



Devereux Student Strengths Assessment, Second Edition (DESSA 2)

Technical Manual

August 2024

Support Portal Version
(Items & Norms Redacted)

Jennifer L. Robitaille, Evelyn S. Johnson, Paul A. LeBuffe, Jack A. Naglieri, & Valerie B. Shapiro*



Copyright © 2024 by Riverside Insights.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of Riverside Insights.

Editorial and Design: Abella Publishing Services, LLC

Riverside Insights
One Pierce Place
Suite 900W
Itasca, IL 60143

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES	7
ACKNOWLEDGMENTS	10
ABOUT THE AUTHORS	12
FOREWORD	15
CHAPTER 1: Introduction	17
Purpose of the Revision	19
Background	20
Description of the DESSA 2 Measures	21
DESSA 2	22
DESSA 2 mini	23
Uses of the DESSA 2 Measures	23
Uses of the DESSA 2	23
Uses of the DESSA 2 mini	24
Values Guiding the Development and Use of the DESSA 2 Measures	25

Qualifications of Users and Raters of the DESSA 2 Measures	.26
Qualifications of Users	.26
Qualifications of Raters.	.27
Restrictions for Use.	.27
 CHAPTER 2: Development and Standardization	.29
Development of the DESSA 2 Items	.30
Feedback from Expert Reviewers	.31
National Standardization.	.32
Selection of the DESSA 2 Standardization Sample	.32
Representativeness of the DESSA 2 Standardization Sample	.33
Organization of the DESSA 2 Items into Scales.	.37
Development of the DESSA 2 mini Items	.40
Norming Procedures	.41
 CHAPTER 3: Psychometric Properties	.44
Reliability	.45
Internal Reliability	.45
Standard Error of Measurement.	.47
Alternate Form Reliability	.47
Test-Retest Reliability	.48
Interrater Reliability	.50
Reliability Study Summary	.53
Validity	.54
Content-Related Validity	.54
Construct-Related Validity	.55
Use of the DESSA 2 mini as a Screener for Social and Emotional Competence.	.64
Validity Study Summary	.66

CHAPTER 4: Administration and Scoring	.67
General Administration Guidelines	.68
Specific Directions for Completing the DESSA 2 Measures	.69
Completing the Ratings	.70
Use of the DESSA 2 Measures with Raters Who Have Limited English Proficiency.	.71
Treatment of Missing or Blank Items	.71
Scoring the DESSA 2 mini	.72
Scoring the DESSA 2	.72
CHAPTER 5: Interpretation	.74
General Interpretation Guidelines.	.75
Considerations Regarding the Use of the DESSA 2 with Students with Special Needs.	.76
Types of Scores Given	.77
Note Regarding Raw Scores	.77
Standard Scores.	.77
T-Score Range Descriptions	.78
The Meaning and Interpretation of the DESSA 2 Scales	.80
The DESSA 2 mini Social-Emotional Total	.80
The DESSA 2 Scales	.80
The DESSA 2 Social-Emotional Composite	.80
Basic Interpretation of the DESSA 2 mini	.81
Basic Interpretation of the DESSA 2	.81
Step 1: The Social-Emotional Composite	.81
Step 2: Examining Scale Scores	.81
Step 3: Identifying Specific Strengths and Need for Instruction Items	.81

Advanced Interpretation of the DESSA 2 Measures	.85
Progress Monitoring	.85
Evaluating Programmatic Outcomes and Impact	.86
Interpretation Example.	.88
Step 1: Examination of the Social-Emotional Composite	.89
Step 2: Examining Scale Scores	.89
Step 3: Individual Item Analysis	.89
Use of the DESSA 2 Measures within a Multi-Tiered System of Support (MTSS).	90
Use of the DESSA 2 Measures at Tier 1	90
Use of the DESSA 2 Measures at Tier 2.	.91
Use of the DESSA 2 Measures at Tier 3.	.91
 APPENDICES	 .93
APPENDIX A: DESSA 2 mini Norms (Educator Ratings)	.94
APPENDIX B: DESSA 2 Norms (Educator Ratings)	.99
APPENDIX C: List of Data Collection Sites by State	102
 REFERENCES	 104
 About Aperture Education, A Riverside Insights Company	 .111

LIST OF TABLES AND FIGURES

CHAPTER 2

TABLE 2.1 DESSA 2 Standardization Sample Characteristics by Grade and Gender

TABLE 2.2 DESSA 2 Standardization Sample Characteristics by Geographic Region and Grade

TABLE 2.3 DESSA 2 Standardization Sample Characteristics by Race and Geographic Region

TABLE 2.4 DESSA 2 Standardization Sample Characteristics by Hispanic/Latinx Ethnicity and Geographic Region

TABLE 2.5 Evidence of DIF for Final Set of DESSA 2 Items

TABLE 2.6 DESSA 2 Raw Score Means and Standard Deviations by Grade Bands

TABLE 2.7 DESSA 2 mini Raw Score Means and Standard Deviations by Grade Bands

TABLE 2.8 DESSA 2 Standard Score Gender Differences by Scale

TABLE 2.9 DESSA 2 mini Standard Score Gender Differences by Form

CHAPTER 3

TABLE 3.1 Internal Reliability (Alpha) Coefficients and Standard Errors of Measurement for the Four DESSA 2 mini Forms

TABLE 3.2 Internal Reliability (Alpha) Coefficients and Standard Errors of Measurement for the DESSA 2 Scales

TABLE 3.3 DESSA 2 mini Alternate Form Reliability: Correlation Coefficients

TABLE 3.4 Four Form Equivalence: DESSA 2 mini Social-Emotional Total *T*-scores

TABLE 3.5	Sample Characteristics for the DESSA 2 mini and DESSA 2 Test–Retest Reliability Study
TABLE 3.6	Test–Retest Reliability and Stability Results for DESSA 2 mini Ratings
TABLE 3.7	Test–Retest Reliability and Stability Results for DESSA 2 Ratings
TABLE 3.8	Sample Characteristics for the DESSA 2 mini and DESSA 2 Interrater Reliability Study
TABLE 3.9	Interrater Reliability Results for DESSA 2 mini Ratings
TABLE 3.10	Interrater Reliability Results for DESSA 2 Ratings
TABLE 3.11	Fit Indices for the DESSA 2 Six-Scale Model and Two Alternative Models
TABLE 3.12	Comparisons Between the DESSA 2 Six-Scale Model and Two Alternative Models
TABLE 3.13	Cumulative Frequencies of the <i>T</i> -score Difference Between the Highest and Lowest DESSA 2 Scale Scores
TABLE 3.14	Sample Characteristics for the DESSA 2 Construct Validity Sample
TABLE 3.15	Results of the DESSA 2 Construct Validity Study (<i>N</i> = 60): Means and Standard Deviations of the DESSA 2 and the SSIS SEL Scales and Composite Scores
TABLE 3.16	Results of the DESSA 2 Construct Validity Study (<i>N</i> = 60): Correlation of the DESSA 2 with the SSIS SEL
TABLE 3.17	Sample Characteristics for the DESSA and DESSA 2 Comparability Study
TABLE 3.18	Results of the Comparability Study (<i>N</i> = 63): Means, Standard Deviations, and Correlations of the DESSA and DESSA 2 Scales and Composite Scores
TABLE 3.19	Means, <i>SDs</i> , <i>Ns</i> , and Correlations of the DESSA 2 with the Four DESSA 2 mini Forms
TABLE 3.20	Means, <i>SDs</i> , <i>Ns</i> , and Correlations of the DESSA 2 Raw Scores with the Four DESSA 2 mini <i>T</i> -scores Corrected for Item Overlap
TABLE 3.21	Rates of Agreement Between Each DESSA 2 mini and the Social-Emotional Composite <i>T</i> -score From the DESSA 2

CHAPTER 4

FIGURE 4.1 DESSA 2 Educator Form Presented in the Online System

CHAPTER 5

TABLE 5.1 Descriptive Categories and Interpretations of the DESSA 2 *T*-scores

TABLE 5.2 Individual Item Analysis Values for the DESSA 2

TABLE 5.3 Interpretation and Guidance for Progress Monitoring

FIGURE 5.1 Relationship of DESSA 2 *T*-scores, Percentile Ranks, and the Normal Curve

FIGURE 5.2 A Sample DESSA 2 Individual Student Rating Report

FIGURE 5.3 Item Level Identification as Shown on the Individual Student Rating Report

FIGURE 5.4 Sample Impact Report for DESSA 2 Data

FIGURE 5.5 Sample Student Movement Report for DESSA 2 Data



ACKNOWLEDGMENTS

The Devereux Student Strengths Assessment, Second Edition (DESSA 2) has evolved over the 15 years since the publication of the first edition of the DESSA. A great many people have contributed to this effort, and we would like to thank them for sharing their knowledge, talents, and expertise with us.

First, we are deeply thankful to the educators, administrators, and staff who have shared their feedback, questions, and concerns about the DESSA over the years. Their insights and experiences have enriched the DESSA 2, the associated strategies and interventions, and the delivery of these tools in our online system. We would also like to thank the thousands of educators who contributed the ratings that were used in the development of the DESSA 2. Without their participation in the national standardization and psychometric studies, the development of the DESSA 2 would not have been possible. A list of participating data collection sites is provided in Appendix C.

We want to thank Vivek Kartha, CEO of Riverside Insights, and Riverside's senior leadership team for sharing our vision of data-driven social and emotional skill development throughout the development of the DESSA 2. Their interest, guidance, and commitment to our mission are essential to all that we do. We would also like to express our sincere thanks to Anand Iyer, General Manager, Aperture Education, for his unwavering support, advocacy, and guidance throughout the development process.

Next, we express our appreciation for our colleagues at Aperture Education who are our partners in the measurement and promotion of students' social and emotional competence and resilience. Stacey Preator, Vice President of Customer Experience, and Kelly Hays, Vice President of Sales and Marketing, shared their knowledge and guidance throughout the development process and the rollout of the DESSA 2. Along with their staff, they encouraged and support to our recruitment efforts. It took the expertise of many individuals to prepare for the release of the DESSA 2. We would especially like to thank Jim Sandherr, Molly Parke, Bri Spoentgen, Alyssa Moore, and Kolter Atkinson for sharing their cross-functional knowledge to ensure the successful launch of this product.

