the skills of listening, focusing attention, self-talk, and assertiveness.
- ***Empathy*** (nine items): A child's ability to identify and label emotions in himself/herself and others and take on others' perspectives.

- ***Emotion Management*** (nine items): A child's ability to cope with strong emotions and express them in socially acceptable ways.
- ***Problem Solving*** (nine items): A child's ability to effectively handle personal challenges and interpersonal conflicts in prosocial ways.

Each of the four DESSA-SSE scale scores is derived from the ratings of the items assigned to that scale. A Social–Emotional Composite score is also included, which is a combination of the four scales. This composite score provides an overall indication of the strength of the child's social–emotional competence. The separate scores on the four DESSA-SSE scales are used to create Individual Student Profiles as well as Classroom/Program Profiles to convey the strengths and needs of the student or groups of students as compared to national norms. This information can also be used to compare ratings across time to monitor progress and evaluate outcomes. More information about these interpretation strategies will be presented in Chapter 5.

Uses of the DESSA-SSE

The DESSA-SSE has been developed to provide a measure of social–emotional competencies taught by the *Second Step* program. Specifically, the DESSA-SSE has been designed to:

- Provide a psychometrically sound, strength-based measure of social–emotional competence of children in kindergarten through 5th grade.
- Describe the social–emotional competence of children. This information can be used at the school or district level as part of a needs assessment for implementing a social–emotional learning program such as the *Second Step* program or for tracking levels of social–emotional competence at the district, school, or class level over time.
- Evaluate the impact of the *Second Step* program and similar social–emotional learning programs by rigorously evaluating outcomes at the child, classroom, grade, school, and district levels.

Qualifications of DESSA-SSE Users and Raters

Qualifications of DESSA-SSE Users

For the purposes of this manual, DESSA-SSE *users* are those who not only administer the DESSA-SSE but also interpret its scores. The guidelines presented here should be considered a general description rather than an exhaustive list of those who may use the DESSA-SSE. In presenting these descriptions, we assume that the titles used by professionals in different settings vary, as do their levels of training and the regulations that govern professional practice in their states. In every case, however, the DESSA-SSE user has responsibility for the proper use and interpretation of DESSA-SSE results. Therefore, DESSA-SSE users should have training in the proper administration, interpretation, and utilization of the DESSA-SSE. This should include knowledge of the interpretation of standardized scores such as *T*-scores and percentiles, the interpretation of scale content and profiles, and communication of the results to the child, parents, and professionals. Typically, DESSA-SSE users will include administrators, counselors, educators, mental health consultants, nurses, program directors and evaluators, pediatricians, psychologists, researchers, school psychologists, and social workers.

Qualifications of DESSA-SSE Raters

A *rater* is any person who completes the items on the DESSA-SSE. There are two main qualifications of a rater: First, the rater must have had sufficient exposure to the child over the 4 weeks prior to completing the DESSA-SSE; and second, raters should also be able to read English at the 6th-grade level. (**Note:** Recommendations for using

the DESSA-SSE with raters who have difficulty reading English are presented in Chapter 4.) Because the scores are a function of the number of times specific behaviors have been noted, a rater's insufficient opportunity to observe the child could yield an erroneously low rating. In general, raters should have contact with the child for 2 or more hours a day for at least 3 days per week for a 4-week period.

Raters generally fall into two categories: (a) parents, guardians, and other adult caregivers who live with the child and (b) teachers, after-school program staff, or other professionals who interact directly with the child on a regular basis. The first group of raters is referred to as "parents" and the second group as "teachers" in the remainder of this manual and on the DESSA-SSE record form.

Restrictions for Use

DESSA-SSE users should follow both the instructions included in this manual and commonly accepted guidelines for test use and interpretation, such as the *American Psychological Association's Standards for Educational and Psychological Testing* (APA, 1999). It is the DESSA-SSE user's responsibility to ensure that completed DESSA-SSE protocols and reports remain secure and are released with parental consent only to professionals who will safeguard their proper use. Copyright law does not permit the DESSA-SSE user to photocopy or otherwise duplicate test items or record forms in any way, even for the purpose of sharing results with parents or multidisciplinary teams. The completed DESSA-SSE Individual Student Profile may be copied and provided to parents after it has been reviewed with them. Because all DESSA-SSE items, norms, and other materials are copyrighted, no DESSA-SSE materials may be reproduced or transmitted in any form or by any means without written permission from the Devereux Foundation.

CHAPTER 2: Development and Standardization

Development of the DESSA-SSE Items

The Devereux Student Strengths Assessment—*Second Step* Edition (DESSA-SSE) was derived from the DESSA (LeBuffe, Shapiro, & Naglieri, 2009). All 36 items on the DESSA-SSE were selected from the 72 items on the DESSA. Therefore, the process used for the development of the original DESSA items is presented here.

A variety of approaches were used to develop the initial set of DESSA items. First, we carefully reviewed the literature on resilience (e.g., Werner & Smith, 1982, 1992), social–emotional learning (e.g., Payton et al., 2000), and positive youth development (e.g., Catalano, Berglund, Ryan, Lonczak, & Hawkins, 1998) to identify behavioral descriptions of resilient children and youth. Second, we examined other strength-based assessments, such as the Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999a, 1999b). This resulted in an initial set of 765 potential items. Next, we reduced the pool of potential items by combining statements with similar meanings, deleting those that were not measurable (e.g., references to nonobservable subjective states or cognitions), and eliminating any that were overly value laden or relatively unmodifiable, such as physical attractiveness.

The item refinement phase resulted in a pool of 156 items, which served as the starting point in the construction of the DESSA. All the items were written to measure observable behaviors that would require little or no inference on the part of the observer. We carefully considered the reading level of the items so that the overall readability level of the DESSA would be as low as possible.

To investigate the usefulness of the initial items and their interrelationships, we conducted a national pilot study. Either parents or teachers completed ratings on 428 students in kindergarten through 8th grade. Of these students, 106 (25%) had already been identified as having significant emotional or behavioral disorders. We reduced the initial pool of 156 items by eliminating any items that had relatively low item–total correlations (i.e., < .60), did not differentiate between those students with known emotional or behavioral disorders and those without by at least half a standard deviation, or were rated by 20% or more of the parents and teachers as unclear or not applicable. This resulted in a set of 81 items that we incorporated into the standardization edition of the DESSA.

National Standardization

The standardization sample of the DESSA also serves as the standardization sample of the DESSA-SSE. Therefore, the process used in standardizing the DESSA is presented here.

We standardized the DESSA through a carefully prescribed method so that the sample would closely represent the United States population on several important dimensions. The data collection procedures also ensured that a wide variety of children and youth were included for the generation of norms. We collected data using both a paper form and an online, computerized version. Both samples were collected simultaneously from April 2005 through March 2006.

We collected ratings from two groups of individuals: (a) parents and other relatives living with the child and (b) teachers and after-school program staff. We obtained teacher ratings from schools across the United States. We obtained parent ratings from these same schools; through recruitment posters at community locations, such as clinics and libraries; and through parenting listserves and Internet forums. Staff from 81 after-school programs participated in the development of the DESSA and hence the DESSA-SSE. To ensure confidentiality, the completed ratings were sent directly to the Devereux Foundation Institute of Clinical Training and Research (ICTR). No

personally identifying information was included in the standardization protocols. A complete list of standardization sites can be found in the DESSA manual.

Representativeness of the DESSA/DESSA-SSE Standardization Sample

A total of 2,494 children and youth who were in kindergarten through 8th grade at the time of the data collection comprised the DESSA/DESSA-SSE standardization sample. Teachers and teacher aides provided ratings on 778 students; parents and other adult relatives living in the home provided ratings on 1,244 children and youth; and after-school and other program staff contributed the remaining 472 ratings.

A comparative analysis of responses obtained through paper and computerized versions utilizing hierarchical regression revealed that, when controlling for demographic variables, the administration format offered no additional predictive validity in explaining variation in DESSA scores (Berkley, 2008). Using the final norms tables, the mean DESSA Social–Emotional Composite *T*-scores were 50.6 and 49.3 (*d*-ratio of 0.1) for the paper and pencil and online samples, respectively. The *d*-ratio indicates that the means of these two samples differed by approximately one-tenth of a standard deviation, which would be interpreted as a negligible effect size. Therefore, in all subsequent analyses, we combined data obtained from both administration formats. Similarly, there were minimal differences between the ratings provided by teachers and after-school staff, so these data sets were combined as well. In all subsequent analyses and descriptions, "parent" refers to a parent or other adult relative living with the child; "teacher" refers to a teacher, teacher aide, or member of an after-school staff.

The DESSA/DESSA-SSE standardization sample closely approximated the K–8 population of the United States with respect to age, gender, geographic region of residence, race, ethnicity, and socioeconomic status. We based the desired characteristics of the standardization sample on the *Statistical Abstract of the United States 2008: The National Data Book* published by the U.S. Census Bureau. In the tables that follow, the total numbers of children included may not sum to 2,494, due to missing data.

Grade and Gender—Table 2.1 presents the numbers and percentages of males and females in each grade from kindergarten through 8th grade. The number of children in each grade ranged from 104 in 8th grade to 492 in kindergarten. The overall mean number of students per grade was 275. These results show that each grade was well sampled. The data also show that the percentages of males and females in the standardization sample as a whole, as well as in each grade, closely approximated the proportions of the U.S. population.

Table 2.1

DESSA-SSE Standardization Sample Characteristics by Grade and Gender

	Males		Females		Total	
	n	%	*n*	%	*n*	%
Kindergarten	256	52.0	236	48.0	492	19.8
1st Grade	186	50.0	186	50.0	372	15.1
2nd Grade	161	50.0	161	50.0	322	13.1
3rd Grade	160	50.0	160	50.0	320	12.9
4th Grade	134	47.5	148	52.5	282	11.4
5th Grade	138	49.1	143	50.9	281	11.3
6th Grade	88	48.9	92	51.1	180	7.2
7th Grade	57	46.7	65	53.3	122	4.9
8th Grade	46	44.2	58	55.8	104	4.2
Total Sample	1,226	49.5	1,249	50.5	2,475	
U.S. %		51.2		48.8		

Note: The U.S. population data are based on the 2006 figures for 5- through 14-year-olds only in "Resident Population by Age and Sex: 1980 to 2006, Table No. 7," *Statistical Abstract of the Unites States: 2008 (127th edition)*: U.S. Census Bureau, 2008.

Geographic Region—We collected data from parents and teachers of students attending 711 schools in all 50 states and the District of Columbia. Table 2.2 shows the numbers and percentages of students by grade level and location, according to the four geographic regions designated by the U.S. Census Bureau: Northeast, South, Midwest, and West. These data show that the DESSA-SSE standardization sample closely approximated the regional distribution of the U.S. population.

Table 2.2

DESSA-SSE Standardization Sample Characteristics by Grade and Geographic Region

	Northeast		South		Midwest		West		Total
	n	%	*n*	%	*n*	%	*n*	%	*n*
Kindergarten	128	26.0	148	30.1	125	25.4	91	18.5	492
1st Grade	73	19.5	143	38.2	76	20.3	82	21.9	374
2nd Grade	63	19.3	132	40.5	63	19.3	68	20.9	326
3rd Grade	62	19.4	155	48.6	69	21.6	33	10.3	319
4th Grade	79	27.8	99	34.9	48	16.9	58	20.4	284
5th Grade	62	22.1	97	34.5	57	20.3	65	23.1	281
6th Grade	28	15.6	44	24.6	45	25.1	62	34.6	179
7th Grade	27	22.1	43	35.2	29	23.8	23	18.9	122
8th Grade	19	18.3	30	28.8	30	28.8	25	24.0	104
Total Sample	541	21.8	891	35.9	542	21.8	507	20.4	2,481
U.S. %		17.4		36.3		22.2		24.0	

Note: The U.S. population data are based on the 2006 figures for 5- through 17-year-olds only in "Resident Population by Age and State: 2006, Table No. 16," *Statistical Abstract of the United States: 2008 (127th edition)*: U.S. Census Bureau, 2008.

Race—Table 2.3 provides the DESSA-SSE standardization sample composition by race and geographic region. Based on information provided on the rating forms, we classified the children according to the five major race categories used by the U.S. Census Bureau: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. The data in Table 2.3 indicate that the racial composition of the total standardization sample closely approximated that of the U.S. population.

Table 2.3

DESSA-SSE Standardization Sample Characteristics by Race and Geographic Region

	American Indian/Alaska Native		Asian		Black/African American		Native Hawaiian/ Pacific Islander		White		Total
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	
Northeast	9	1.7	7	1.3	190	35.6	0	0	327	61.4	533
South	16	1.8	22	2.5	200	22.3	3	0.3	424	47.3	665
Midwest	3	0.6	12	2.2	71	13.0	2	0.4	453	83.1	541
West	13	2.6	24	4.7	20	3.9	9	1.8	354	69.8	420
Total	41	1.9	65	3.0	481	22.3	14	0.6	1,558	72.2	2,159
U.S. %[1]		1.2		4.0		15.4		0.2		76.3	

Note: The U.S. race data are based on the 2006 figures for 5- through 14-year-olds only in "Resident Population by Race, Hispanic Origin, and Age: 2000 and 2006, Table No. 8," *Statistical Abstract of the Unites States: 2008 (127th edition)*: U.S. Census Bureau, 2008.
[1] U.S. figures do not add up to 100% due to "Two or more Races" not being included.

Hispanic Ethnicity—The proportions of children of Hispanic ethnicity included in the DESSA-SSE standardization sample are presented in Table 2.4. These data, based on the number of participants who reported Hispanic ethnicity, show that the composition of the standardization sample closely approximated that of the U.S. population.

Table 2.4

DESSA-SSE Standardization Sample Characteristics by Hispanic Ethnicity and Geographic Region

	Hispanic		Non-Hispanic		Total
	n	%	*n*	%	*n*
Northeast	35	6.4	510	93.6	545
South	259	28.9	638	71.1	897
Midwest	23	4.2	522	95.8	545
West	133	26.2	374	73.8	507
Total	450	18.0	2,044	82.0	2,494
U.S. %		19.9		80.1	

Note: The total U.S. Hispanic population data are based on the 2006 figures for 5- through 14-year-olds only in "Resident Population by Race, Hispanic Origin, and Single Years of Age: 2006, Table No. 9," *Statistical Abstract of the Unites States 2008 (127th edition)*: U.S. Census Bureau, 2008.

Socioeconomic Status—To assess the socioeconomic status of the DESSA-SSE standardization sample, we determined the number of students eligible to receive either free or reduced-price lunches. Of the entire sample of 2,494 students, 550 (22%) were eligible to receive free or reduced-price lunches. This very closely approximated the 25% of children

and youth living in poverty. This figure is slightly higher than the 19% of families in 2005 whose income was $25,000 or less (U.S. Census Bureau, 2008, Table 685) and would qualify for the free school lunch program.

Organization of DESSA-SSE Items into Scales

The primary purpose of the DESSA-SSE is to provide teachers, parents, after-school staff, and other professionals concerned with the mental health of children and youth with a useful and meaningful set of scales that both reflect current social–emotional functioning and can be used to evaluate the impact of social–emotional learning programs such as the *Second Step* program.

Utilizing the standardization data set, we organized DESSA-SSE items into logically derived and statistically validated scales based on the social–emotional competency framework provided by the *Second Step* program. First, we reviewed the definitions and descriptions of the four *Second Step* program skill areas provided in the *Second Step* program (Committee for Children, 2011). Next, items were assigned to one of the four scales based on their relationship to the definitions and descriptions of the four skill areas: *Skills for Learning, Empathy, Emotion Management,* and *Problem Solving*. These initial item–scale assignments were reviewed by the Research and Program Development staff of the Committee for Children. Based on their feedback, a few items from the initial mapping were deleted and others substituted in their place.

We then used a series of statistical analyses to further refine and simplify the scales based on the following goals: (a) to identify the best scale solution from both psychometric and interpretability perspectives; (b) to shorten the DESSA-SSE as much as possible without compromising breadth of coverage; (c) to simplify the administration, scoring, and interpretation of the DESSA-SSE; and (d) to ensure that the constructs were measured reliably by the scales. We examined the corrected item–total correlations to ensure that each item correlated highly with the scale to which it was assigned. Next, we examined each item for age trends and found very few that showed a progression of scores across age. Those few items that did show age trends were eliminated, simplifying the scales and avoiding the necessity of age norms. The elimination of any item that showed an age trend also enables the use of the entire standardization sample, which includes children in Grades 6 through 8 even though the *Second Step*: *Skills for Social and Academic Success* program encompasses kindergarten through 5th grade. This process resulted in a final set of 36 items, with nine items assigned to each of the four scales. Finally, based upon the sum of the standard scores of all four scales, we also created a Social–Emotional Composite, which provides an overall estimate of a child's social and emotional competencies.

Norming Procedures—The first step in preparation of the norms was to determine if any trends existed in the data. We examined the DESSA-SSE scale raw scores for age, rater, and gender differences. Table 2.5 presents the raw-score means for the four DESSA-SSE scales in one-grade intervals. These data are also presented in Figure 2.1 (Parent Raters) and Figure 2.2 (Teacher Raters). It is apparent that there is only minor variability across grades in these means, indicating an absence of age trends across the K–8 range; therefore, we constructed the norms for all grades combined. The lack of age trends also enabled us to use the full standardization sample in constructing the scales.

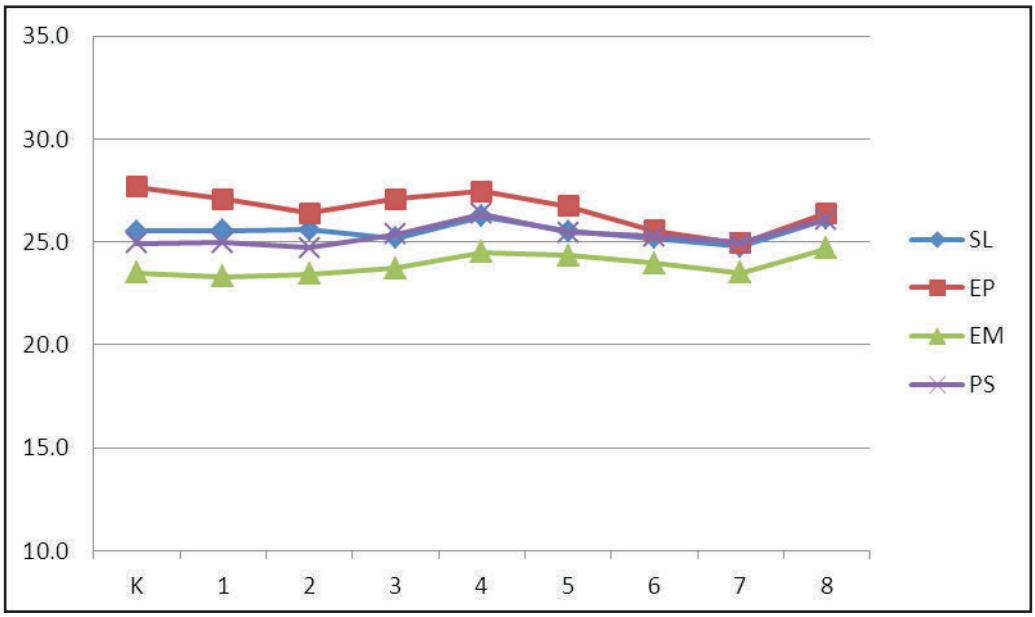
Table 2.5

DESSA-SSE Mean Raw Scores and Standard Deviations by Grade

	Parents				Teachers			
	SL	EP	EM	PS	SL	EP	EM	PS
K	25.5 (5.1)	27.7 (5.3)	23.5 (4.8)	24.9 (4.7)	24.5 (7.1)	24.4 (6.7)	23.8 (6.3)	23.4 (6.5)
1	25.5 (5.0)	27.1 (5.1)	23.3 (5.0)	24.9 (5.0)	23.4 (6.4)	23.2 (5.9)	22.3 (5.8)	22.4 (5.9)
2	25.6 (5.1)	26.4 (5.2)	23.4 (4.8)	24.7 (4.6)	23.8 (6.7)	23.5 (6.2)	22.8 (5.6)	23.0 (6.0)
3	25.2 (5.0)	27.1 (5.2)	23.7 (4.8)	25.4 (5.0)	23.2 (7.1)	23.4 (6.8)	22.3 (6.6)	22.1 (7.2)
4	26.2 (5.4)	27.4 (5.0)	24.5 (5.0)	26.4 (4.8)	26.1 (6.5)	25.1 (6.3)	24.2 (6.1)	24.5 (6.3)
5	25.5 (5.7)	26.7 (5.5)	24.3 (5.4)	25.4 (5.0)	23.5 (8.3)	22.5 (7.4)	22.2 (7.4)	22.5 (7.6)
6	25.2 (6.2)	25.5 (5.8)	24.0 (5.5)	25.3 (5.7)	24.8 (6.6)	23.9 (6.1)	23.0 (5.8)	24.0 (6.0)
7	24.8 (6.3)	24.9 (5.6)	23.5 (5.4)	25.0 (4.5)	22.9 (7.1)	21.1 (6.6)	21.9 (5.7)	22.4 (6.7)
8	26.1 (5.4)	26.4 (5.2)	24.7 (4.9)	26.1 (5.6)	22.5 (7.2)	21.0 (6.5)	21.8 (6.0)	22.1 (6.2)

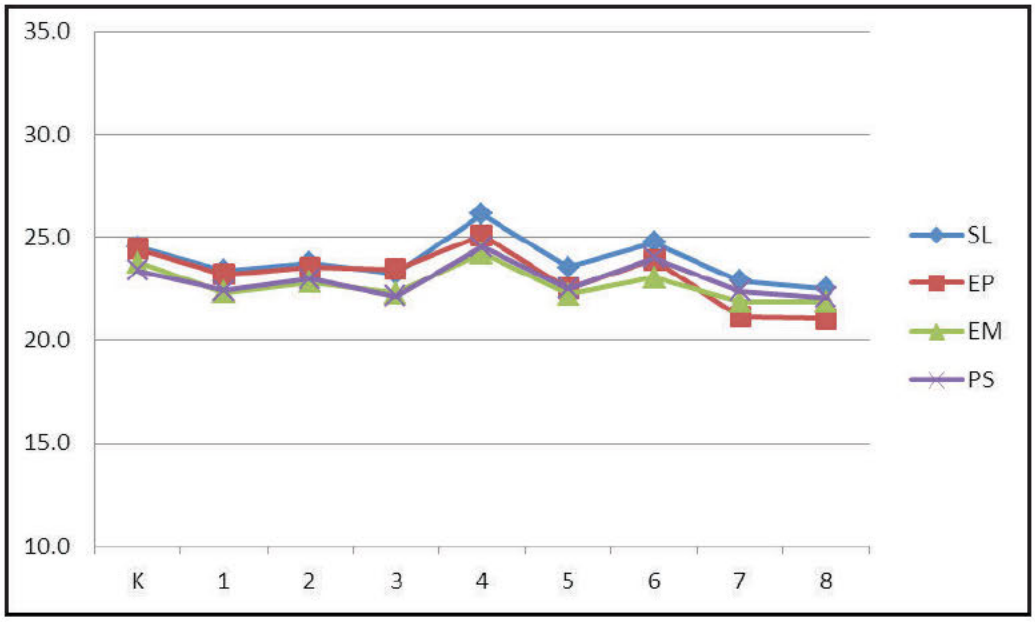
Note: SL = Skills for Learning; EP = Empathy; EM = Emotion Management; and PS = Problem Solving

Figure 2.1
DESSA-SSE Mean Raw Scores by Grade for Parent Raters



Note: SL = Skills for Learning; EP = Empathy; EM = Emotion Management; and PS = Problem Solving

Figure 2.2
DESSA-SSE Mean Raw Scores by Grade for Teacher Raters



Note: SL = Skills for Learning; EP = Empathy; EM = Emotion Management; and PS = Problem Solving

We also examined the Social–Emotional Composite raw scores for rater differences. There were significant differences between the ratings provided by parents and teachers. Consequently, we prepared separate norms for parents and teachers. This is to be expected, as behavior often differs across environments and in the presence of different adults.

Mean-score differences also indicated gender differences, which reflect real disparities in how boys and girls behave. Table 2.6 presents the raw-score mean differences between boys and girls by grade and by rater. The values in Table 2.6 were obtained by subtracting the mean raw score obtained by boys from the mean raw score obtained by girls at each grade for each scale. A positive value indicates that girls scored higher than boys. As shown in Table 2.6, for both parent and teacher raters, the mean scale raw scores for girls are consistently one-half to five points higher than those for boys.