Educators and administrators would not have access to the DESSA 2 without the efforts of Aperture Education’s digital products team. We appreciate Jessica Adamson, Vice President of Product and Partnerships, JD Duhadway, Vice President of Engineering, and their teams of product owners, developers, engineers, user experience designers, and others for making the online system easy to use and informative. Similarly, we are indebted to Aperture Education’s customer experience team and professional learning team, who help educators and staff become experts in using the DESSA 2 to improve student outcomes. We believe assessment data only has value if it leads to improved outcomes for students, and as such, we are especially grateful to Sara Prewett, Senior Researcher, and her team for developing the DESSA strategies and intervention programs that educators and staff can use to improve students’ social and emotional competencies.

One member of Aperture Education’s research and development team deserves our special thanks. Emily Parker, Research Associate, was responsible for the collection of our national standardization sample, a daunting task that she undertook with enthusiasm. She also collected data for several psychometric studies and spent countless hours running analyses. Emily helped with literature reviews, prepared the manual, and provided valuable insights throughout development during our team meetings. Without Emily’s dedication, the DESSA 2 would not have been possible, and for that we are truly indebted.

Our sincere thanks to all of you who contributed to the development of the DESSA 2.

— Jennifer, Evelyn, Paul, Jack, and Valerie



ABOUT THE AUTHORS

Jennifer L. Robitaille, M.S. is the Director of Research and Development for Aperture Education. She earned her B.A. in psychology from Bloomsburg University of Pennsylvania and her master's degree in research psychology from Villanova University. Jennifer has focused her work on strength-based approaches to the measurement and promotion of social and emotional competence and resilience in children, youths, and the adults who care for them. She has provided consultation, professional development, and evaluation services to early care and education, school, and out-of-school-time settings using the strength-based assessment suite, including the *Devereux Early Childhood Assessment (DECA) for Infants, Toddlers, and Preschoolers* and the *Devereux Student Strengths Assessment (DESSA) for grades K–12*. Jennifer is an author of the *DESSA assessment suite and the Educator Social-Emotional Reflection and Training (EdSERT)* program, a professional development program focused on enhancing educators' knowledge and effective use of social and emotional teaching practices. Her primary research interests focus on the measurement of social and emotional competence and resilience in children and youths, the promotion of educator well-being and social and emotional teaching practices, and program implementation.

Evelyn S. Johnson, Ed.D. is the Vice President, Research and Development for Riverside Insights and Professor Emeritus at Boise State University. She earned her Doctor of Education degree from the University of Washington, Seattle in Special Education and Measurement in 1999. She worked as a Research Scientist for the Center for Disease Control and as a Research Associate for the National Research Center on Learning Disabilities at the University of Kansas before joining the faculty at Boise State University, where she also served as the CEO of Lee Pesky Learning Center. During her tenure at Boise State, Dr. Johnson's research focused on: (1) screening and assessment to improve educational outcomes for students with learning disabilities, (2) improving instructional practice through teacher observation and feedback, and (3) self-regulated learning approaches to improve outcomes for students with disabilities. She served as the principal investigator for a federally funded project RESET, which developed reliable and valid teacher observation measures of evidence-based practices for students with

disabilities. In 2016, she received a Fulbright Scholars Award to Qatar, where she worked with a team of researchers in the Middle East to develop the *Test of Early Arabic Literacy Skills (TEALS)*. Dr. Johnson's research has consistently focused on the role of assessments in improving instructional decisions and learning outcomes.

Paul LeBuffe, M.A. is a senior research consultant for Aperture Education. He is a graduate of St. Mary's College of Maryland and received his master's degree in experimental psychology from Bryn Mawr College. For the past 30 years, Paul's career has focused on strength-based approaches to promoting social and emotional competence and resilience in children, youths, and the adults who care for them. Believing that such approaches should be data-driven, Paul has authored many widely used, strength-based assessments of behaviors related to children and adolescents' social and emotional strengths and needs, including the *Devereux Early Childhood Assessment for Infants (DECA-I)*, *Toddlers (DECA-T)*, *Preschoolers (DECA-P2)*, and most recently, the *Devereux Student Strengths Assessment (DESSA and DESSA 2) for grades K–12*. These assessments have been adopted by school districts, out-of-school-time programs, Head Start, and other early care and education programs both in the United States and internationally and the results have been used by professionals and parents to promote their students' social and emotional competence, foster their resilience, and build the skills they will need for school and life success. In addition to the assessments mentioned above, Paul has authored numerous research articles and chapters on social and emotional competence assessment, resilience, and related topics. Paul has presented to varied audiences, including psychologists, teachers, school administrators, out-of-school-time professionals, and parents throughout the United States and internationally on promoting resilience in children, youths, and adults.

Jack A. Naglieri, Ph.D. is a research professor at the University of Virginia and senior research scientist consultant for Aperture Education. He started his career as a school psychologist in 1975, earned a Ph.D. in 1979, and held university positions at Northern Arizona University, The Ohio State University, and George Mason University. Throughout these years, he focused on applied educational and psychological research and the development of psychological and educational measurement tools. He has published 23 books, 293 scholarly papers, and 51 tests and rating scales. Dr. Naglieri is the author of *Naglieri Nonverbal Ability Test* and the *Naglieri General Ability Tests: Verbal, Nonverbal and Quantitative*. He is also well known for his PASS neurocognitive theory of intelligence as measured by the *Cognitive Assessment System*, 2nd Edition, and the intervention handouts book entitled *Helping Children Learn*, 2nd Edition. He also authored *The Autism Spectrum Rating Scales*, *Comprehensive Executive Function Inventory: Child* and *Comprehensive Executive Function Inventory: Adult*, *The Devereux Student Strengths Assessment (DESSA)*, and the *DESSA-mini*. Dr. Naglieri has consistently emphasized the role tests play in accurate diagnosis, educational interventions, and especially equitable assessment based on sound theory and test construction.

Valerie B. Shapiro, Ph.D.* is an associate professor jointly appointed in Social Welfare and Public Health at the University of California, Berkeley. Dr. Shapiro was recently selected as a William T. Grant Foundation Scholar. Dr. Shapiro’s research is in the prevention of mental, emotional, and behavioral problems in children and youths through the adoption, implementation, and sustainability of effective prevention practices. To promote the use of evidence in prevention practice, her scholarship focuses on how to (1) set the stage for communities to adopt and sustain a science-based approach to prevention, (2) implement programs successfully, (3) assess youth outcomes in routine practice, and (4) use big data for continuous improvement in the implementation of systematic social and emotional learning. She serves as the chair of the Coalition for the Promotion of Behavioral Health, which produced the National Academy of Medicine Discussion Paper entitled “Unleashing the Power of Prevention” as part of the Grand Challenges for Social Work and Society Initiative and sits on the board of the National Prevention Science Coalition. Dr. Shapiro was the sole recipient of the 2014 Prytanean Faculty Prize, became a Fellow of the Society of Social Work & Research in 2017, and has co-authored several practice tools, such as the DESSA, used to promote the well-being of children nationally. She serves on the California Department of Education State SEL Advisory Group and recently co-authored UNESCO guidance for the assessment, monitoring, and evaluation of social and emotional learning worldwide. 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. She is a licensed social worker and Department of Education certified school social worker.

*This work is derivative of the Devereux Student Strengths Assessment (DESSA) and DESSA-mini, to which Valerie Shapiro made authorial contributions. Valerie Shapiro is not responsible for the final form of this derivative work.



FOREWORD

As part of Riverside Insights, Aperture Education offers a continuum of strength-based behavior rating scales and aligned instructional strategies for the assessment and promotion of students' social and emotional competencies from kindergarten through 12th grade. Collectively these tools, referred to as the DESSA System, reflect our commitment to *data-driven social and emotional skill development*, which has three key elements.

First, like academic achievement, the social and emotional competence of each student should be assessed – and when indicated, differentiated – and individualized social and emotional instruction should be provided. Although contextual factors including school culture and climate play an important role in facilitating or inhibiting both the acquisition and demonstration of social and emotional competencies, individual assessment is critically important. Only by assessing and addressing each individual student's social and emotional competencies, reinforcing their existing strengths, and remediating any skill deficits can we ensure that each student has the skills that they need to be successful in school and in life. Given that educational equity has been defined as “mean(ing) that **every student** has access to the resources and educational rigor they need” (Jagers et al., 2018, emphasis added) and is “achieved when **all students** receive the resources they need so they graduate prepared for success” (Center for Public Education, 2016, emphasis added), the assessment of social and emotional competencies accompanied by differentiated instruction is essential to promoting educational equity.

A second key element of *data-driven social and emotional skill development* is supporting educators in understanding DESSA data. The reporting features of the online system that delivers the DESSA encourage the aggregation of DESSA data at various levels (e.g., classroom, grade, site, program/district) and the disaggregation of data by important student and program characteristics. These powerful data analytic tools enable educators to generate and explore hypotheses about program impact on diverse groups of students, deepening understanding and further supporting effective practice and educational equity efforts.

The third core element of *data-driven social and emotional skill development* is the use of assessment data in both formative (student progress) and summative (program efficacy)

evaluations to continuously improve practice and optimize outcomes. The DESSA assessments provide advanced interpretation techniques to support these important activities.

Since the publication of the first edition of the DESSA for grades K–8 in 2009, the science of social and emotional learning has expanded dramatically, as have educational policy and public interest in this area. Our thinking about social and emotional skill development has continued to evolve as we reflect on what we have learned from our research, the findings of other researchers, and, most importantly, the experience of thousands of educators who have used the DESSA. We hope that the publication of the DESSA 2 will support and extend current efforts by communities to recognize the importance of social and emotional competence in ensuring the well-being, resilience, and success of all students. The authors and staff of Aperture Education and Riverside Insights welcome opportunities to collaborate with students, educators, parents, and organizations that share this goal. We can be reached through the Aperture Education website, www.ApertureEd.com.



Chapter 1

INTRODUCTION

CHAPTER 1

Introduction



The Devereux Student Strengths Assessment, Second Edition (DESSA 2) is a behavior rating scale completed by educators and assesses social and emotional competencies in students grades K–8. It is accompanied by the DESSA 2 mini, a screener and progress monitoring tool of overall social and emotional competence. We collectively refer to these two assessment tools as the DESSA 2 (or DESSA 2 measures) throughout this manual for simplicity, unless otherwise noted. The DESSA 2 is part of the DESSA System, a comprehensive suite of assessment tools and instructional strategies that enable education professionals to measure and strengthen the social and emotional competencies of students K–12.