To evaluate the practical significance of the mean raw-score differences between sexes, we calculated *d*-ratios, a measure of effect size, which are also presented in Table 2.6. This statistic is computed by subtracting one mean from the other and dividing that difference by the average standard deviation for the two groups being contrasted. According to Cohen (1988), *d*-ratio values of less than .2 are negligible. Those between .2 and .5 reflect a small effect size. Those between .5 and .8 indicate a medium effect size, and *d*-ratios greater than .8 indicate a large effect size. Table 2.6 presents the *d*-ratios comparing boys' and girls' scores on each scale averaged across the K–5 grade levels. For example, the first column indicates that, for parent raters, on the Skills for Learning scale, averaged across grades, the *d*-ratio comparing boys' and girls' ratings was .34. All of the *d*-ratios presented in Table 2.6 would be classified as small, with the exception of Problem Solving as rated by teachers, which just met the criterion (.50) for a medium effect size. The data in this table indicate that, as a group, girls consistently show more behaviors related to social and emotional competence than boys, but the magnitude of this difference tends to be small.

Girls in the DESSA-SSE standardization sample earned higher scores than boys on each scale. In order to preserve these important differences in social–emotional competencies, we constructed the raw score to *T*-score norms conversion tables based on both genders combined. Consequently, it can be expected that girls will, on average, earn higher scores on the DESSA-SSE than boys. This reflects the natural differences between the genders and establishes a single set of social–emotional competency expectations that applies equally to both genders.

After determining that norms would be constructed by rater, we examined the distributions of raw scores for normality. The cumulative frequency distributions for the scales all approached normality, but they were slightly positively skewed. For this reason, we decided to compute norms using normalization procedures. This was accomplished by fitting the obtained frequency distribution for each scale to normal probability standard scores via the obtained percentile ranks. We eliminated minor irregularities in raw score to standard score progressions by smoothing, and we followed these procedures for all the scales. For the four scales and the Social–Emotional Composite, we computed standard scores (*T*-scores with a mean of 50 and a standard deviation of 10) based on percentile score distributions separately for teacher and parent raters. We based the Social–Emotional Composite *T*-score on the percentile distribution of the sum of the four *T*-scores corresponding to the DESSA-SSE scales for each case. We selected the *T*-score metric because of its familiarity to professionals and because it facilitates interpretation of the results and comparison to scores obtained from other, similar scales.

Table 2.6

DESSA-SSE Raw Score Gender Differences by Grade

	Parent Raters				Teacher Raters			
	SL	EP	EM	PS	SL	EP	PS	SL
K	1.8	2.5	1.2	1.8	1.4	1.2	0.8	1.5
1	2.0	2.3	2.0	2.4	0.9	0.5	0.5	1.1
2	2.1	3.5	1.7	2.0	3.8	3.8	2.3	3.6
3	1.1	1.3	0.4	1.2	3.3	3.3	2.6	3.5
4	2.2	1.9	1.5	2.3	5.0	4.0	3.8	4.8
5	1.2	0.9	0.5	1.6	4.9	4.0	4.6	4.5
Mean	1.7	2.1	1.2	1.9	3.2	2.8	2.4	3.2
d-ratios	0.34	0.41	0.24	0.40	0.48	0.44	0.39	0.50

Note: SL = Skills for Learning; EP = Empathy; EM = Emotion Management; and PS = Problem Solving

CHAPTER 3: Psychometric Properties

Reliability

The reliability of an assessment tool like the DESSA-SSE is defined as "the consistency of scores obtained by the same person when reexamined with the same test on different occasions, or with different sets of equivalent items, or under other variable examining conditions" (Anastasi, 1988, p. 102). DESSA-SSE scale reliability was assessed using several methods. First, the internal reliability coefficient for each scale was computed. Second, test-retest reliability of each scale was assessed. Finally, interrater reliability (two raters evaluating the same student) for each scale was determined.

Internal Reliability

Internal reliability (or internal consistency) refers to the extent to which the items on the same scale or assessment instrument measure the same underlying construct. We determined internal consistency using Cronbach's alpha (Cronbach, 1951). The internal reliability coefficients were based on the individuals in kindergarten through 5th grade included in the DESSA-SSE standardization sample. Parents provided ratings on 978 children; teachers provided ratings on 1098 children. The internal consistency estimates for each scale were calculated according to rater and are presented in Table 3.1. The results indicate that the DESSA-SSE scales have excellent internal reliability.

The Social–Emotional Composite reliability was computed using the formula provided by Nunnally and Bernstein (1994) for the reliability of a linear combination. The Social–Emotional Composite coefficients for parent raters (.96) and teacher raters (.98) both well exceed the .90 value for a total score suggested by Bracken (1987) and also meet the desirable standard described by Nunnally (1978, p. 246).

The internal reliability coefficients for the four social–emotional competence scales range from a low of .82 (Emotion Management—Parent Raters) to a high of .93 (Skills for Learning—Teacher Raters). All of the reliability coefficients exceed the .80 minimum suggested by Bracken (1987).

Table 3.1

Internal Reliability (Alpha) Coefficients for the DESSA-SSE Scales by Rater

	Raters	
Scales	**Parents**	**Teachers**
Social–Emotional Composite	**.96**	**.98**
Skills for Learning	**.85**	**.93**
Empathy	**.87**	**.92**
Emotion Management	**.82**	**.90**
Problem Solving	**.86**	**.92**

Standard Errors of Measurement—The standard error of measurement (SE_M) is an estimate of the amount of error in observed scores, expressed in standard score units (i.e., *T*-scores). We obtained the SE_M for each of the DESSA-SSE scale *T*-scores directly from the internal reliability coefficients using the formula

$$SE_M = SD\sqrt{1 - reliability}$$

where SD is the theoretical standard deviation of the *T*-score (10) and the appropriate reliability coefficient is used.

The SE_Ms for each DESSA-SSE scale are presented in Table 3.2 according to rater. Note that the values of the SE_M vary with the size of the reliability coefficient—the higher the reliability, the smaller the standard error of measurement.

Table 3.2
Standard Errors of Measurement for the DESSA-SSE *T*-Scores by Rater

	Raters	
Scales	Parents	Teachers
Social–Emotional Composite	2.00	1.40
Skills for Learning	3.87	2.65
Empathy	3.61	2.83
Emotion Management	4.24	3.16
Problem Solving	3.74	2.83

Test-Retest Reliability

The correlation between scores obtained for the same child on two separate occasions is another indicator of the reliability of an assessment instrument. The correlation of this pair of scores is the test-retest reliability coefficient (*r*), and the magnitude of the obtained value informs us about the degree to which random changes influence the scores (Anastasi, 1988). To investigate the test-retest reliability of the DESSA-SSE, a group of teachers (n = 38) and a group of parents (n = 54) rated the same child on two different occasions separated by an interval of 4 to 8 days. Demographic information on this diverse convenience sample is provided in Table 3.3.

Table 3.3
Sample Characteristics for the DESSA-SSE Test-Retest Reliability Study

	Parent Sample		Teacher Sample	
	n	%	*n*	%
Size of Sample	54		38	
Age (grade)				
Mean	3.6		4.2	
SD	2.6		2.6	
Gender				
Boys	31	57%	17	45%
Girls	23	43%	20	53%
Missing	0	0%	1	2%
Race				
American Indian/Alaskan Native	0	0%	1	3%
Asian	2	4%	0	0%
Black/African American	3	6%	11	29%
Native Hawaiian/Pacific Islander	0	0%	1	3%
White	51	94%	22	58%
Other	2	4%	1	3%
Hispanic Ethnicity	2	4%	2	5%
Region of Residence				
Northeast	19	35%	5	13%
South	15	28%	0	0%
Midwest	15	28%	23	60%
West	2	4%	10	27%
Other/Missing	3	5%	0	0%

The results of this study are shown in Table 3.4. All of the correlations are significant ($p < .01$) and high in magnitude ranging from $r = .84$ (Skills for Learning and Empathy—Parent Raters) to $r = .94$ (Skills for Learning and Problem Solving—Teacher Raters). The correlation coefficients for the Social–Emotional Composite are $r = .87$ for Parent Raters and $r = .95$ for Teacher Raters. These findings indicate that the DESSA-SSE scales have good test-retest reliability.

Table 3.4

Test-Retest Reliability Coefficients for Two DESSA-SSE Ratings by the Same Teacher or Same Parent for the Same Child Over a 4- to 8-day Interval

	Raters	
Scales	**Parents**	**Teachers**
Social–Emotional Composite	**.87***	**.95***
Skills for Learning	**.84***	**.94***
Empathy	**.84***	**.90***
Emotion Management	**.86***	**.91***
Problem Solving	**.85***	**.94***

*$p < .01$

Interrater Reliability

The correlation between scores obtained for the same child at approximately the same time by two different raters is an indicator of the interrater reliability of an assessment instrument. The magnitude of the correlations between these scores tells us about the degree of similarity in the different raters' perceptions of the child's behavior. We examined the interrater reliability of the DESSA-SSE by comparing ratings obtained from two parents who live in the same household with the child ($n = 51$) or two teachers, or a teacher and teacher aide, who either work in the same classroom or see the same child in different classrooms for core academic subjects ($n = 51$). In these studies, the sample size (n) refers to the number of *unique pairs* of adults rating a child. Demographic information on these two samples is presented in Table 3.5.

Table 3.5

Sample Characteristics for the DESSA-SSE Interrater Reliability Study

	Parent Sample		Teacher Sample	
	n	%	*n*	%
Size of Sample	51		51	
Age (Grade)				
Mean	3.2		2.5	
SD	2.6		2.3	
Gender				
Boys	26	51%	29	57%
Girls	25	49%	22	43%
Race				
American Indian/Alaskan Native	2	4%	1	2%
Asian	2	4%	1	2%
Black/African American	3	6%	16	31%
Native Hawaiian/Pacific Islander	0	0%	0	0%
White	47	92%	29	57%
Other	0	0%	1	2%
Hispanic Ethnicity	7	14%	6	12%
Region of Residence				
Northeast	13	25%	28	55%
South	14	27%	0	0%
Midwest	12	24%	10	20%
West	12	24%	13	25%

The correlations of a set of ratings obtained for the same children by two parents or two teachers (or a teacher and a teacher aide) are provided in Table 3.6. These results indicate that pairs of parents or pairs of teachers who saw the children in the same environment rated the children very similarly. All the correlations are significant ($p < .01$) and moderate to high in magnitude. The Social–Emotional Composite correlations are .77 for Parent Raters and .81 for Teacher Raters. For the four individual scales, the values range from .68 (Emotion Management—Parent Raters; Empathy—Teacher Raters) to .85 (Problem Solving—Teacher Raters).

Table 3.6

Interrater Reliability Coefficients for Two DESSA-SSE Ratings by Two Parents or Two Teachers for the Same Child

	Raters	
Scales	Parents	Teachers
Social–Emotional Composite	.77*	.81*
Skills for Learning	.76*	.80*
Empathy	.71*	.68*
Emotion Management	.68*	.70*
Problem Solving	.76*	.85*

*$p < .01$

Stability of DESSA-SSE Ratings

The correlation coefficients reported for the test-retest and interrater reliability studies indicate that the pairs of raters in each study ranked the children similarly. However, the coefficients do not indicate the actual similarity in the mean scores. Tables 3.7a and 3.7b provide the pretest and posttest mean scale scores and standard deviations received by the children in the test-retest study by parents and teachers, respectively.

Table 3.7a

DESSA-SSE Pretest and Posttest Mean Scale Scores and Standard Deviations—Parent Raters

Scales	Pretest	Posttest
Skills for Learning	45.8 (9.3)	46.4 (9.7)
Empathy	47.1 (10.4)	47.6 (10.4)
Emotion Management	46.9 (9.2)	47.4 (9.6)
Problem Solving	47.0 (9.4)	47.7 (10.1)

Table 3.7b

DESSA-SSE Pretest and Posttest Mean Scale Scores and Standard Deviations—Teacher Raters

Scales	Pretest	Posttest
Skills for Learning	47.9 (11.1)	47.7 (11.1)
Empathy	46.7 (10.1)	46.3 (10.8)
Emotion Management	47.4 (11.7)	46.7 (12.9)
Problem Solving	47.6 (11.5)	47.7 (12.2)

For Parent Raters, on average, the absolute value of the test-retest difference on the four social–emotional competence scales was less than one *T*-score point (0.6). The results for teacher raters were very similar. On the four social–emotional competence scales, the mean absolute value of the test-retest difference was 0.35 for Teacher Raters. These results demonstrate that the DESSA-SSE ratings are very stable across a 4- to 8-day interval for both Parent and Teacher Raters.