The DESSA 2 is a revision of the Devereux Student Strengths Assessment (DESSA) and DESSA-mini, which were published in 2009 and 2011, respectively (LeBuffé et al., 2009/2014; Naglieri et al., 2011). The first edition of the DESSA was developed to meet the burgeoning need for a practical, norm-referenced measure of social and emotional competence in school and out-of-school time (OST) settings. Upon publication, the DESSA received favorable reviews by experts in the field (Atlas, 2010; Denham et al., 2010; Haggerty et al., 2011; Malcomb, 2010; Merrell & Gueldner, 2010; Tsang et al., 2012). The DESSA and DESSA-mini have demonstrated strong psychometric qualities as evidenced in the technical manuals and numerous studies, which are summarized elsewhere (Hwang et al., 2023; LeBuffé et al., 2018; Naglieri et al., 2013; Shapiro et al., 2017). Studies have shown that children who receive typical or high scores on the DESSA are less likely to have behavior problems (Shapiro & LeBuffé, 2006; Shapiro et al., 2017) and more likely to have academic success (Chain et al., 2017). Recent studies have provided evidence of measurement invariance of the DESSA-mini over time and across subgroups, suggesting that the DESSA-mini measures the same construct of social and emotional competence (a) within and across academic years, raters, and grade levels (Lee et al., 2023b) and (b) across diverse subgroups of students based on gender (female and male), race and ethnicity (Asian/Asian American, Black/African American, Hispanic/Latinx, and White), family income level (eligible for free or reduced price lunch and not eligible), disability status (receiving special education and not receiving), and language (English language learner [ELL] and non-ELL) (Lee et al., 2023a; Lee et al., 2024).

Since publication, the DESSA and DESSA-mini have been widely adopted to assess the social and emotional competence of over half a million children annually. The tools have been used with a variety of students reflecting diversity across racial/ethnic groups, socio-economic status, ability, and risk contexts, and have considerable use in research (for a review, see Hwang et al., 2023). They have also been used as part of implementing or evaluating a variety of social and emotional, resilience-promoting, and youth development interventions. Collectively, there is a strong empirical and practical foundation for the use of the DESSA and DESSA-mini to screen, assess, monitor progress, and evaluate growth in social and emotional competence over time and among diverse populations.

Purpose of the Revision

Despite the success of the first editions of the DESSA and DESSA-mini, it is best practice to periodically review and revise assessment tools when necessary (AERA, 2014). We embarked on this revision for several reasons. First, our collective understanding of the skills needed for success in today's world has evolved since the publication of the DESSA in 2009. The CASEL framework has undergone revisions, and the field has seen rapid growth in research, curricula, and social and emotional learning standards. As a result, a primary goal of the revision was to update the item content and scale structure to align with these changes. Second, the United States student population has changed since the national norms were originally collected during the 2005–2006 school year. In addition to student demographics changing, it has been well documented that mental health concerns for children and youth have been steadily rising over the last decade (Centers for Disease Control, 2021; LeBrun-Harris et al., 2022). The COVID-19 pandemic ushered in a new set of challenges for youth, and recent data suggests it may have exacerbated mental health concerns (Centers for Disease Control, 2021). As a result, we aimed to update the normative sample to ensure scores would be representative of K–8 school-age students today. And finally, the educators and staff across the United States who have used the DESSA in practice over the past 15 years shared feedback with us about specific items that were confusing or difficult behaviors to rate. They also helped us understand the need to simplify and reduce the time to administer the assessment and interpret its results, given the realities and time constraints present in classrooms today. As such, we aimed to address areas of improvement they have identified while still retaining what works well.

The DESSA 2 retains the strengths of the DESSA while offering the following improvements:

- Updated national norms
- Updated item content
- Shorter at 40 items
- Simplified item wording and removal of items that caused confusion or were hard to rate
- Reduced number of scales to simplify interpretation and planning
- Enhanced instructions to address common rater questions
- Updated response options for consistency across the DESSA assessment suite

Background

Social and emotional skill development is not only an integral part of education and human development but is broadly considered a path to personal well-being and global citizenship (Chatterjee Singh & Duraiappah, 2020). Decades of research have demonstrated that social and emotional and other positive youth development initiatives in schools and OST programs can (1) improve students' social and emotional skills and relationships, perceptions of school climate, and academic performance and (2) reduce student anxiety and undesirable behavior (Durlak et al., 2022; Mahoney et al., 2018). In addition, these initiatives can contribute to continuous improvement in education and youth development systems, when implemented well and systemically, with a favorable cost-benefit ratio (i.e., they can save more than they cost; Payton et al., 2008; Gullotta, 2015; Belfield et al., 2015).

To identify and support the social and emotional skill development of all students, a strength-based approach to assessment is needed that can assess student learning and provide actionable information to continuously improve initiatives. Information about an individual student's social and emotional competencies has the potential to inform instruction in ways that give each young person what they need to thrive, prevent problems before they occur, and invite multiple stakeholders into collaborative conversations. Aggregating information about student social and emotional competencies to the classroom, site, program, or district level can help inform local decision making and planning in ways that lead to greater coherence and thoughtful resource allocation. This also opens useful feedback loops for understanding the extent to which all young people are achieving social and emotional skill goals. The DESSA 2 is an assessment tool that provides these essential functions in the implementation of social and emotional and other youth development initiatives for students in grades K–8.

The DESSA tools have origins in the strand of applied developmental psychology known as *resilience theory*, which explores how individuals attain “good outcomes in spite of serious threats to adaptation or development” (Masten, 2001, p. 228). *Resilience* has been conceptualized as a dynamic process (i.e., rather than a stable trait) that represents the capacity to “adapt successfully to significant challenges that threaten the function, viability, or development” of an individual (Masten, 2018, p. 16). Studies of resilient individuals have identified a consistent set of attributes and assets that contribute to resilient outcomes (Masten, 2014). These *protective factors* have been defined (Masten & Garmezy, 1985) as characteristics that moderate or buffer the negative effects of risk factors. Garmezy (1985) suggested that protective factors could be divided into three categories: (1) community systems, such as high-quality schools, (2) a supportive family, and (3) individual attributes (e.g., physical health, intelligence, problem-solving skills). Some protective factors – such as social and emotional skills, positive school climate, and supportive relationships – are also *promotive factors*, which lead to desirable outcomes regardless of risk level (Masten et al., 2021). The DESSA 2 is used to evaluate behaviors related to social and emotional competencies – a subset of malleable individual attributes that act as protective factors in the face of adversity and promotive factors for the positive well-being of all students. Since all young people can experience adverse events and stressors, building social and emotional competence can help to promote resilience and the healthy development of all youth (Shapiro, 2015). To be clear, the DESSA 2 is intended for use in systems in which adults both provide meaningful opportunities for young people to build social and emotional

competence, and simultaneously take responsibility for addressing and alleviating adversities that create an excessive or disparate need for resilience.

We use the term *social and emotional competence* to refer to an individual's ability to develop healthy identities, manage emotions, achieve personal and collective goals, feel and show empathy for others, establish and maintain supportive relationships, and make responsible and caring decisions (CASEL, 2020). We conceptualize a competence continuum ranging from a complete lack of proficiency to full proficiency in the execution of prosocial behavior. Our goal is to help identify and nurture the social and emotional strengths of youths, while simultaneously improving the relationships and environments that provide the contexts for their development (Shapiro, 2015). As consistent with CASEL's definition of social and emotional learning (<https://casel.org/fundamentals-of-sel/>), this involves empowering young people and adults to co-create thriving schools and contribute to safe, healthy, and just communities (Ozer et al., 2021). The DESSA 2 is intended to support whole-child education, the creation of trauma-informed schools, the growing emphasis of schools and OST providers to help promote equity and excellence (e.g., Jagers et al., 2018), and the related need for the assessment of social and emotional competence in routine educational practice.

The rapid growth of social and emotional learning research, curricula, and programs, accompanied by the adoption of learning standards for K–12 education by more than 20 states (CASEL, 2021), creates an ongoing need for an aligned assessment system. Some school districts seek an assessment system as a means of determining whether all students have met standards or otherwise acquired the prerequisite “noncognitive” skills for school and life success. Some districts and OST programs desire a formative assessment that can identify each student's social and emotional strengths and needs, inform instruction and programming, and gauge progress over time (Shapiro et al., 2017; Shapiro et al., 2024). Finally, schools and OST programs that have invested heavily in developing and/or implementing social and emotional, resilience-building, or youth development programs have a need for summative assessment to evaluate and continuously improve impact. The DESSA 2 mini and DESSA 2 were developed in response to these various needs.

Description of the DESSA 2 Measures

The DESSA 2 mini and DESSA 2 are standardized, norm-referenced behavior rating scales used to assess the social and emotional competence of students in grades K–8. We chose this method for several reasons. First, behavior rating scales are the most prevalent method used to assess behavior in schools (Elliott et al., 2015); they are well suited to evaluate the frequency of behaviors across several areas; and they can be “cheap, quick, reliable, and in many cases, remarkably predictive of objectively measured outcomes” (Duckworth & Yeager, 2015, p. 239). The DESSA 2 mini and DESSA 2 can be completed by educators or staff at schools and youth-serving agencies, including OST, social service, and mental health programs. The assessment tools are entirely strength-based, meaning that the items query positive behaviors (e.g., contribute to group or team goals) rather than maladaptive ones (e.g., annoy others). 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 four weeks.