Reliability Study Summary

The results of the several reliability studies of the DESSA-SSE indicate that the instrument is reliable for assessing children's social–emotional competencies. The results of the internal consistency analysis demonstrate that the DESSA-SSE meets standards suggested by Bracken (1987). The test-retest study shows that raters rank the children's scores on the DESSA-SSE similarly over time. This similarity in mean scores over time in the absence of specific social–emotional instruction is important in enabling the examination of changes in scores as a result of instruction. The results of the interrater reliability study show that different parents and teachers also rank children's scores similarly. The stability studies further indicate that not only the rankings but also the actual mean scale scores received by the children at different points in time are quite similar.

Validity

The validity of a test "concerns what the test measures and how well it does so" (Anastasi, 1988, p. 139). More specifically, validity "is the degree to which all the accumulated evidence supports the intended interpretation of test scores for the proposed purpose" (APA, 1999, p. 11). According to the *Standards for Educational and Psychological Testing* (APA, 1999), the sources of validity evidence can be conceptualized in various ways. We investigated the validity of the DESSA-SSE in regard to *test content* (content validity), *internal structure* (presented in Chapter 2 on the development of the DESSA-SSE), and *test–criterion relationships* (criterion validity).

Content-Related Validity

This type of validity assesses the degree to which the domain measured by the test is represented by the test items. With respect to the DESSA-SSE, content-related validity addresses how well the 36 items represent the domain of behavioral characteristics related to social–emotional competence, especially those presented in the *Second Step* program.

As detailed in Chapter 2, we based the items on the DESSA-SSE on a thorough review of the literature on social–emotional competence, positive youth development, and resilience in school-aged children. We also based the items on the DESSA-SSE, in part, on our earlier publication, the Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999a, 1999b), which has its own research base (see DECA Manual). From the initial pool of items on the DESSA, we were able to find nine items that reflected each of the four constructs in the *Second Step* program. This provided sufficient coverage of these constructs and also ensured the high internal consistency documented in the previous section on reliability.

Criterion-Related Validity

This type of validity measures the degree to which the scores on the assessment predict either an individual's performance on an outcome or criterion measure or the status or group membership of the individual. As a measure of behaviors related to social–emotional competence, scores on the DESSA-SSE should predict social–emotional functioning of school-aged children. To test this hypothesis, we obtained DESSA-SSE ratings on two samples of students. First, we obtained a sample of children who were reported by their parent or teacher to be receiving special education services under the seriously emotionally disturbed (SED) classification. The children in this SED sample ($N = 78$) were matched to a comparison group (referred to as the Regular Education or RE sample; $N = 78$) selected from the national standardization sample. Matching variables included: type of rater, gender of the child, and age of the child. Table 3.8 provides descriptive information on both samples and shows that the two groups were demographically similar.

Table 3.8

Sample Characteristics for the DESSA-SSE Criterion Validity Study

	SED Sample		Regular Education Sample	
	n	%	*n*	%
Size of Sample	78		78	
Rater				
Parent	38	49%	38	49%
Teacher	40	51%	40	51%
Age (Grade)				
Mean	3.36		3.36	
SD	2.58		2.58	
Gender				
Boys	51	65%	51	65%
Girls	27	35%	27	35%
Race				
American Indian/Alaskan Native	0	0%	2	2.6%
Asian	1	1.3%	0	0%
Black/African American	16	20.5%	16	20.5%
Native Hawaiian/Pacific Islander	0	0%	0	0%
White	54	69.2%	54	69.2%
Other	3	3.8%	0	0%
Hispanic Ethnicity	9	11.5%	9	11.5%
Region of Residence				
Northeast	19	24.4%	27	34.6%
South	18	23.1%	21	26.9%
Midwest	27	34.6%	14	17.9%
West	14	17.9%	14	17.9%
Other/Missing	0	0%	2	2.6%

We compared the RE and SED groups using analysis of variance (ANOVA) procedures to contrast the four social–emotional competence scale scores. An independent *t*-test was used to compare the Social–Emotional Composite scores for the two groups. The results of these analyses are provided in Table 3.9, which documents that there were large and significant differences between the mean scores of the SED and RE samples on all DESSA-SSE scales. The mean standard score differences and other results reported in Table 3.9 clearly show that the ratings of the two groups differed significantly, despite the similarity in demographic characteristics. All scale comparisons were significant ($p < .001$).

In addition to being statistically significant, the means of the two groups on each scale differed by at least 90% of a standard deviation (*d*-ratios range from .90 to 1.5). The *d*-ratio is a measure of the size of the difference between the mean scores of two groups, expressed in standard deviation units. According to commonly accepted guidelines for interpreting *d*-ratios (Cohen, 1988), *d*-ratios of .2, .5 and .8 are interpreted as small, medium and large, respectively. Therefore, all of the effect sizes reported in Table 3.9 would be characterized as large. The Social–

Emotional Composite also differentiated between the two samples in this study (t (155) = 8.2, p < .001; d = 1.3). These results provide strong evidence of the validity of the DESSA-SSE scales in discriminating between groups of students identified as seriously emotionally disturbed and their regular education peers.

Table 3.9

Mean *T*-Scores, Standard Deviations, and Difference Statistics for the DESSA-SSE Criterion Validity Study

	SED Sample		Regular Education Sample
Skills for Learning			
Mean	37.8		48.4
SD	7.9		9.8
F Value		55.9*	
d-ratio		1.2	
Empathy			
Mean	39.0		47.4
SD	8.5		10.1
F Value		31.9*	
d-ratio		0.9	
Emotion Management			
Mean	35.6		48.5
SD	8.0		9.7
F Value		83.7*	
d-ratio		1.5	
Problem Solving			
Mean	37.0		48.2
SD	8.2		9.9
F Value		59.5*	
d-ratio		1.2	
Social–Emotional Composite			
Mean	35.9		47.8
SD	8.0		10.1
t Value		8.2*	
d-ratio		1.3	

*p < .001

Validity Study Summary

The content-related evidence provided in this chapter related the DESSA-SSE items to both the research and practice literatures on social–emotional competence in children. The results of the criterion-related validity studies demonstrated that DESSA-SSE scores do differentiate between groups of children with and without the special education designation of serious emotional disturbance.

CHAPTER 4: Administration and Scoring

General Administration Guidelines

The DESSA-SSE can be completed by a parent or family member (this includes parents, stepparents, foster parents, guardians, or other relatives who live with the child), a teacher (this includes teachers, teacher aides, instructional assistants, etc.), after-school program staff, and staff from social service, mental health, or child welfare programs. The person who actually completes the DESSA-SSE and provides the ratings is referred to as the "rater." The person who administers, scores, and interprets the DESSA-SSE ratings is referred to as the "user." The qualifications of raters and users were described in Chapter 1. The following general guidelines for completing the DESSA-SSE should be reviewed with rater:

1. The rater should complete the DESSA-SSE during a quiet time when there are few distractions.
2. The rater should base the ratings on direct observations of the child, considering only behaviors that he/she has actually seen. The rater should not consider behaviors that were reported to occur in other classrooms or settings.
3. The rater should consider only those behaviors that have occurred in the past 4 weeks.
4. When completing the DESSA-SSE, the rater should avoid comparing the child being rated to other children. The rating should be based solely on the number of times the child being rated exhibited the behaviors, not how frequently the child exhibits the behavior in comparison to other children in the classroom.
5. The rater should answer every item. An inability to complete the items indicates that the rater may not know the child well, and another rater should be used.

Specific Directions for Completing the DESSA-SSE Record Form

The DESSA-SSE Record Form is used for the administration and scoring of the rating scale. There is one form, which is used for all children in kindergarten through 5th grade. The same form is used for all raters.

A ballpoint pen works best when completing the DESSA-SSE Record Form. The DESSA-SSE is a multipart, carbonless form, and the rater should be told to press firmly so that the information provided on the front page will transfer to the inside page.

Demographic Information

The top of the front page (see Figure 4.1) provides spaces (A) to record demographic information about the child being rated, including name, gender, date of birth, age, school/organization, classroom/program, and grade. There are also spaces to record the rater's name, relationship to the child (teacher, mother, etc.), and the date of the rating. Raters should complete all of the information at the top of the front page of the DESSA-SSE Record Form.

Completing the Ratings

The front page of the DESSA-SSE Record Form contains the following directions to the rater:

> This form describes a number of behaviors seen in some children. Read the statements that follow the phrase: ***During the past 4 weeks, how often did the child…*** and place a check mark in the box underneath the word that tells how often you saw the behavior. Please answer each question carefully. There are no right or wrong answers. If you wish to change your answer, put an ☒ through it and fill in your new choice as shown to the right. Please do not skip any items.

Figure 4.1
DESSA-SSE Record Form Page 1

DESSA
DEVEREUX STUDENT STRENGTHS ASSESSMENT

The Devereux Student Strengths Assessment

Second Step Edition (DESSA-SSE)

secondSTEP

For Grades K-5

(A)

Paul A. LeBuffe, Jack A. Naglieri, & Valerie B. Shapiro

Child's Name: Sean Evans Gender: Male Date of Birth: 7/15/03 Age: 8
School/Organization: Central Elementary Classroom/Program: 3-A Grade: 3
Person Completing This Form: Mrs. Evans Relationship to Child: Mother Date of Rating: 10/14/11

This form describes a number of behaviors seen in some children. Read the statements that follow the phrase: ***During the past 4 weeks, how often did the child...*** and place a check mark in the box under the word that tells how often you saw the behavior. Please answer each question carefully. There are no right or wrong answers. If you wish to change your answer, put an ☒ through it and fill in your new choice as shown to the right. Please do not skip any items.

Never	Rarely	Occasionally	Frequently	Very Frequently
☒	☑	☐	☐	☐

Item #	During the past 4 weeks, how often did the child...	Never ✓	Rarely ✓	Occasionally ✓	Frequently ✓	Very Frequently ✓
1.	remember important information?	☐	☐	☐	☐	☑
2.	keep trying when unsuccessful?	☐	☐	☐	☐	☑
3.	cope well with insults and mean comments?	☑	☐	☐	☐	☐
4.	take steps to achieve goals?	☐	☐	☐	☑	☐
5.	get along with different types of people?	☐	☐	☐	☑	☐
6.	seek out additional knowledge or information?	☐	☐	☐	☑	☐
7.	take an active role in learning?	☐	☐	☐	☑	☐
8.	respect another person's opinion?	☐	☐	☑	☐	☐
9.	do routine tasks or chores without being reminded?	☐	☐	☐	☑	☐

(B)

[redacted]

Kaplan #20398

The 36 items that comprise the DESSA-SSE are also on the front page. The rater responds to each item by placing a checkmark in the appropriate box (B) underneath the words "Never," "Rarely," "Occasionally," "Frequently," or "Very Frequently."

Use of the DESSA-SSE with Raters Who Have Limited English Proficiency

If the rater has difficulty reading and completing the DESSA-SSE because of limited English proficiency, the DESSA-SSE items may be read to him/her. The person reading the DESSA-SSE for the rater should try not to influence the ratings. The items should be read in an even, calm tone and explanations of the items or examples should not be given. The person reading the DESSA-SSE should also not provide any feedback or react in any way to the rater's responses.

Scoring the DESSA-SSE

Once the form is completed, scoring the DESSA-SSE is simple. All of the scoring is done on the DESSA-SSE Record Form. Complete scoring directions are given below and also on page 2 of the Record Form.

Step 1: Recording the DESSA-SSE Item Raw Scores

The DESSA-SSE user should review the Record Form and make sure that all of the demographic information was provided and that all 36 items were completed. If any information or items were left blank, the rater should be asked to complete the information. When the information on the Record Form is complete, the DESSA-SSE user tears off the perforated strip at the left edge of the form, breaks the seal on the top and bottom of the form, and opens the form from left to right. All of the scoring is performed on the inside page of the DESSA-SSE Record Form as described below.