DESSA 2

The DESSA 2 is comprised of 40 items organized into conceptually derived scales that provide information about six CASEL-aligned social and emotional competencies. Standard scores can be used to calibrate each student's competence in each of the six dimensions and guide school and program-wide, class-wide, and individual strategies to promote those competencies. The scale names, scale definitions, and sample items are as follows:

- ***Optimistic Thinking*** (7 items): Optimistic Thinking is the belief and demonstration of confidence, hopefulness, and positive thinking regarding oneself, others, and one's life situations in the past, present, and future.
 - accept that making mistakes is part of learning?
 - speak positively about their future potential?
 - believe they can achieve their goals?
- ***Self-Awareness*** (7 items): Self-Awareness is the ability to understand emotions, thoughts, and values and how they influence one's behavior; recognize strengths and limitations; and develop healthy identities and a sense of purpose.
 - describe the emotion they were feeling?
 - show an awareness of their personal strengths?
 - demonstrate a sense of who they are and what is important to them?
- ***Self-Management*** (7 items): Self-Management is the ability to manage emotions and behaviors across different situations and environments and to demonstrate agency as one works to set and achieve personal and collective goals.
 - stay calm when faced with a challenge?
 - think before they acted?
 - persist to achieve a goal?
- ***Social Awareness*** (6 items): Social Awareness is the understanding of social norms for behavior; the ability to empathize with, respect, and take the perspectives of others; and the feeling of connection and belonging with family, peers, schools, and community groups.
 - respond to others' feelings in kind and safe ways?
 - contribute to creating a positive learning environment?
 - respect a person's right to have a different perspective?
- ***Relationship Skills*** (7 items): Relationship Skills are the abilities to establish and maintain healthy and positive relationships, including effective communication, collaborative problem solving, negotiating conflict, and demonstrating helpful and supportive behaviors.
 - cooperate with others to solve a problem?
 - listen to others?
 - resolve conflicts positively?
- ***Responsible Decision Making*** (6 items): Responsible Decision Making is the ability to make careful, reliable, and constructive choices about personal and social behavior that are appropriate across diverse situations; to consider the personal, social, and collective impact of one's actions; and to demonstrate curiosity and open-mindedness to learning.

- prepare for school, activities, or upcoming events?
- accept responsibility for their actions?
- demonstrate openness to new situations, experiences, and people?

Each of the six DESSA 2 scale scores is derived from the ratings of the items assigned to that scale. A Social-Emotional Composite (SEC) score is also included, which is based on a combination of the scores received on the six scales. This composite score provides an overall indication of the strength of the student's social and emotional competence. The separate scores on the six DESSA 2 scales are used to create individual student rating reports as well as classroom and group reports, to convey the strengths and needs of the student and/or groups of students as compared to national norms (see Chapter 2 for a further explanation of the importance of norms). The DESSA 2 yields information that can also be used to compare ratings across raters and/or environments and across time to monitor progress and evaluate outcomes. More information about these interpretation strategies will be presented in Chapter 5.

DESSA 2 mini

The DESSA 2 mini is comprised of four 8-item parallel forms that are designed to be used on a universal (i.e., school- or program-wide) basis to determine the need for social and emotional interventions. Each of the four DESSA 2 mini forms contains a sampling of the six scales found on the DESSA 2 (see Chapter 2 for more details on the construction of the DESSA 2 mini). The DESSA 2 mini yields a single score, the Social-Emotional Total (SET) score, which provides an indication of the strength of the student's social and emotional competence based on a comparison to national norms and can be used to compare ratings between educator raters across time to monitor progress toward improving social and emotional competence.

Uses of the DESSA 2 Measures

Uses of the DESSA 2

The DESSA 2 has been developed to provide a measure of social and emotional competence that can be used to guide the implementation of strategies to promote positive youth development. Specifically, the DESSA 2 has been designed to:

- Provide a psychometrically sound, strength-based measure of social and emotional competence in students.
- Prioritize areas for social and emotional growth for individual students and groups of students.
- Inform the selection or design of strategies and interventions within multi-tiered systems of support (universal, targeted, indicated) to promote social and emotional competence for all students.
- Ensure that all students have the social and emotional competence they need to succeed and are meeting social and emotional learning standards set by the state, district, or organization.

- Identify students with the greatest need for social and emotional instruction, prevent problems before they emerge, and promote positive developmental outcomes.
- Provide meaningful information on student strengths for inclusion in individual education and service plans, as required by federal, state, and funder regulations.
- Identify social and emotional disparities between socio-demographic groups that can be subjected to a root-cause analysis and addressed to promote equitable outcomes for all students.
- Facilitate collaboration between students, parents, and education professionals by providing a means of comparing ratings of the same students using the DESSA 2 and the DESSA Middle School Edition Student Self-Report (DESSA-MSE SSR) to identify similarities and meaningful differences.
- Facilitate progress monitoring for individuals and groups of students by evaluating change over time at the individual scale level.
- Enable the evaluation and continuous improvement of social and emotional and positive youth development programs by rigorously evaluating outcomes at the individual student, classroom/group, school, district/program, and community levels.
- Serve as a sound research tool to advance science and support public policy development regarding the promotion of social and emotional competence.
- Provide a common language and approach to those involved in promoting positive youth development, including educators, administrators, policymakers, community members, mental health and social service professionals, parents, and students.
- Serve as a tool for conducting a baseline assessment to document the need for social and emotional and youth development programming.

Uses of the DESSA 2 mini

The DESSA 2 mini has been developed to provide a measure of social and emotional competence, which can be used to support social and emotional, positive youth development, and mental health promotion initiatives. More specifically, the DESSA 2 mini was designed to:

- Provide a psychometrically sound, brief, strength-based measure of social and emotional competence in students for screening purposes.
- Ensure that all students have the social and emotional competence they need to succeed and are meeting social and emotional learning standards set by the state, district, or organization.
- Identify students with the greatest need for social and emotional instruction, prevent problems before they emerge, and promote positive developmental outcomes.
- Identify social and emotional disparities between socio-demographic groups that can be subjected to a root-cause analysis and addressed to promote equitable outcomes for all students.
- Help determine the appropriate strategy and intervention approach for strengthening the social and emotional competence of each student.
- Serve as a tool for conducting a baseline assessment to document the need for social and emotional and youth development programming.

- Provide a method for monitoring the progress of individual students or groups of students receiving social and emotional instruction and differentiated interventions.
- Provide a method for evaluating outcomes at the individual student, classroom/group, school, district/program, and community levels.
- Serve as a sound research tool to advance science and support public policy development regarding the promotion of social and emotional competence.

Values Guiding the Development and Use of the DESSA 2 Measures

The overarching goal of the DESSA 2 mini and DESSA 2 are to inform the promotion of social and emotional competence and resilience of students. Five characteristics shape our approach to achieving this goal. First, the measures are strength-based. This orientation is important to the dual goals of mental health promotion and challenging behavior prevention in that it enables the proactive identification of strengths and weaknesses in social and emotional development before the occurrence of significant social and emotional challenges emerge (LeBuffe et al., 2021; LeBuffe & Shapiro, 2004). If practitioners wait until undesirable behaviors emerge before offering social and emotional instruction, they may have missed the opportunity to prevent the enormous costs of mental, emotional, and behavioral problems, and their remediation, to students, their families, schools, and society (O’Connell et al., 2009). Strength-based approaches also clearly list positive skills that students can work to develop as needed to achieve their own personal goals.

The second key characteristic of the DESSA 2 is to be justice-promoting. In this commitment, we intend to affirm the diversity of students by including their voices in decision making through the accompanying assessment tool – the DESSA-MSE SSR for middle school-aged students – and contribute to equity for all. To fulfill this commitment, the DESSA 2 was standardized on a sample of students who reflect the regional, gender, and racial/ethnic diversity of the United States. Analyses were conducted prior to publication to examine how the tool detects and/or presents differences between sociodemographic subgroups, which are transparently reported in Chapter 2. Our strength-based approach, described in this chapter, aims to prevent the stigmatization and pathologization of students as a result of the assessment process. Similarly, our preventative orientation advances the call for a reorganization of community resources to promote population health rather than waiting for a meaningful subsection of students to experience hardship and rationing cost-intensive interventions. Furthermore, Chapters 4 and 5 describe our approach to scoring and interpretation, which centers on educational institutions taking responsibility for social and emotional instruction and building students’ capacity to develop their social and emotional competency (e.g., providing high-quality, evidence-based social and emotional instruction), rather than presuming that low DESSA 2 scores are the fault or responsibility of the student themselves. Chapter 5 stresses the importance of including the voice of students and their families in the process of interpreting DESSA 2 scores, setting goals, making decisions, and setting the expectation that the DESSA 2 be used

alongside the DESSA-MSE SSR and in conjunction with climate surveys and other approaches to risk assessment, such that basic needs and threats to developmental outcomes are not missed and the promise of structural and environmental strategies are not overlooked.

The third defining characteristic is the use of an assessment process that merges all we know about a student with norm-referenced data to help understand the individual, and ultimately guide intervention decisions. In common with the positions of other professional organizations, we believe that measures of social and emotional competence have maximum value when they lead to improved outcomes for young people (National Association for the Education of Young Children, 1987). As a result, the DESSA 2 was designed to yield *actionable* insights to inform the selection and implementation of evidence-based social and emotional instruction programs or strategies intended to be integrated into routine practice in schools, OST programs, and at home.

The fourth foundational characteristic of the DESSA 2 is a commitment to strong psychometric qualities. The assessment tool meets or exceeds the standards promulgated by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (AERA, 2014), including large, diverse standardization samples that approximate the population of school-age youths with respect to important demographic characteristics, good to excellent reliability, and sufficient validity data to support the intended uses of the scales. These are important attributes for defensible decision making with and on behalf of young people. Detailed information on the psychometric characteristics of the DESSA 2 measures is provided in Chapter 3.

The fifth foundational characteristic of the DESSA 2 is the focus on educators, including teachers and OST providers, as not only the raters (i.e., the people providing the ratings) but also as the users of that information (i.e., the people who interpret the assessment results and use them to inform instruction). This focus on empowering educators to be the consumers of test results was originally in response to a resource deficit; the lack of mental health consultants in public schools and OST environments (e.g., NASP, 2011). The strength-based orientation of the DESSA 2 makes its use by individuals who are not mental health professionals appropriate in that the scales do not yield scores with pejorative labels (e.g., “extreme risk”) or diagnoses (e.g., anxious/depressed). Appropriate usage is encouraged through simple directions, on-demand training, and a best-practice model that positions the assessment as part of routine educational practice.

Qualifications of Users and Raters of the DESSA 2 Measures

Qualifications of Users

For the purposes of this manual, DESSA 2 mini and DESSA 2 *users* are those who not only administer the assessment tool 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 2. 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 2 user has responsibility for the proper use and interpretation of DESSA 2 results.

Because DESSA 2 results can be used to make decisions that shape the experiences of students, DESSA 2 users should have training in the proper administration, interpretation, and utilization of the DESSA 2. This should include knowledge of the interpretation of standardized scores such as *T*-scores and percentile ranks, the interpretation of scale content and profiles, and how to communicate the results to families, allied professionals, and young people themselves. Typically, DESSA 2 users will include educators, administrators, coaches, program directors, and evaluators. The DESSA 2 can also be used by counselors, social workers, psychologists, and other professionals in education, behavioral health, child welfare, and juvenile justice settings to gain a better understanding of a student's social and emotional strengths and needs.

Qualifications of Raters

A *rater* is any person who completes the items on the DESSA 2 measures. There are two main qualifications of a rater: first, the rater must have had sufficient exposure to the student over the four weeks prior to completing the DESSA 2; and second, raters should also be able to read English at the sixth-grade level. (Recommendations for using the DESSA 2 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 student could yield an erroneously low rating. In general, we recommend that raters should have contact with the student for two or more hours for at least three days per week for a four-week period.

Raters of the DESSA 2 measures include educators. This group typically includes teachers, teacher aides, assistant teachers, instructional assistants, OST program staff, and other professionals who interact directly with the student on a regular basis.

Reasonable concerns exist as to whether a rater's biases may shape a student's assessment scores. Rater bias may artificially inflate or suppress assessment scores relative to the actual frequency of behavior. A large amount of rater bias is problematic because scores could be less precise than are desired for educational decision making. Studies have shown that rater-specific bias on the first edition of the DESSA may be less than expected (Shapiro et al., 2016), perhaps due to the strength-based nature of the items (see Chapter 2), and is further reduced when educators are prepared for implementation through training. Features intended to mitigate bias are described in depth by Mahoney and colleagues (2022).

Restrictions for Use

DESSA 2 users should follow both the instructions included in this manual and all commonly accepted guidelines for test use and interpretation, such as the Standards for Educational and Psychological Testing (AERA, 2014). It is the DESSA 2 user's responsibility to ensure that completed DESSA 2 ratings and reports remain secure and are released with consent only to professionals who will safeguard their proper use. Copyright law does not permit the DESSA

2 user to photocopy or otherwise duplicate test items or record forms in any form, even for the purpose of sharing results. The completed DESSA 2 Individual Student Rating Report may be copied and provided to students, parents, and multi-disciplinary teams after it has been reviewed with them. Because all DESSA 2 items, norms, and other materials are copyrighted, no DESSA 2 materials may be reproduced or transmitted in any form or by any means without written permission from Riverside Insights.



Chapter 2

DEVELOPMENT AND STANDARDIZATION

CHAPTER 2

Development and Standardization



This chapter presents the development and standardization of the DESSA 2 and the four forms of the DESSA 2 mini, which were developed simultaneously using the same item pool and standardization sample.

Development of the DESSA 2 Items

Item development for the Devereux Student Strengths Assessment, Second Edition (DESSA 2) began with a review of the eight social and emotional constructs measured by the original Devereux Student Strengths Assessment (DESSA; LeBuffe et al., 2009/2014). As part of this review, we consulted the literature and research on social and emotional competence, which has grown extensively since the publication of the DESSA in 2009. This construct review process resulted in updated definitions of the social and emotional competencies we aimed to measure with the DESSA 2.

Second, we reviewed the existing 72 items on the DESSA. These items were originally developed through a thorough review of the literature on resilience (e.g., Werner & Smith, 1982, 1992), social and emotional learning (e.g., Payton et al., 2000), and positive youth development (e.g., Catalano et al., 2002). Feedback on the items was solicited from DESSA users and staff at Aperture Education who regularly support DESSA users. We reviewed input and suggestions for item revisions collected from DESSA users over the past decade. We also examined data from DESSA ratings completed in the online DESSA system over the past several years to identify items that are most often skipped by educator raters. For example, the item “pass up something they wanted, or do something they did not like, to get something better in the future” was found to be confusing to educator raters and skipped much more frequently compared to the other items. Based on these methods, several items from the original DESSA were removed from consideration for inclusion in the revised item pool.

Educator raters also reported confusion completing items that included references to both school and home settings (e.g., “serve an important role at home or school”, “do routine tasks or chores without being reminded”). Some original DESSA items were intentionally written this

way to meet the need for only one printed record form applicable to both educators and parents (prior to the availability of an online system) and to aid social and emotional instructional planning across home and school environments. Given the availability of an online system for administering and reporting DESSA 2 ratings, we decided to revise or write new items in a manner that allowed for the specification of the environment where necessary. For example, we revised the original item “do routine tasks or chores without being reminded” to “do tasks or homework without being reminded” for educator raters and “do chores or homework without being reminded” for parent/guardian raters. Although we developed the item pool for both educator and parent/guardian raters simultaneously, as of the date of publication, the DESSA 2 is an educator-completed assessment tool. The addition of a parent/guardian form is planned.

Then, we reviewed, and in some cases revised, the remaining original DESSA items to make them clearer or to reduce the reading level. For example, the item “get things done in a timely fashion” was revised to “get things done on time.” Items were considered for developmental appropriateness for elementary and middle school students, as well as from an equity perspective to help ensure appropriateness across diverse groups of students. As a result, additional items were deleted (e.g., “wait for their turn”) or revised (e.g., the item “show the ability to decide between right and wrong” was reworded to “do the right thing in a difficult situation”).

Next, we compared the content of the remaining items to our updated scale definitions. We conducted literature reviews and new items were written to cover content areas not previously emphasized on the original DESSA (e.g., developing healthy identities) or content areas where we wanted to expand coverage or test new items. Finally, we ensured all the items were written to measure observable behaviors that would require little or no inference on the part of the observer. We also carefully considered the reading level of the items so that the overall readability of the DESSA 2 would be as low as possible, remaining at a sixth-grade reading level. This item-development phase resulted in a pool of 78 items, 34 of which were retained from the original DESSA.

Feedback from Expert Reviewers

Prior to standardization, the construct definitions and 78 items were reviewed by four experts in the fields of education and social and emotional development, including individuals with expertise in the related fields of educational equity, special education, and school psychology and mental health. Items were conceptually organized by the social and emotional competency domain it intended to measure (i.e., Self-Management) and experts were asked to review each item and indicate whether the item was relevant to its aligned domain and whether the item was appropriate from an equity perspective. Reviewers also indicated whether the items were developmentally appropriate for elementary and middle school students and whether any important social and emotional skills were missing from the item pool.

Feedback obtained from the reviewers was very consistent. The reviewers positively evaluated the item pool from the perspectives of social and emotional skill content coverage, developmental appropriateness, and equity across diverse groups of students. Based on feedback, wording revisions were made to four items. For example, the item “show respect for others in a game or competition” was simplified to “show respect for others.” Reviewers also indicated

eight items that might be difficult for educators to rate (e.g., “consider the impact of decisions on others”). As a result, these items were eliminated. Lastly, reviewers also identified *collective goal setting and efficacy* as an important social and emotional content area missing from our scale definitions and item pool; three new items were written to measure this skill set. This process resulted in a set of 73 items that were incorporated into the standardization edition of the DESSA 2.

National Standardization

In accordance with standards promulgated by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (AERA, 2014), we normed the DESSA 2 through a carefully prescribed method to ensure the data collection procedures resulted in a large, diverse standardization sample that closely approximated the United States population of students in kindergarten through eighth grades with respect to important demographic characteristics. A discussion of the psychometric characteristics of the DESSA 2 measures is provided in Chapter 3.

We collected data using an online rating form. The sample was collected between September 2023 and February 2024. Ratings were obtained from educators of elementary and middle school students (kindergarten through eighth grades) from school districts and out-of-school-time (OST) programs across the United States. Schools and programs were recruited through a variety of methods including invitations to existing DESSA users, advertising through national organizations such as the National Association of School Psychologists (NASP) and the American Educational Research Association (AERA), and posting the study opportunity on websites, listservs, and social media. No personally identifying information was included in the standardization protocols, which were reviewed and approved by the Devereux Advanced Behavioral Health’s Institutional Review Board.

Selection of the DESSA 2 Standardization Sample

Ratings of elementary and middle school students in kindergarten through eighth grade completed by classroom teachers, teacher aides, or OST program staff were eligible for inclusion in the DESSA 2 standardization sample. Students receiving special education services were also eligible for inclusion. We eliminated ratings with too much missing data (defined as missing responses on two or more of the final set of 40 items) and ratings with the same item response across all items (e.g., the rater answered “Almost Always” for all items). Prior to finalizing, the sample was trimmed to achieve representativeness to U.S. Census data regarding grade, gender, race, Hispanic/Latinx ethnicity, geographic region of residence, socioeconomic status, and special education status.

Representativeness of the DESSA 2 Standardization Sample

A total of 1,550 students in grades K–8 (ages 5–14) comprised the DESSA 2 standardization sample. As with the original DESSA, classroom teachers, teacher aides, and OST and other program staff contributed ratings. The sample closely approximated the population of 5- through 14-year-olds in the United States with respect to grade, gender, geographic region of residence, race, Hispanic/Latinx identity, and socioeconomic status. We based the desired characteristics of the standardization sample on the most current national estimates (2018–2022) from the American Community Survey (ACS) published by the U.S. Census Bureau. In the tables that follow, the total number of students included may not sum to 1,550 due to missing data.

Grade and Gender

Table 2.1 presents the numbers and percentages of males and females in the DESSA 2 standardization sample in each grade from grades K–8, presented relative to the composition of the U.S. population. The number of students in each grade ranged from 133 in first grade to 241 in third grade. The overall mean number of students per grade was 172. 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 well as in each grade, closely approximate the proportions of the U.S. population.

TABLE 2.1
DESSA 2 Standardization Sample Characteristics by Grade and Gender

	Males		Females		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Kindergarten	89	49.7	90	50.3	179	11.5
1st Grade	65	48.9	68	51.1	133	8.6
2nd Grade	74	51.7	69	48.3	143	9.2
3rd Grade	117	48.5	124	51.5	241	15.5
4th Grade	106	52.2	97	47.8	203	13.1
5th Grade	102	50.7	99	49.3	201	13.0
6th Grade	73	48.7	77	51.3	150	9.7
7th Grade	77	51.3	73	48.7	150	9.7
8th Grade	84	56.0	66	44.0	150	9.7
Total Sample	787	50.8	763	49.2	1,550	100.0
U.S. %		51.2		48.8		

Note: The U.S. population data are based on the 2018–2022 estimates for 5- through 14-year-olds in “Table S0101: Age and Sex, 2022 American Community Survey 5-Year Estimates”, U.S. Census Bureau, 2022. Retrieved June 2024, from <https://data.census.gov/cedsci/>.