When the rater places a checkmark in a box on the front page to indicate the rating for a given item, that checkmark transfers to a corresponding box on the inside page of the Record Form. (See (C) on Figure 4.2.) The boxes on the inside page contain numbers that are the raw score values associated with each rating as follows:

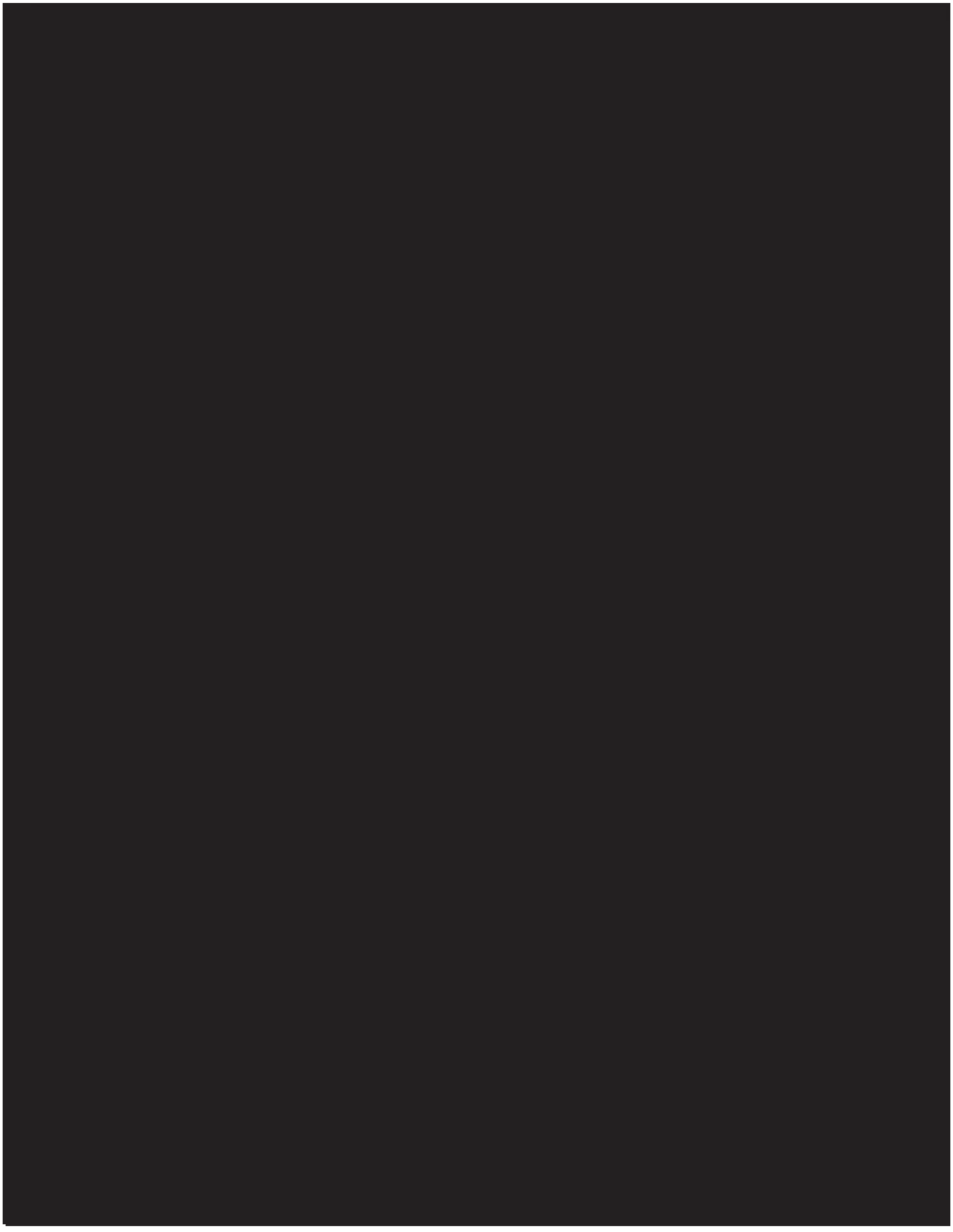
Never = 0 Rarely = 1 Occasionally = 2 Frequently = 3 Very Frequently = 4

The first step in scoring the DESSA-SSE is to copy the raw score value (0–4) from the box that was checked to the empty box on the same line (D).

- *Treatment of Missing Items*—As indicated above, the rater should be encouraged to complete *all* of the DESSA-SSE items. Although the DESSA-SSE items were developed so that they could be rated by parents, teachers, and after-school staff, there may be some unusual instances in which a rater cannot rate a specific behavior. For instance, a rater may be unable to provide a rating for item #3—"cope well with insults and mean comments" if the rater has never observed the child being teased or bullied. In circumstances such as these, the following guidelines apply:

1. There can be no more than two (2) items left blank on the entire DESSA-SSE.
2. There can be no more than one (1) item left blank on any individual scale.
3. If the above two conditions are met, the value that appears in the *grey shaded rectangular box* on the inside page of the Record Form should be used as the item raw score for that item. This value is the typical or most common score for this item in the national standardization sample. Using this score minimizes biasing the scale either up or down. The user should circle the rating on the inside page of the Record Form to note that the rating was not provided by the rater and was originally left blank.
4. If there are three (3) or more blank items in total, or two (2) or more blank items on any individual scale, the DESSA-SSE should be discarded and another rater found.

Figure 4.2
DESSA-SSE Record Form Page 3



Step 2: Calculating the DESSA-SSE Scale Raw Scores

The Scale Raw Scores for the four scales (Skills for Learning, Empathy, Emotion Management, and Problem Solving) are obtained by adding the raw scores for all of the items that comprise each scale. On the inside page of the Record Form, the boxes where the item raw score values were copied are arranged in four columns, one for each scale. (See Figure 4.2.) To compute the Scale Raw Scores for the four scales, add the item raw scores in each column and enter the sums in the boxes (E) provided at the bottom of the inside page. Then copy these Scale Raw Scores on the first row of the Scale Score Summary Table (F) as shown in Figure 4.2.

Step 3: Determining DESSA-SSE *T*-Scores and Percentiles

The norms tables for the DESSA-SSE can be found on the DESSA-SSE Individual Student Profile (see Figure 4.3), which can be downloaded from the DESSA Tools page found in the K–5 Teaching Guide section of http://www.secondstep.org and also in Appendix A of this manual. These norms tables are used to determine the *T*-scores and percentile scores for each scale. Both the Individual Student Profile and the norms tables in Appendix A work the same way. To determine the *T*-score and percentile score for each scale, first determine the appropriate norms table (parent or teacher rater) to use. It is very important that the correct table is used. In Appendix A, Table 1 is for ratings obtained from a parent, and Table 2 is for ratings obtained from a teacher or after-school or program staff.

Then, using the appropriate table, find the Scale Raw Score in the column of numbers under the appropriate scale name. The corresponding *T*-score is found on the same row in the far left column labeled "*T*-Score." Similarly, the percentile score is found on the same row in the far right column labeled "Percentile Score." For example, in the Parent norms table, the Skills for Learning Scale Raw Score of 32 corresponds to a *T*-score of 62 and a percentile score of 88. (The meaning of these scores is explained in Chapter 5.) The *T*-score and percentile score for each scale should be recorded on the second and third lines of the Scale Score Summary Table.

Step 4: Determining the *T*-Score and Percentile Score for the Social–Emotional Composite

To calculate the *T*-score and percentile score for the Social–Emotional Composite, begin by adding the *T*-scores for four scales as indicated by the "+" and "=" signs on the Scale Score Summary Table. The sum of these four *T*-scores is treated as the *Scale Raw Score* for the Social–Emotional Composite. The corresponding *T*-score and percentile score are determined in exactly the same way as described for the four scales in Step 3. On the appropriate norms table, find the Social–Emotional Composite raw score (that is the sum of the four scale *T*-scores) in the column of numbers under "Social–Emotional Composite." The corresponding *T*-score and percentile score are found on the same row in the far left and far right columns, respectively. Record these values on the Scale Score Summary Table. The sum of the four scale *T*-scores is used to determine the Social–Emotional Composite Scale Raw Score so that each scale has an equal contribution to or influence on this summary scale.

Step 5: Creating the Individual Student Profile

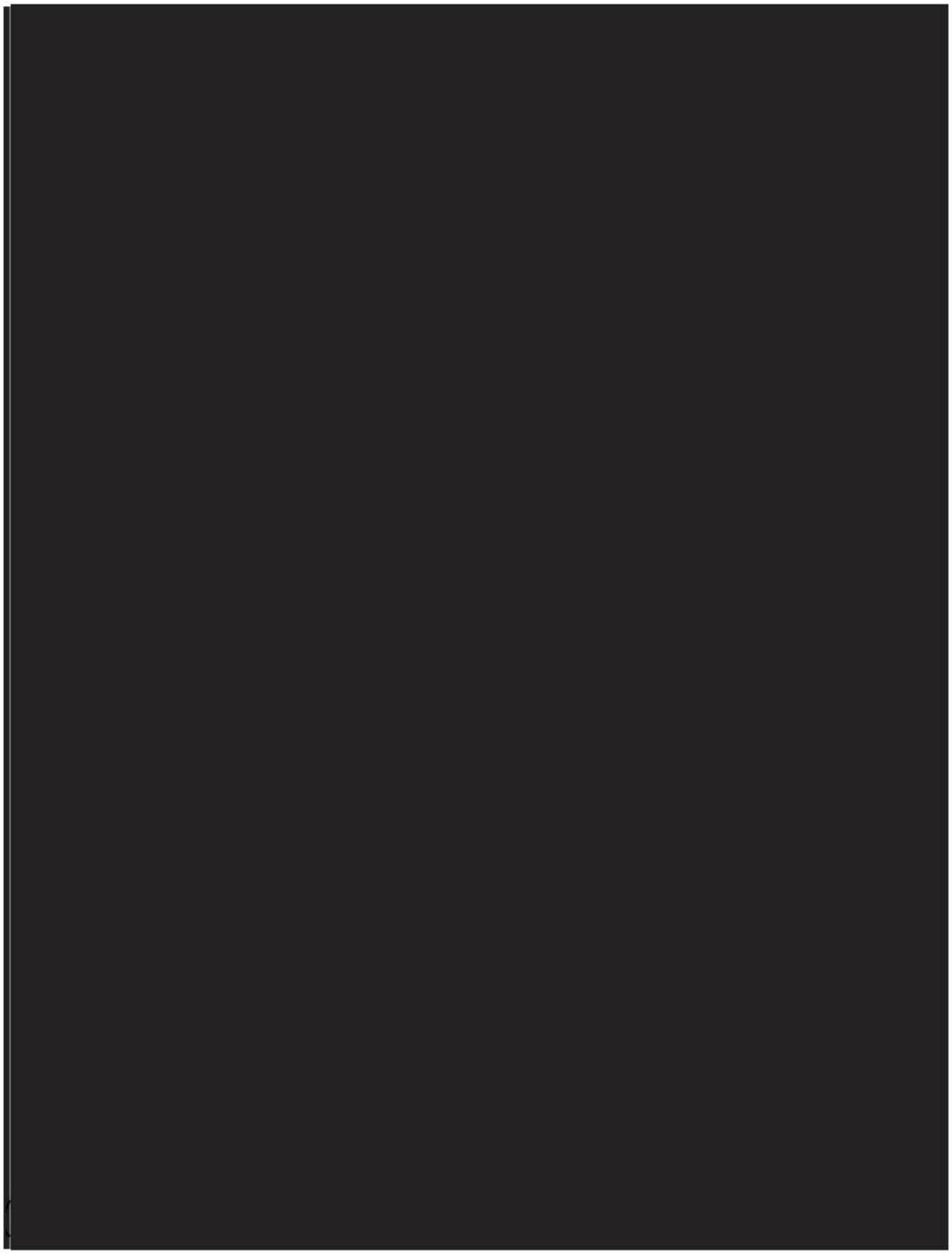
The DESSA-SSE Individual Student Profile (see G in Figure 4.3) is used to graphically display the child's scores on the five DESSA-SSE scales. To create the DESSA-SSE Individual Student Profile, circle the raw score that the child received on each scale, making sure to use the appropriate version (parent or teacher) of the Individual Student Profile. Then connect the four scale *T*-scores to create the child's scale score profile. Do not connect the score for the Social–Emotional Composite to the other four scales.

Step 6: Determining the Description for Each Scale

An interpretation key is provided at the bottom of the Individual Student Profile (H). This key indicates the preferred terminology for describing various DESSA-SSE scores. For each scale, high scores (*T*-scores of 60 and above) are referred to as *strengths*. This range of scores is indicated by gray shading on the Individual Student Profile. *T*-scores that fall between 41 and 59 inclusive are described as *typical*. Low scores (*T*-scores of 40 and below) are

described as a *need for instruction.* This range of scores is indicated by orange shading on the Individual Student Profile. The appropriate description for each scale should be recorded on the Scale Score Summary Table. When the Individual Student Profile has been created and all of the information on the Scale Score Summary Table provided, the scoring of the DESSA-SSE is completed.

Figure 4.3
DESSA-SSE Individual Student Profile



CHAPTER 5: Interpretation

Effective interpretation of any scale demands that the user be familiar with what is being measured, the scores that are provided, and how these scores should be interpreted. When interpreting DESSA-SSE scores, the DESSA-SSE user should always consider the following general guidelines. First, the DESSA-SSE user should have a thorough understanding of the meanings and appropriate uses of the various standard scores and profiles. These topics are discussed later in this chapter.

Second, the DESSA-SSE user should remember that the DESSA-SSE scores that a child receives are based on the ratings provided by a single adult. Therefore, the scores reflect the unique interactions between the child and that adult. A different rater who sees the child in a different context may well provide somewhat different ratings.

Third, always consider the child's and family's cultural heritage and family background when interpreting DESSA-SSE findings. Although we took many steps during the development of the DESSA-SSE to avoid items that might elicit different responses from various racial and ethnic groups, cultural differences in the prevalence and meaning of specific DESSA-SSE items might exist, as they would with any assessment. Therefore, the DESSA-SSE user should be sensitive to cultural differences when interpreting the DESSA-SSE.

The Center for Mental Health Services of the federal Substance Abuse and Mental Health Services Administration has published Cultural Competence Standards (Center for Mental Health Services, 2001). Among the provider competencies, the following are particularly relevant to DESSA-SSE users:

- An understanding of psychosocial stressors and traumas such as war, immigration, socioeconomic status, racism, and discrimination for various groups.
- Differences in the meaning of specific behaviors across different groups.
- Nuances of language and the meaning of items.
- Differences between "culturally acceptable" behaviors and behavioral concerns across different groups.
- Who constitutes the family in various groups.

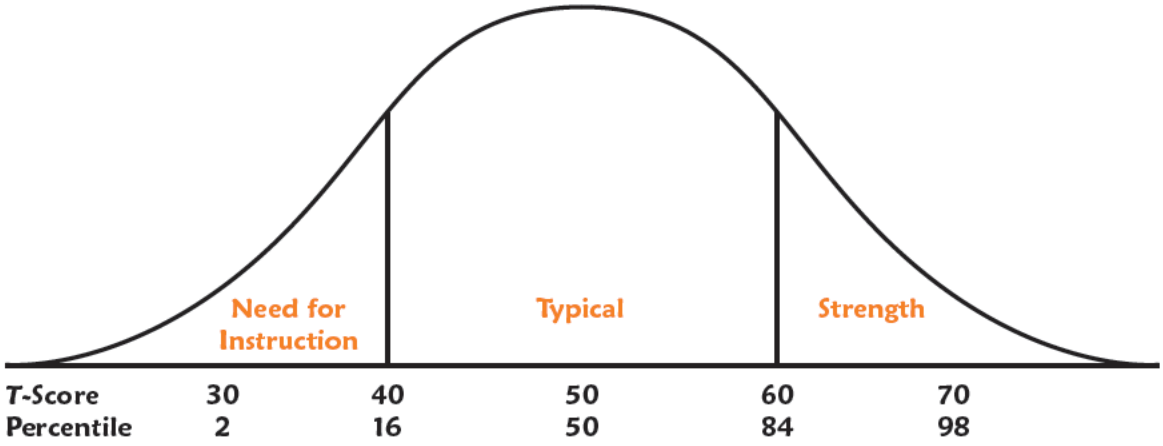
Knowledge of the child's and family's culture will result in more sensitive interpretations of DESSA-SSE findings and more useful recommendations to both parents and teachers. More specific issues regarding interpretation of the DESSA-SSE are provided in the remainder of this chapter. This will include a summary of the types of scores the scale yields, the mechanics of how these scores should be examined, and methods for their interpretation.

Types of Scores Given

Raw Scores

The raw score for each DESSA-SSE scale provides little information about the overall level of the child's performance. Raw scores are converted to standard scores so that the separate scales of the DESSA-SSE can be directly compared. Standard scores also enable the comparison of a given child's behavior to that of the other children in the standardization sample. The DESSA-SSE provides two standard scores, percentile scores and *T*-scores. Figure 5.1 shows the relationships between percentile scores, *T*-scores, the normal distribution, and the *T*-score range descriptions for the DESSA-SSE scales. These standard scores and descriptions are described below.

Figure 5.1
Relationship of DESSA-SSE *T*-Scores, Percentile Scores, and the Normal Curve



Percentile Scores

DESSA-SSE raw scores are converted to percentile scores using Appendix A - Table 1 (for Parent Raters) and Table 2 (for Teacher Raters) or the Individual Student Profile. Percentile scores compare the child's behavior to that of other children who have been rated using the DESSA-SSE. The percentile score indicates the percentage of children in the standardization sample who earned the same or lower raw score. For example, if a child earns a percentile score of 65, that means that 65% of the children in the standardization sample earned the same or a lower raw score. DESSA-SSE percentile scores range from a minimum of 1 to a maximum of 99.

Percentile scores are easy to understand, but they do have a significant disadvantage—they cannot be easily compared and cannot be used in mathematical computations. The principal problem with percentile scores is that differences between the scores do not have the same meaning across the 1 through 99 scale. That is, a five-point difference between percentile ranks of 90 and 95 is a much greater distance on the normal curve than a five-point difference between percentile ranks of 50 and 55. This means that comparing two DESSA-SSE scales using percentile scores might lead the practitioner to conclude that a significant difference exists when it does not. Consequently, although percentile scores are useful for describing the relative standing of a child versus the other children in the standardization sample, they should not be used to compare a child's scores across DESSA-SSE scales, because their meaning changes at different points on the normal distribution. It is important to remember that these scores should ***never*** be averaged or used in mathematical computations. Only DESSA-SSE *T*-scores should be used for that purpose.

T-Scores

Each DESSA-SSE *T*-score is a standard score set to have a mean of 50 and standard deviation of 10. Like the percentile scores, *T*-scores are based on the ratings received by the children in the standardization sample. In contrast to percentile scores, however, DESSA-SSE *T*-scores have the same meaning throughout their range. The five-point difference between the *T*-scores of 50 and 55 is equivalent to the five-point difference between *T*-scores of 40 and 45. In both cases, the difference between these sets of scores is one half of a standard deviation. For this reason, *T*-scores should always be used when reporting the DESSA-SSE results and when comparing scores earned on the various scales. On the DESSA-SSE, *T*-scores range from 28 to 72.

T-Score Range Descriptions for the DESSA-SSE Scales

The DESSA-SSE raw score and the corresponding percentile and *T*-score reflect strengths related to social and emotional competence in children, and therefore, high scores are desirable. For example, when rating how often a child "keeps trying when unsuccessful" or "offers to help somebody," the higher the score the better. Consequently, high scale scores are desirable as well.

For clarity and consistency, and to aid in the communication of results, we recommend using the following *T*-score range descriptions when reporting DESSA-SSE results. The term "need for instruction" should be used to describe DESSA-SSE scale *T*-scores of 28 to 40 inclusive. Scores of 40 or less mean that the child was rated as showing few behaviors associated with the particular social–emotional strength. Children with scores in this range can be considered at risk for exhibiting or developing social–emotional problems. On each scale, approximately 16% of the children in the standardization sample received scores in the need for instruction range.

Scale scores of 41 to 59 inclusive should be described as "typical." Approximately 68% of the children in the standardization sample received scores in the typical range.

DESSA-SSE scale *T*-scores of 60 to 72 inclusive should be described as "strengths." Approximately 16% of the children in the standardization sample received scale scores in the strength range

The various descriptions and their relationship to DESSA-SSE *T*-scores are summarized in Table 5.1. The DESSA-SSE user should keep in mind that these are guidelines for the categorization and interpretation of DESSA-SSE scores and should not be rigidly applied, overinterpreted, or reified. Although the DESSA-SSE scales have very high internal reliability (see Table 3.1), and consequently minimal standard errors of measurement (see Table 3.2), DESSA-SSE users should take measurement error into account when interpreting DESSA-SSE scores. This is particularly important when the *T*-score obtained by the child is close to the thresholds presented above.

Table 5.1

Descriptive Categories and Interpretations of DESSA-SSE *T*-Scores

60 and above	Strength
41–59	Typical
40 and below	Need for Instruction

The Meaning and Interpretation of the DESSA-SSE Scales

The DESSA-SSE Scales

The following brief descriptions are to aid in the interpretation of the DESSA-SSE scales. More thorough information on the content and meaning of these scales is presented in Chapter 1 and in the *Second Step* program materials.

Skills for Learning: A child's ability to use the skills of listening, focusing attention, self-talk, and assertiveness.

Empathy: A child's ability to identify and label emotions in himself/herself and others and take on others' perspectives.

Emotion Management: A child's ability to cope with strong emotions and express them in socially acceptable ways.

Problem Solving: A child's ability to effectively handle personal challenges and interpersonal conflicts in prosocial ways.

The Social–Emotional Composite: This scale gives an overall indication of the child's social and emotional competencies. This scale is the most reliable overall indicator of strengths within the DESSA-SSE. Because it characterizes the child's social and emotional strengths with a single number, the Social–Emotional Composite is particularly useful in outcome measurement and program evaluation.

Interpretation of the DESSA-SSE

Interpretation of the DESSA-SSE results proceeds in a step-wise fashion from the most general indicator of the child's social and emotional status, the Social–Emotional Composite, to the more specific information provided by the four scales.

Step 1: The Social–Emotional Composite

First, examine the Social–Emotional Composite *T*-score and note the corresponding range description (i.e., strength, typical, need for instruction). This is the broadest and the most reliable index of the child's social and emotional well-being. Research with the Devereux Student Strengths Assessment (LeBuffe, Shapiro, & Naglieri, 2009) as well as the analogous Total Protective Factors Scales from the Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999a) indicates that high scores on these strength-based summary scales are associated with children who are functioning well in academic and other environments. These children tend to have few behavioral concerns and are likely to be resilient when faced with risk and adversity. The Social–Emotional Composite *T*-score is a highly reliable indicator of the child's overall social and emotional functioning and serves as the starting point in interpreting the DESSA-SSE. The score a child receives on the Social–Emotional Composite Scale also provides a frame of reference for examining the four specific scale scores.

Step 2: Examining Scale Scores

Examine the four separate DESSA-SSE scales and note the *T*-scores and corresponding strength, typical, and need for instruction ranges. Examination of the separate DESSA-SSE scale *T*-scores provides useful information about the specific strengths and needs of the child in relation to the social–emotional competencies taught by the *Second Step* program. Examination of the DESSA-SSE Individual Student Profile is particularly useful at this step, as the visual depiction of the scale scores can make patterns easier to discern.

Use of the DESSA-SSE in Program Evaluation

An important use of the DESSA-SSE is evaluating the impact of the *Second Step* program and similar social–emotional learning programs in a scientifically rigorous manner. Results from this evaluation can assist schools and other programs in documenting the positive effects of the *Second Step* program. The approach outlined below has a significant advantage in being able to measure outcomes at the individual child, classroom, grade, building, and even district level. The foundation of this approach is to compare the *T*-scores obtained by a child at two points in time. The essential question is whether the amount of change between the earlier (pretest) and subsequent (posttest) ratings exceeds that which can be expected by chance variation—that is, whether the amount of change is statistically significant.

Changes in a child's *T*-scores over time can be evaluated when a period of at least 4 weeks between the ratings has elapsed, so that the latter rating represents a new sample of behaviors. Whenever possible, the same rater should

be used for both the pretest and the posttest rating. It is essential, however, that the same type of rater (parent or teacher) be used at both administrations.

The statistical significance of the difference between pretest and posttest scores can be determined using the method described by Atkinson (1991). This approach involves the comparison of the obtained posttest score with a range of scores that represents the variability expected by both regression to the mean and measurement error based on the pretest score. To obtain the values needed to assess the significance of the pretest–posttest score differences, we calculated the standard error of prediction (SEp). The standard error of prediction is used instead of the standard error of measurement because we are concerned about the predictability (or consistency) between the pretest and posttest scores. See Atkinson (1991) for more details or Naglieri, LeBuffe, and Pfeiffer (1993) for more discussion.