Geographic Region and Grade

We collected data from students attending 38 schools and OST programs across 19 U.S. 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 (Northeast, Midwest, South, and West) designated by the U.S. Census Bureau. These data show that the DESSA 2 standardization sample closely approximated the regional distribution of the U.S. population.

Race and Geographic Region

[Table 2.3](#) provides the DESSA 2 standardization sample composition by race within each geographic region. Based on information provided by educators on the rating forms, we classified students according to the six 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, White, Two or More Races, and Some Other Race. Data on race was available for 1,174 of the 1,550 students in the standardization sample. The remaining 376 ratings missing race included 65 responses for which the educator selected “Don’t Know” and 311 responses that were left blank. 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 2.2

DESSA 2 Standardization Sample Characteristics by Geographic Region and Grade

	Northeast		Midwest		South		West		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Kindergarten	14	7.8	43	24.0	44	24.6	78	43.6	179	11.5
1st Grade	20	15.0	13	9.8	22	16.5	78	58.6	133	8.6
2nd Grade	27	18.9	47	32.9	32	22.4	37	25.9	143	9.2
3rd Grade	36	14.9	32	13.3	62	25.7	111	46.1	241	15.5
4th Grade	45	22.2	34	16.7	86	42.4	38	18.7	203	13.1
5th Grade	30	14.9	13	6.5	84	41.8	74	36.8	201	13.0
6th Grade	36	24.0	32	21.3	65	43.3	17	11.3	150	9.7
7th Grade	61	40.7	14	9.3	67	44.7	8	5.3	150	9.7
8th Grade	45	30.0	19	12.7	71	47.3	15	10.0	150	9.7
Total Sample	314	20.3	247	15.9	533	34.4	456	29.4	1,550	100.0
U.S. %		15.9		21.0		39.0		24.1		

Note: The U.S. population data are based on the 2018–2022 estimates for 5- through 14-year-olds in “Table S0101: Age and Sex, 2021 American Community Survey 5-Year Estimates”, U.S. Census Bureau, 2022. Retrieved June 2024, from <https://data.census.gov/cedsci/>.

TABLE 2.3
DESSA 2 Standardization Sample Characteristics by Race and Geographic Region

	American Indian/Alaska Native		Asian		Black/African American		Native Hawaiian/ Other Pacific Islander		White		Two or More Races		Some Other Race		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Northeast	3	1.3	9	3.9	153	65.7	0	0.0	41	17.6	7	3.0	20	8.6	233	19.8
Midwest	34	16.0	2	0.9	12	5.7	0	0.0	156	73.6	6	2.8	2	0.9	212	18.1
South	0	0.0	21	4.8	126	28.8	0	0.0	277	63.2	10	2.3	4	0.9	438	37.3
West	2	0.7	20	6.9	20	6.9	3	1.0	218	74.9	24	8.2	4	1.4	291	24.8
Total Sample	39	3.3	52	4.4	311	26.5	3	0.3	692	58.9	47	4.0	30	2.6	1,174	100.0
U.S. %		1.0		5.0		13.5		0.2		57.7		13.5		9.0		

Note: The U.S. population data are based on the 2018–2022 estimates for 5- through 14-year-olds in “Tables B01001A, B, C, D, E, F, G: Sex by Age (Race), 2021 American Community Survey 5-Year Estimates”, U.S. Census Bureau, 2022. Retrieved June 2024, from <https://data.census.gov/cedsci/>.

Hispanic/Latinx Ethnicity and Geographic Region

The proportions of students of Hispanic/Latinx ethnicity included in the DESSA 2 standardization sample by geographic region are presented in Table 2.4. Educator raters were asked whether the student was of Hispanic/Latinx ethnicity. Data show that the Hispanic/Latinx composition of the standardization sample closely approximated that of the U.S. population and that Hispanic/Latinx students in the sample lived in all four regions of the United States.

Socioeconomic Status

To assess the socioeconomic status of the DESSA 2 standardization sample, we determined the number of students eligible to receive either free or reduced-price lunches. Based on the information provided by educators on the rating forms, eligibility data was available for 1,056 of the 1,550 students in the standardization sample. Of this sample of 1,056 students, educators reported that 607 (57.5%) were eligible to receive free or reduced-price lunches. This approximated the 53.3% of K–12 students in the U.S. eligible to receive free or reduced-price lunches in the 2022–2023 academic year (U.S. Department of Education, 2024).

Special Education Services

To assess the education status of the DESSA 2 standardization sample, we determined the number of students receiving special education services. Based on the information provided by educators on the rating forms, eligibility data were available for 1,507 of the 1,550 students in the standardization sample. Of this sample of 1,507 students, 303 (20.1%) were receiving special education services. This closely approximated the 15.2% of 3- to 21-year-old students in the United States being served under the Individuals with Disabilities Education Act (IDEA) during the 2022–2023 academic year (U.S. Department of Education, 2023).

TABLE 2.4

DESSA 2 Standardization Sample Characteristics by Hispanic/Latinx Ethnicity and Geographic Region

	Hispanic/Latinx		Not Hispanic/Latinx		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Northeast	93	29.6	221	70.4	314	21.1
Midwest	26	11.0	211	89.0	237	15.9
South	85	16.4	432	83.6	517	34.8
West	148	35.3	271	64.7	419	28.2
Total Sample	352	23.7	1,135	76.3	1,487	100.0
U.S. %		25.3		74.7		

Note: The U.S. population data are based on the 2018–2022 estimates for 5- through 14-year-olds in “Tables B01001I: Sex by Age (Hispanic or Latino), 2021 American Community Survey 5-Year Estimates”, U.S. Census Bureau, 2022. Retrieved June 2024, from <https://data.census.gov/cedsci/>.

Organization of the DESSA 2 Items into Scales

The primary purpose of the DESSA 2 is to provide educators, parents and guardians, and other professionals concerned with the social and emotional competence of students, as well as the students themselves, with a useful and meaningful set of scales that both (a) reflect current social and emotional functioning and (b) lead to strategies and interventions to promote social and emotional competencies. We have aligned our items across the DESSA K–12 assessment suite with the descriptions of core social and emotional competencies provided by the Collaborative for Academic, Social, and Emotional Learning (CASEL; www.casel.org). This framework is widely reflected in state and school district educational standards as well as social and emotional learning curricula, and it is, therefore, familiar to many educators and administrators.

On the original DESSA (LeBuffe et al., 2009/2014), three of the five core social and emotional competencies suggested by CASEL (Self-Awareness, Self-Management, and Responsible Decision Making) had been subdivided to yield more specific social and emotional competencies that were intended to simplify understanding and intervention (e.g., Responsible Decision Making was subdivided into the two scales of Personal Responsibility and Decision Making), and emphasize optimistic thinking as an important social and emotional competency (Ciarrochi et al., 2015). This process yielded eight preliminary first-order scales (Self-Awareness, Self-Management, Social Awareness, Relationship Skills, Personal Responsibility, Decision Making, Goal-Directed Behavior, and Optimistic Thinking).

As described at the start of this chapter, development of the DESSA 2 began with a construct and item review process that resulted in updated definitions of the social and emotional competencies we aimed to measure. In addition to updating our content coverage, we also refined our alignment to the CASEL Framework. Specifically, scale definitions were revised to yield six preliminary first-order scales. These include the five core competencies defined by CASEL (Self-Awareness, Self-Management, Social Awareness, Relationship Skills, and Responsible Decision Making) and the retaining of the Optimistic Thinking scale. This revision provides clearer alignment to the CASEL Framework, state social and emotional learning standards, and existing social and emotional learning curricula, while also reducing complexity for educators as they review and plan instruction based on DESSA 2 data.

We organized the DESSA 2 items into these logically derived and defined scales based, in part, on the CASEL Framework. Then, we used a series of statistical analyses to further refine and simplify the scales based on the following goals: (1) to identify the best scale solution, from both psychometric and interpretability perspectives; (2) to shorten the DESSA 2 as much as possible without compromising breadth of coverage; (3) to simplify the administration, scoring, and interpretation of the DESSA 2; and (4) to ensure that the constructs were measured reliably by the scales. Because we received consistent feedback from educator raters and administrators that the original DESSA containing 72 items was too time-intensive at about 5 to 8 minutes to complete, we aimed to shorten the DESSA 2 to approximately 40 to 50 items total, so long as the above goals were still able to be met.

To achieve the goals outlined above, we examined item performance using a combination of Classical Test Theory (CTT) and Item Response Theory (IRT) techniques. We dropped items based on reviewing the following criteria: First, we examined the corrected item-total

correlations to ensure that each item correlated highly with the scale to which it was assigned. Second, to avoid potential ceiling effects on any scale, which would impact the ability of the measure to detect a change, we examined each item’s mean raw score for evidence of potential ceiling effects (defined as an item mean score of 2.9 or higher; possible range of 0 to 4). To simplify the scales and avoid the necessity of age norms, we examined each item for evidence of age trends. These results are described in more detail later in this chapter. To minimize potential bias, we examined differential item functioning (DIF) between groups of students (Black or African American, White, and Hispanic/Latinx students). DIF analyses were conducted in flex-MIRT (Cai, 2022), which provides an overall test for DIF (total X2), a test of DIF with respect to the slope (X2a), and a test of the threshold parameters, assuming α is fixed across groups (X2c|a). An item was flagged for DIF with a significant overall Wald X2 test (adjusted for family-wise Type I error with Bonferroni correction). Using the “all-others-as-anchors” iterative approach (Woods, 2009), items not identified as exhibiting DIF were constrained as equal across groups to form a new latent trait ability metric. The DIF test was then repeated until no additional items were identified as exhibiting DIF. The magnitude of DIF was determined by the discrepancies between threshold parameters and slope, with extreme differences (>1) signifying a considerable effect. No items in the item pool indicated considerable effects; only negligible or small effects were observed. [Table 2.5](#) displays the DIF results for the final selected items.