Posttest confidence ranges were calculated for each DESSA-SSE scale and are presented in Appendix B, Tables 1 (for parents) and 2 (for teachers). To determine if significant change has occurred, the pretest and posttest scores should be compared using the following method:

Using the appropriate table based on the rater, find the pretest DESSA-SSE *T*-score in the first column labeled "Pretest Obtained Score." Read across the table to the column that corresponds to the DESSA-SSE scale being evaluated. If the posttest DESSA-SSE *T*-score falls within the posttest range provided in the table, there has been no significant change in the child's score. If, however, the posttest score falls above the posttest range, we can conclude that the child's score has shown significant improvement. If the posttest score falls below the range provided, then we conclude that the score has shown significant worsening. For example, for a parent rating and a pretest score of 39 on Skills for Learning, the posttest score would have to be 51 or higher to conclude that there has been reliable (statistically significant) improvement in the child's score. With extreme pretest scores, there may be values in the posttest range that exceed possible scores on the DESSA-SSE (i.e., less than 28 or greater than 72).

Clinical Meaningfulness of Statistically Reliable Change

Using the "dual criterion" approach recommended by Jacobsen and Truax (1991), we can also determine the clinical meaningfulness of the change in the child's *T*-scores. Whereas statistical significance "refers to real differences as opposed to ones that are illusory, questionable or unreliable," clinical significance refers to "its ability to make a difference in people's lives" (Jacobsen & Truax, 1991, p. 12). It is important to note that the second criterion of clinical meaningfulness is only addressed when there has been statistically reliable improvement in the child's scores.

When statistically reliable change has occurred, the second criterion, the clinical meaningfulness of the change, is determined by the examination of the value of the posttest *T*-scores. Clinically meaningful improvement can be further divided into optimal outcomes and favorable outcomes.

An optimal outcome is found when a child with a pretest score in the need for instruction or typical range shows reliable change in a positive direction, as determined using Appendix B, Tables 1 and 2, *and* the posttest *T*-score falls in the strength range. A favorable outcome occurs when a child with a pretest *T*-score in the need for instruction or typical range shows reliable improvement, but the posttest *T*-score is below 60.

Ultimately, the best possible outcome for a child is that he or she has all of the DESSA-SSE social and emotional competency scales rated in the strength range. Conversely, the worst outcome for a child is to have all of the DESSA-SSE scales rated in the need for instruction range.

This dual-criterion approach to examining the effectiveness of curricula, interventions, and strategies to help children develop social and emotional competencies is a very flexible and powerful tool. This approach enables

the DESSA-SSE user to look at the effectiveness of the *Second Step* program on a scale-by-scale and child-by-child basis. For instance, we might observe that a child with two scales in the need for instruction range at pretest showed significant improvement in one area but not the other. By using this method, we can determine which children benefited in which areas.

This approach can also be used for program evaluation and quality improvement efforts by aggregating findings across children, classrooms, grades, schools, and so forth. In this instance we might note, for instance, that of all the children with a need for instruction in a given area, 80% showed reliable improvement in comparison to a second area in which only 60% of children showed improvement. Such data can inform quality improvement and professional development activities.

The use of the Social–Emotional Composite (SEC) is recommended for overall program evaluation because it is the most general and reliable indicator of a child's social–emotional competence. We recommend that programs examine the effectiveness of the *Second Step* program in two ways using the SEC. First, the percentage of children with SEC *T*-scores in the need for instruction, typical, and strength range should be determined at both pretest and posttest. Obviously, at posttest we would expect to see a greater percentage of children with social and emotional strengths and fewer children with a need for instruction. Second, using the pretest–posttest comparison table in Appendix B, a school can readily determine what percentage of children showed a statistically reliable improvement in their overall social and emotional competence. It is important to look at both aspects due to ceiling effects. That is, children who at pretest had high scores on the SEC may not show significant improvement because they are already in the strength range. Reporting both the percentage of children in each range at pretest and posttest as well as the percentage of children who showed statistically reliable change presents a balanced view of program effectiveness.

The authors of the DESSA-SSE and the staff of the Committee for Children believe that this tool will assist schools, after-school programs, and others using the *Second Step* program to promote the social–emotional competencies of children in order to facilitate their success in school and life.

References

American Psychological Association. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Atkinson, L. (1991). Three standard errors of measurement and the Wechsler memory scale-revised. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 136–138.

Berkley, A. (2008). *Equivalence of computerized and paper-and-pencil administration of a measure of resilience in school-age children*. Unpublished manuscript, George Mason University.

Bracken, B. A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment, 5*, 313–326.

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (1998). *Positive youth development in the United States. Research findings on evaluations of the positive youth development programs*. Washington DC: United States Department of Health and Human Services National Institute for Child Health and Human Development.

Center for Mental Health Services. (2001). *Cultural competence standards in managed care mental health service: Four underserved/underrepresented racial/ethnic groups*. Substance Abuse and Mental Health Services Administration (SAMHSA), U.S. Department of Health and Human Services. Retrieved October 16, 2008, from http://mentalhealth.samhsa.gov/publications/allpubs/ sma00-3457/default.asp.

Cohen, J. (1988). *Statistical power analysis in the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Committee for Children. (2011). *Second step: Skills for social and academic success*. Seattle, WA: Author.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432.

Epstein, M. H. (2004). *Behavioral and emotional rating scale examiner's manual* (2nd ed.). Austin, TX: PRO-ED.

Individuals with Disabilities Education Improvement Act, 20 U.S.C. § 1400 *et seq.* (2004). (Reauthorization of the Individuals with Disabilities Education Act of 1990).

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12–19.

Kanner, A. D., Coyne, J. C., Schaefer, C., & Lazarus, R. S. (1981). Comparison of two modes of stress measurement: Daily hassles and uplifts versus major life events. *Journal of Behavioral Medicine, 4,* 1–39.

LeBuffe, P. A., & Naglieri, J. A. (1999a). *Devereux early childhood assessment user's guide*. Lewisville, NC: Kaplan Press.

LeBuffe, P. A., & Naglieri, J. A. (1999b). *Devereux early childhood assessment technical manual.* Lewisville, NC: Kaplan Press.

LeBuffe, P. A., & Shapiro, V. B. (2004). Lending "strength" to the assessment of preschool social–emotional health. *The California School Psychologist, 9,* 51–61.

LeBuffe, P. A., Shapiro, V. B., & Naglieri, J. A. (2009). *The Devereux student strengths assessment (DESSA) assessment, technical manual, and user's guide.* Lewisville, NC: Kaplan Press.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50,* 741–749.

Naglieri, J. A., LeBuffe, P. A., & Pfeiffer, S. I. (1993). *Devereux behavior rating scale- school form.* San Antonio: The Psychological Corporation.

National Association for the Education of Young Children. (1987). *Standardized testing of young children through 8 years of age (position statement).* Washington, DC: Author.

Nickerson, A. B., & Fishman, C. (2009). Convergent and divergent validity of the Devereux Student Strengths Assessment. *School Psychology Quarterly, 24,* 48–59.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Payton, J. W., Wardlaw, D. M., Graczyk, P. A., Bloodworth, M. R., Tompsett, C. J., Weissberg, R. P. (2000). Social and emotional learning: A framework for promoting mental health and reducing risk behavior in children and youth. *Journal of School Health, 70,* 179–185.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior assessment system for children-second edition: Manual.* Circle Pines, MN: American Guidance Service.

U.S. Census Bureau. (2008). *Statistical abstract of the United States* (127th ed.). Washington, DC: Author.

Werner, E. E., & Smith, R. S. (1982). *Vulnerable but invincible: A study of resilient children.* New York: McGraw-Hill.

Werner, E. E., & Smith, R. S. (1992). *Overcoming the odds: High-risk children from birth to adulthood.* Ithaca, NY: Cornell University Press.

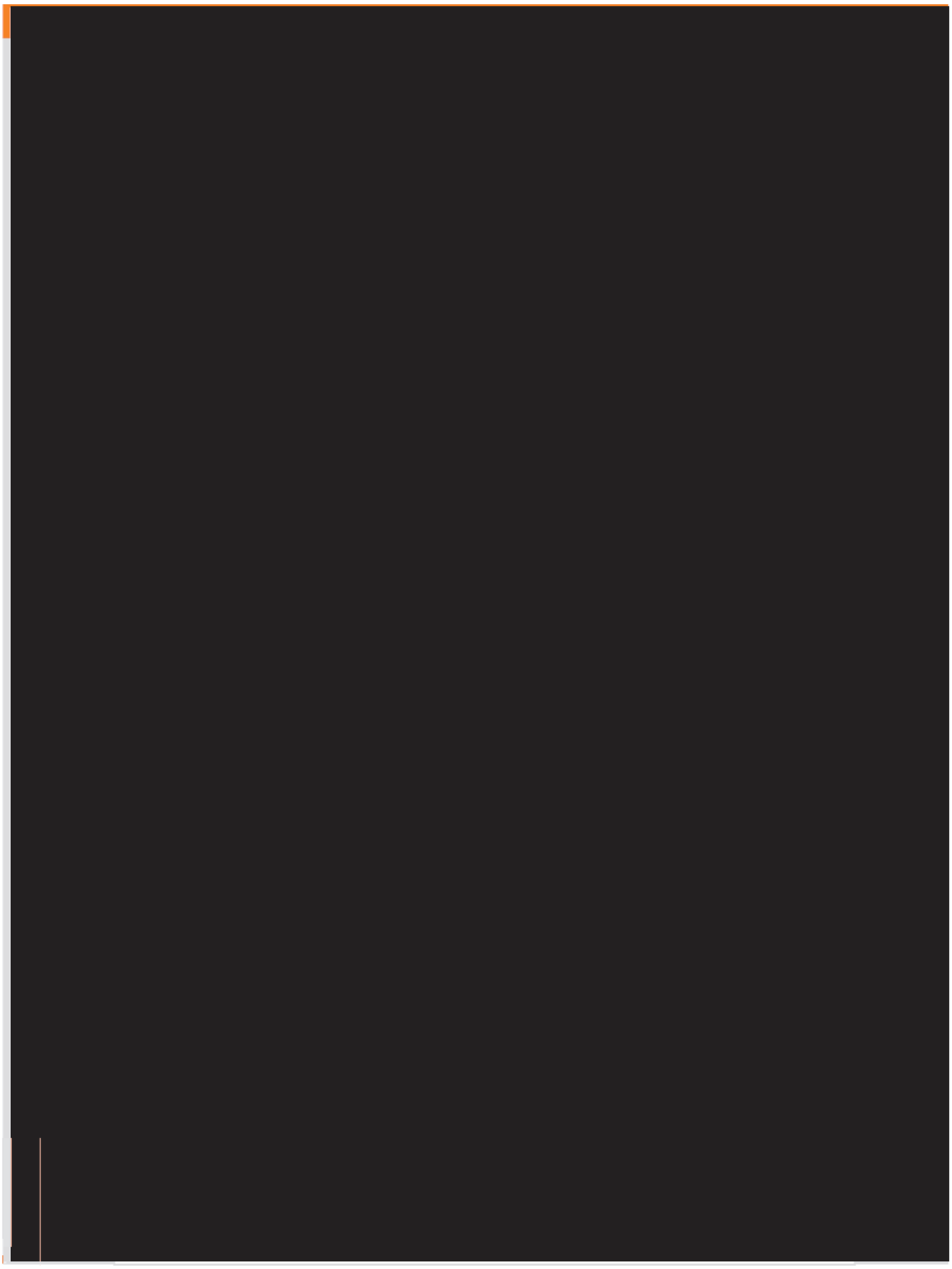
Work, W. C., Cowen, E. L., Parker, G. R., & Wyman, P. A. (1990). Stress resilient children in an urban setting. *Journal of Prevention, 11,* 3–17.

Wright, M. O. D., & Masten, A. S. (2005). Resilience processes in development: Fostering positive adaptation in the context of adversity. In S. Goldstein & R. Brooks (Eds.), *Handbook of resilience in children* (pp. 17–37). New York: Kluwer Academic/Plenum Press.