TABLE 2.5
Evidence of DIF for Final Set of DESSA 2 Items

Scale/Item	Evidence of DIF Black vs. White Students	Evidence of DIF Hispanic/Latinx vs. White Students
Optimistic Thinking		
Item 3	<i>ns</i>	<i>ns</i>
Item 7	<i>ns</i>	<i>ns</i>
Item 15	<i>ns</i>	<i>ns</i>
Item 23	<i>ns</i>	<i>ns</i>
Item 27	<i>ns</i>	<i>ns</i>
Item 34	<i>ns</i>	<i>ns</i>
Item 40	<i>ns</i>	<i>ns</i>
Self-Awareness		
Item 1	<i>ns</i>	<i>ns</i>
Item 6	<i>ns</i>	<i>ns</i>
Item 8. ask questions when they did not understand something?	Yes	<i>ns</i>
Item 21	<i>ns</i>	<i>ns</i>
Item 25	<i>ns</i>	<i>ns</i>
Item 29	<i>ns</i>	<i>ns</i>
Item 32	<i>ns</i>	<i>ns</i>

Scale/Item	Evidence of DIF Black vs. White Students	Evidence of DIF Hispanic/Latinx vs. White Students
Self-Management		
Item 5	<i>ns</i>	<i>ns</i>
Item 14	<i>ns</i>	<i>ns</i>
Item 20	<i>ns</i>	<i>ns</i>
Item 22	<i>ns</i>	<i>ns</i>
Item 30	<i>ns</i>	<i>ns</i>
Item 37	<i>ns</i>	<i>ns</i>
Item 39	<i>ns</i>	<i>ns</i>
Social Awareness		
Item 9	<i>ns</i>	<i>ns</i>
Item 11	<i>ns</i>	<i>ns</i>
Item 16	<i>ns</i>	<i>ns</i>
Item 19	<i>ns</i>	<i>ns</i>
Item 26	<i>ns</i>	<i>ns</i>
Item 33	<i>ns</i>	<i>ns</i>
Relationship Skills		
Item 4	<i>ns</i>	<i>ns</i>
Item 12	<i>ns</i>	<i>ns</i>
Item 17	<i>ns</i>	<i>ns</i>
Item 24	<i>ns</i>	<i>ns</i>
Item 28	<i>ns</i>	<i>ns</i>
Item 35	<i>ns</i>	<i>ns</i>
Item 38	<i>ns</i>	<i>ns</i>
Responsible Decision Making		
Item 2	<i>ns</i>	<i>ns</i>
Item 10	<i>ns</i>	<i>ns</i>
Item 13	<i>ns</i>	<i>ns</i>
Item 18. gather information before making an important decision?	<i>ns</i>	Yes
Item 31	<i>ns</i>	<i>ns</i>
Item 36	<i>ns</i>	<i>ns</i>

Note. *ns* indicates nonsignificant DIF.

In addition to the previously described methods of item evaluation and scale assignment, we assessed each item and scale's performance through IRT techniques. Our primary interest in carrying out these analyses was to either confirm the item- and scale-level conclusions drawn from the techniques described in the previous section (i.e., CTT techniques) or to refine our conclusions using the additional information gained from the IRT analyses. Analyses were completed in flexMIRT version 3.65 (Cai, 2022) and in R using the *ltm* (Rizopoulos, 2006) and *mirt* (Chalmers, 2012) packages. Graded Response Modeling (GRM) models were fit for each iteration of the six DESSA 2 scales. The primary information reviewed to evaluate the items and scales were:

- The Test Information Curve (TIC) for each scale, which indicated how precisely the scale measured the social and emotional construct (e.g., Self-Management) across different levels of the construct.
- The Item Information Curve (IIC) for each item, which indicated how much information each item contributed to the scale across different levels of the construct, compared to the other items on the scale.
- The model summary statistics, which estimated item difficulty and how well each item discriminated among students exhibiting similar levels of the construct (e.g., how well a Self-Management item discriminated between two students with similar competence in Self-Management).

With the techniques described above, we were able to refine our item- and scale-level decisions.

A total of 33 of the 73 items on the standardization form were eliminated through these methods, resulting in a final set of 40 items comprising the six scales. As is typically the outcome in item selection, in a few instances items were retained despite not meeting every a priori criteria. Of the final 40 items, 14 were retained from the original DESSA. Based upon the sum of the standard scores of all six scales, we also created a composite score referred to as the Social-Emotional Composite (SEC), which provides an overall estimate of the student's social and emotional competencies.

Development of the DESSA 2 mini Items

The eight items on each of the four forms of the DESSA 2 mini were selected from the 40 items on the DESSA 2. We followed a method designed to maximize the reliability and equivalence of each of the four forms. Items with the highest corrected item-total correlation with the DESSA 2 SEC *T*-score were rank-ordered, and the first 32 items (eight items for four forms) were identified. The 32 items were then organized by their scale assignment on the DESSA 2 (e.g., Self-Awareness) to examine representation across the six constructs measured by the DESSA 2. Four combinations of items were obtained that (1) had as equal as possible representation across the six DESSA 2 constructs, and (2) had a high average correlation with the DESSA 2 SEC *T*-score. Internal reliability coefficients as well as raw score means and standard deviations were computed for each of the combinations. Item combinations were modified until these values were sufficiently similar to determine the composition of the four DESSA 2

mini forms. Calculation of the raw score to *T*-score normative conversion tables was based on the values obtained for the four forms of the DESSA 2 mini for the educator raters who participated in the DESSA 2 standardization study.

Norming Procedures

The initial step in the preparation of the norms was to determine if any trends existed in the data. We first examined the DESSA 2 scale raw scores for grade and gender differences. [Table 2.6](#) presents the raw score means and standard deviations for the six DESSA 2 scales in three-grade intervals. The SEC is not included in this analysis because it is a derived scale based on the sum of the *T*-scores of the six scales. We then repeated the analysis with the four DESSA 2 mini raw scores. Minor variations in mean raw scores were observed across the three grade bands on the DESSA 2 and DESSA 2 mini forms. To evaluate the practical significance of these mean raw score differences, we calculated *d*-ratios, a measure of effect size. 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 0.2 are negligible, between 0.2 and 0.5 are small, between 0.5 and 0.8 are medium, and those greater than 0.8 are large. Across the 18 possible grade band and scale comparisons for the DESSA 2, 14 were categorized as negligible, 4 were small, and no medium or large effect sizes were observed. Effect sizes ranged from 0.01 to 0.31. Scale raw score means differed by 1.6 raw score points or less for all comparisons. A similar pattern of results was found for the four DESSA 2 mini forms. [Table 2.7](#) presents the raw score means and standard deviations for the four forms of the DESSA 2 mini in three-grade intervals. Minor variations in mean raw scores were observed with all means differing by 1.3 raw score points or less. Across the 12 possible grade band comparisons, 11 were negligible, 1 was small, and no medium or large effect sizes were observed. Effect sizes ranged from 0.01 to 0.23. Given that the scale raw score differences observed across grade bands on both the DESSA 2 and DESSA 2 mini were all negligible to small, we constructed the norms for all grades combined.

TABLE 2.6
DESSA 2 Raw Score Means and Standard Deviations by Grade Bands

Scales	Grades K-2 (<i>n</i> = 455)		Grades 3-5 (<i>n</i> = 645)		Grades 6-8 (<i>n</i> = 450)	
	Mean	SD	Mean	SD	Mean	SD
Optimistic Thinking	17.3	5.9	17.1	5.8	16.4	5.6
Self-Awareness	17.9	5.9	17.5	5.4	16.7	5.5
Self-Management	18.1	6.2	18.1	6.0	17.8	5.8
Social Awareness	16.8	5.1	16.1	4.8	15.2	5.0
Relationship Skills	19.1	5.9	18.8	5.5	17.9	5.6
Responsible Decision Making	15.6	5.1	15.6	5.0	14.9	5.0

TABLE 2.7**DESSA 2 mini Raw Score Means and Standard Deviations by Grade Bands**

Forms	Grades K-2 (<i>n</i> = 455)		Grades 3-5 (<i>n</i> = 645)		Grades 6-8 (<i>n</i> = 450)	
	Mean	SD	Mean	SD	Mean	SD
Form A	21.1	6.8	20.9	6.4	20.0	6.5
Form B	21.0	6.7	20.9	6.4	20.1	6.5
Form C	21.3	6.8	20.8	6.4	19.8	6.4
Form D	21.1	6.6	20.9	6.4	20.0	6.3

Next, we examined mean score differences across the DESSA 2 scales and the SEC by gender. There were statistically significant differences between the ratings for male and female students on all six scales and the SEC, which is consistent with research examining social and emotional skills of children and youth in practice (Kim et al., 2015). [Table 2.8](#) presents the *T*-scale means, standard deviations, and *d*-ratios by scale for male and female students using norms based on all genders combined. The mean-scale *T*-scores for female students are consistently three to four points higher than those for male students. To evaluate the practical significance of these mean-scale *T*-score differences, we calculated *d*-ratios which are presented in [Table 2.8](#). We observed all of these *d*-ratios to be small (0.26–0.38). We next repeated this analysis using the Social-Emotional Total (SET) score on the four forms of the DESSA 2 mini. As shown in [Table 2.9](#), the mean *T*-scores for female students are consistently three to four points higher than those for male students, with effect sizes found to be small (0.32–0.35). The data in [Tables 2.8](#) and [2.9](#) indicate that, as a group, female students consistently show more behaviors related to social and emotional competence than males, but the magnitude of this difference is small.

TABLE 2.8**DESSA 2 Standard Score Gender Differences by Scale**

Scales	Males (<i>n</i> = 787)		Females (<i>n</i> = 763)		Male/Female <i>d</i> -ratio
	Mean	SD	Mean	SD	
Optimistic Thinking	49.5	9.6	52.1	10.0	–0.26
Self-Awareness	49.5	9.5	52.2	10.0	–0.27
Self-Management	48.9	9.8	52.7	9.9	–0.38
Social Awareness	49.4	9.7	52.9	9.7	–0.37
Relationship Skills	49.1	9.8	52.6	10.0	–0.35
Responsible Decision Making	49.2	9.8	52.5	10.0	–0.33
Social-Emotional Composite	48.0	9.7	51.4	10.1	–0.35

Note: All comparisons are significant at $p < .001$.

TABLE 2.9**DESSA 2 mini Standard Score Gender Differences by Form**

Forms	Males (<i>n</i> = 787)		Females (<i>n</i> = 763)		Male/Female <i>d</i> -ratio
	Mean	<i>SD</i>	Mean	<i>SD</i>	
Form A	48.9	9.9	52.2	10.1	–0.33
Form B	49.0	9.7	52.2	10.1	–0.32
Form C	48.9	9.7	52.4	10.0	–0.35
Form D	48.9	9.9	52.4	9.8	–0.35

Note: All comparisons are significant at $p < .001$.