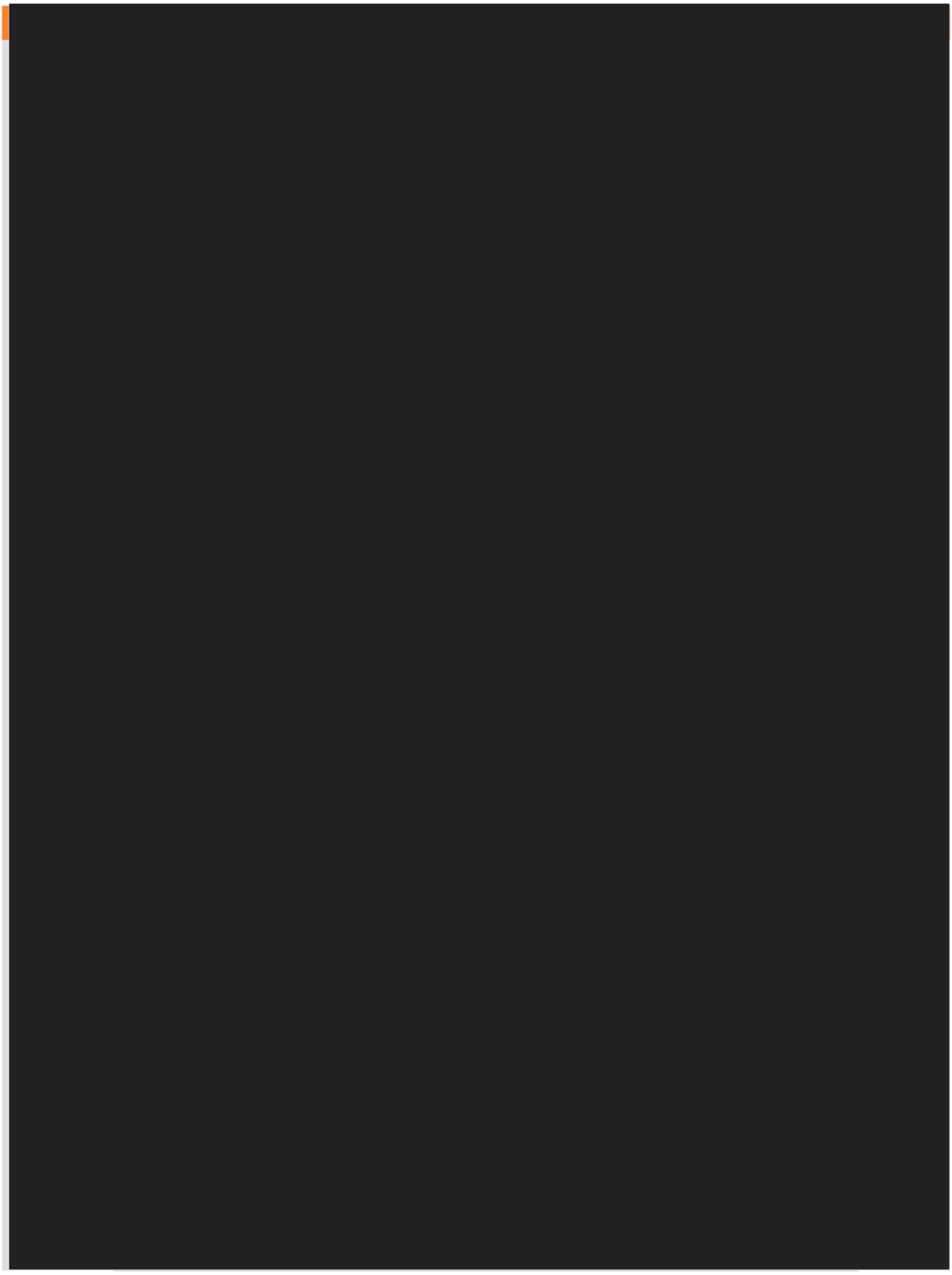
About Devereux

Established in 1912, Devereux is the nation's largest independent, not-for-profit provider of treatment, habilitation, and special education services for children, adolescents, and adults who have a wide range of emotional, behavioral, and developmental challenges. Headquartered in Villanova, PA, and providing services in 11 states, Devereux offers a full continuum of campus- and community-based programs to more than 20,000 individuals annually. The purpose of Devereux is to provide high-quality human services to children, adolescents, adults, and families with special needs. Devereux provides services in a caring and humane way to foster human potential and to contribute to the person's health, social, psychological, and educational well-being. More information about Devereux can be found on its website, www.Devereux.org.

Appendix A—Table 1



Appendix A—Table 2



Appendix B—Table 1

Values Needed for Significance When Comparing DESSA–SSE *T*–Scores Obtained Before and After Intervention for Parent Raters ($p = .05$)

Pretest Obtained Score	Skills for Learning Posttest Range	Empathy Posttest Range	Emotion Management Posttest Range	Problem Solving Posttest Range	Social–Emotional Composite Posttest Range
72	59–79	60–79	57–79	59–79	65–77
71	58–78	59–78	56–78	58–78	64–76
70	57–77	58–77	56–78	57–77	63–75
69	57–76	57–76	55–77	56–76	62–74
68	56–75	56–75	54–76	56–75	61–73
67	55–75	55–74	53–75	55–75	60–72
66	54–74	55–73	52–74	54–74	59–71
65	53–73	54–73	51–73	53–73	58–70
64	52–72	53–72	51–73	52–72	57–69
63	51–71	52–71	50–72	51–71	56–68
62	51–70	51–70	49–71	50–70	55–68
61	50–69	50–69	48–70	50–69	54–67
60	49–68	49–68	47–69	49–69	53–66
59	48–68	48–67	46–68	48–68	52–65
58	47–67	48–66	46–68	47–67	51–64
57	46–66	47–66	45–67	46–66	51–63
56	45–65	46–65	44–66	45–65	50–62
55	44–64	45–64	43–65	44–64	49–61
54	44–63	44–63	42–64	44–63	48–60
53	43–62	43–62	41–63	43–63	47–59
52	42–62	42–61	41–63	42–62	46–58
51	41–61	41–60	40–62	41–61	45–57
50	40–60	41–59	39–61	40–60	44–56
49	39–59	40–59	38–60	39–59	43–55
48	38–58	39–58	37–59	38–58	42–54
47	38–57	38–57	37–59	37–57	41–53
46	37–56	37–56	36–58	37–56	40–52
45	36–56	36–55	35–57	36–56	39–51
44	35–55	35–54	34–56	35–55	38–50
43	34–54	34–53	33–55	34–54	37–49
42	33–53	34–52	32–54	33–53	36–49
41	32–52	33–52	32–54	32–52	35–48
40	32–51	32–51	31–53	31–51	34–47
39	31–50	31–50	30–52	31–50	33–46
38	30–49	30–49	29–51	30–50	32–45
37	29–49	29–48	28–50	29–49	32–44
36	28–48	28–47	27–49	28–48	31–43
35	27–47	27–46	27–49	27–47	30–42
34	26–46	27–45	26–48	26–46	29–41
33	25–45	26–45	25–47	25–45	28–40
32	25–44	25–44	24–46	25–44	27–39
31	24–43	24–43	23–45	24–44	26–38
30	23–43	23–42	22–44	23–43	25–37
29	22–42	22–41	22–44	22–42	24–36
28	21–41	21–40	21–43	21–41	23–35

Appendix B—Table 2

Values Needed for Significance When Comparing DESSA–SSE *T*–Scores Obtained Before and After Intervention for Teacher Raters ($p = .05$)

Pretest Obtained Score	Skills for Learning Posttest Range	Empathy Posttest Range	Emotion Management Posttest Range	Problem Solving Posttest Range	Social–Emotional Composite Posttest Range
72	63–78	63–78	61–78	63–78	67–76
71	63–77	62–77	60–78	62–77	66–75
70	62–76	61–76	59–77	61–76	65–74
69	61–75	60–75	58–76	60–75	64–73
68	60–74	59–74	57–75	59–74	63–72
67	59–73	58–73	57–74	58–73	62–71
66	58–72	57–72	56–73	57–72	61–70
65	57–71	56–71	55–72	56–71	60–69
64	56–70	55–71	54–71	55–70	59–68
63	55–69	54–70	53–70	54–70	58–67
62	54–68	53–69	52–69	53–69	57–66
61	53–67	52–68	51–69	53–68	56–65
60	52–66	52–67	50–68	52–67	55–64
59	51–65	51–66	49–67	51–66	54–63
58	50–65	50–65	49–66	50–65	53–62
57	49–64	49–64	48–65	49–64	52–61
56	49–63	48–63	47–64	48–63	51–60
55	48–62	47–62	46–63	47–62	51–59
54	47–61	46–61	45–62	46–61	50–58
53	46–60	45–60	44–61	45–60	49–57
52	45–59	44–60	43–60	44–59	48–56
51	44–58	43–59	42–60	43–59	47–55
50	43–57	42–58	41–59	42–58	46–54
49	42–56	41–57	40–58	41–57	45–53
48	41–55	40–56	40–57	41–56	44–52
47	40–54	40–55	39–56	40–55	43–51
46	39–53	39–54	38–55	39–54	42–50
45	38–52	38–53	37–54	38–53	41–49
44	37–51	37–52	36–53	37–52	40–49
43	36–51	36–51	35–52	36–51	39–48
42	35–50	35–50	34–51	35–50	38–47
41	35–49	34–49	33–51	34–49	37–46
40	34–48	33–48	32–50	33–48	36–45
39	33–47	32–48	31–49	32–47	35–44
38	32–46	31–47	31–48	31–47	34–43
37	31–45	30–46	30–47	30–46	33–42
36	30–44	29–45	29–46	30–45	32–41
35	29–43	29–44	28–45	29–44	31–40
34	28–42	28–43	27–44	28–43	30–39
33	27–41	27–42	26–43	27–42	29–38
32	26–40	26–41	25–43	26–41	28–37
31	25–39	25–40	24–42	25–40	27–36
30	24–38	24–39	23–41	24–39	26–35
29	23–37	23–38	22–40	23–38	25–34
28	22–37	22–37	22–39	22–37	24–33

Appendix C—DESSA-SSE Classroom/Group Profile

The use of the Classroom/Group Profile is helpful by indicating areas where many children have Strengths, Typical scores, or Areas of Need. This interpretive and planning tool is available at www.secondstep.org and is designed so that each child in a given group is a row and each scale on the DESSA-SSE is a column. Each cell in the profile then presents the T-score and associated descriptive range obtained by the child on a particular scale with green indicating a Strength, blue indicating a Typical score, and red indicating a Need for Instruction. For example, as the Classroom/Group Profile below shows, in this classroom nearly a third of the children had a Need for Instruction in Skills for Learning, whereas less than a quarter evidenced a Need in Empathy, Emotion Management, or Problem Solving. These results can be used to inform the delivery of the Second Step program.

By empirically and reliably identifying common social and emotional needs and strengths of children in a given group, the DESSA-SSE enables teachers and other childserving professionals to be more strategic and intentional in their efforts to promote resilience at the universal level. In addition to identifying common Needs for Instruction, the Classroom/Group Profile also identifies common Strengths, which can be incorporated into strategies to support strength-based approaches.

The DESSA-SSE Classroom/Group Profile is also a useful tool for evaluating outcomes at the universal level. Changes in the Classroom/Group Profile over time can be noted. As the Second Step program is implemented, one would expect to see fewer red cells (indicating a Need for Instruction) and more blue (Typical) and green (Strength) cells. It is particularly important to note changes in the color distribution of the cells in the Social and Emotional Composite column over time, as it is the most general indicator of the children's social and emotional competencies.

DESSA-SSE Classroom Profile

Teacher's Name: Mr. Holland **Date:** 5/31/2013

First	Last	Birth Date	SL	EP	EM	PS	SEC
Joe	Green		56	62	61	59	61
Sarah	Smith		68	43	36	54	51
Andrew	Diaz		46	37	62	49	49
Matt	Brown		37	51	59	63	53
James	Johnson		50	48	32	43	43
Mary	Williams		48	34	44	46	43
Michael	Jones		34	42	49	51	44
Robert	Davis		29	55	34	42	39
Linda	Miller		52	68	50	65	60
Jennifer	Wilson		54	50	55	33	48
Betty	Taylor		58	53	66	46	56
Charles	Jackson		36	30	50	30	35
Carol	White		51	65	57	45	55
Kimberly	Robinson		60	51	55	63	58
Annie	Walker		31	56	49	48	46
Rebecca	King		44	35	46	67	48
Daniel	Lopez		53	48	31	54	46
Stephanie	Baker		32	42	41	34	36
			SL	EP	EM	PS	SEC
Number of children GREEN			2	3	3	4	2
Number of children BLUE			10	11	11	11	13
Number of children RED			6	4	4	3	3