Female students in the DESSA 2 standardization sample earned higher scores than males on each scale. To preserve these noteworthy differences in social and emotional competencies, we constructed the raw-score-to-*T*-score norms conversion tables based on all genders. Consequently, it can be expected that female students will, on average, earn slightly higher scores on the DESSA 2 and the DESSA 2 mini forms than male students. This reflects natural differences commonly observed and establishes a single set of social and emotional competency expectations that applies equally to all students. Our sample was insufficient to provide additional guidance to shape expectations for the social and emotional competencies of nonbinary students relative to their peers.

Next, we examined the distributions of raw scores for normality. The cumulative frequency distributions for the DESSA 2 scales all approached normality, but they were slightly positively skewed. For this reason, we decided to compute norms using a Gaussian Kernel Density Estimation (KDE) approach. Unlike traditional parametric methods, which require assumptions about the data's distribution, KDE is nonparametric which means it can effectively handle data with complex or multimodal distributions, providing a more accurate reflection of the standardization sample. The KDE method provides a smooth estimate of the probability density function (PDF) for the data and helps in generating norms that are robust and less affected by outliers. For the six scales and the SEC, we converted raw scores to *z* scores, and then converted *z* scores to standard scores (*T*-scores with a mean of 50 and a standard deviation of 10). This same approach was followed for constructing the norms for each of the four forms of the DESSA 2 mini separately. 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.



Chapter 3

PSYCHOMETRIC PROPERTIES

CHAPTER 3

Psychometric Properties



As described in Chapter 1, a foundational characteristic of the DESSA suite of assessments is a commitment to strong psychometric properties. The DESSA 2 mini and DESSA 2 were developed to meet or exceed the standards promulgated by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (AERA, 2014). Chapter 2 of this manual describes the large, diverse standardization sample that approximates the population of elementary and middle school-age students in the United States. This chapter will focus on evidence of reliability and validity to support the intended uses of the measures. Together, these important attributes allow for defensible decision making based on students' social and emotional competence.

Reliability

The reliability of an assessment tool like the DESSA 2 mini or the DESSA 2 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). Evidence for the reliability of the DESSA 2 mini and the DESSA 2 were explored using several methods. First, we computed the internal reliability coefficients and the standard errors of measurement for each scale. Second, we computed alternate form reliability for each of the four forms of the DESSA 2 mini. Third, we assessed the test–retest reliability (the same rater evaluating the same student over time) and stability of each scale. Lastly, we assessed the interrater reliability (two raters evaluating the same student) for each scale.

Internal Reliability

Internal reliability (or internal consistency) refers to the extent to which the items on the same scale or instrument are correlated and can be considered to measure the same underlying construct. We determined internal consistency using Cronbach's alpha (Cronbach, 1951). The internal reliability coefficients were based on the students included in the DESSA 2 standardization sample ($N = 1,550$).

Table 3.1 presents the internal consistency estimates for each form of the DESSA 2 mini. The results indicate that the DESSA 2 mini forms have excellent internal reliability. Each of the reliability coefficients exceed the .90 value for a total score suggested by Bracken (1987) and meet the desirable standard described by Nunnally (1978).

Table 3.2 presents the internal consistency estimates for each of the six scales and the Social-Emotional Composite (SEC) score of the DESSA 2. The results indicate that the DESSA 2 scales and total score have excellent internal reliability. The reliability coefficient for the SEC exceeds the .90 value for a total score suggested by Bracken (1987) and meets the desirable standard described by Nunnally (1978). The internal reliability coefficients for the six DESSA 2 scales range from a low of .90 (Self-Awareness) to a high of .93 (Self-Management and Social Awareness), exceeding the .80 desirable standard suggested by Bracken (1987).

TABLE 3.1
Internal Reliability (Alpha) Coefficients and Standard Errors of Measurement for the Four DESSA 2 mini Forms

Form	Alpha Coefficient	SE _M
Form A	.93	2.59
Form B	.93	2.65
Form C	.93	2.65
Form D	.92	2.77

TABLE 3.2
Internal Reliability (Alpha) Coefficients and Standard Errors of Measurement for the DESSA 2 Scales

Scales	Alpha Coefficient	SE _M
Social-Emotional Composite	.98	1.30
Optimistic Thinking	.92	2.90
Self-Awareness	.90	3.13
Self-Management	.93	2.61
Social Awareness	.93	2.70
Relationship Skills	.92	2.81
Responsible Decision Making	.91	2.98

Standard Error 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). As such, the SE_M provides an estimate of the amount of fluctuation in DESSA 2 mini and DESSA 2 scores that can be expected by chance; the larger the SE_M , the greater the amount of chance fluctuation. We obtained the SE_M for each of the four forms of the DESSA 2 mini and the six scales and SEC of the DESSA 2 from the internal reliability coefficients using this formula,

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

where SD is the theoretical standard deviation of the T -score (i.e., 10) and the appropriate reliability coefficient is used. The SE_M values for each form of the DESSA 2 mini are also presented in Table 3.1. The SE_M values for each of the six scales and SEC for the DESSA 2 are also presented in Table 3.2. 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. Because the reliability coefficients in Tables 3.1 and 3.2 meet or exceed recommended standards in the field, the SE_M values are relatively small, indicating that the amount of error observed in a student's DESSA 2 mini or DESSA 2 scores is low.

Alternate Form Reliability

Because there are four forms of the DESSA 2 mini, it is essential to establish the alternate form reliability of this measure. This process calibrates the degree of equivalence of different versions of a measure with different items that are intended to measure the same construct (Rosenthal & Rosnow, 2008). Alternate form reliability analyses were conducted on the students included in the DESSA 2 standardization sample ($N = 1,550$). Alternate form reliability coefficients provided in Table 3.3 indicate that the DESSA 2 mini forms have excellent alternate form reliability. Importantly, the similarity of the form means and standard deviations generated by rating the same students (displayed in Table 3.4) provides evidence of the equivalence and interchangeability of the four forms. Therefore, any differences in T -scores obtained with different forms of the DESSA 2 mini across time or across raters are not due to inherent differences in the DESSA 2 mini forms themselves.

TABLE 3.3

DESSA 2 mini Alternate Form Reliability: Correlation Coefficients

		Form A T -score	Form B T -score	Form C T -score
Form B T -score	r	.94		
Form C T -score	r	.94	.93	
Form D T -score	r	.94	.93	.93

Note: All correlations are significant at the $p < .01$ level (2-tailed).

TABLE 3.4**Four Form Equivalence: DESSA 2 mini Social-Emotional Total T-scores**

Forms	Mean	SD	N
Form A	50.6	10.1	1,550
Form B	50.6	10.0	1,550
Form C	50.6	10.0	1,550
Form D	50.6	10.0	1,550

Test-Retest Reliability

The correlation between scores obtained for the same students on two separate occasions is another indicator of the reliability of an 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 2 mini and the DESSA 2, a group of educators ($N = 53$) from ten schools rated the same K–8 students on two different occasions separated by an interval of four to eight days. Demographic information on this sample is provided in [Table 3.5](#).

TABLE 3.5**Sample Characteristics for the DESSA 2 mini and DESSA 2 Test-Retest Reliability Study**

	Educator Sample ($N = 53$)	
	n	%
Grade		
Kindergarten	8	15.1
1st	7	13.2
2nd	8	15.1
3rd	6	11.3
4th	5	9.4
5th	5	9.4
6th	2	3.8
7th	5	9.4
8th	7	13.2
Gender		
Male	28	52.8
Female	25	47.2

	Educator Sample (N = 53)	
	<i>n</i>	%
Race		
American Indian/Alaskan Native	2	3.8
Asian	1	1.9
Black/African American	6	11.3
Native Hawaiian/Other Pacific Islander	0	0
White	33	62.3
Two or More	2	3.8
Don't Know/Missing	9	17.0
Ethnicity		
Hispanic/Latinx	7	13.2
Not Hispanic/Latinx	43	81.1
Don't Know	3	5.7
Region of Residence		
Northeast	6	11.3
Midwest	19	35.8
South	8	15.1
West	20	37.7
Free or Reduced-Price Lunch Eligibility		
Yes	11	20.8
No	21	39.6
Don't Know	21	39.6

The results of this study are shown in [Table 3.6](#) for the DESSA 2 mini forms and [Table 3.7](#) for the DESSA 2. All correlations are significant ($p < .01$) and high in magnitude. For the DESSA 2 mini forms, the coefficients range from $r = .86$ on Form D to $r = .91$ on Form C. The coefficient for the DESSA 2 SEC score was .91, while the coefficients ranged from $r = .87$ (Optimistic Thinking) to $r = .91$ (Social Awareness and Relationship Skills) across the six DESSA 2 scales. These findings indicate that the DESSA 2 mini and DESSA 2 have good test–retest reliability.

TABLE 3.6
Test-Retest Reliability and Stability Results for DESSA 2 mini Ratings

Forms	<i>r</i>	First Rating		Second Rating		T-score Difference
		Mean	SD	Mean	SD	
Form A	.89	47.6	10.1	47.4	10.2	0.2
Form B	.88	45.9	9.8	46.6	9.9	–0.7
Form C	.91	48.0	9.9	48.0	10.5	0.0
Form D	.86	47.3	10.2	47.9	10.4	–0.6

Note: All correlations are significant at $p < .01$.

TABLE 3.7**Test-Retest Reliability and Stability Results for DESSA 2 Ratings**

Scales	<i>r</i>	First Rating		Second Rating		<i>T</i> -score Difference
		Mean	<i>SD</i>	Mean	<i>SD</i>	
Social-Emotional Composite	.91	46.2	9.5	46.3	9.9	–0.1
Optimistic Thinking	.87	44.9	9.7	46.1	9.9	–1.2
Self-Awareness	.89	46.8	9.6	46.2	10.0	0.6
Self-Management	.88	48.6	10.6	49.2	10.6	–0.6
Social Awareness	.91	47.9	10.5	48.6	10.6	–0.7
Relationship Skills	.91	48.6	10.4	48.1	10.7	0.5
Responsible Decision Making	.88	48.4	10.4	48.0	10.3	0.4

Note: All correlations are significant at $p < .01$.

Stability of the Test–Retest Ratings

The correlation coefficients reported above for the test–retest reliability study indicate that educators ranked students similarly across the two DESSA 2 ratings completed about one week apart. However, the coefficients do not describe the actual similarity in the scores. To examine score stability across one week, the second rating *T*-score for each student on each scale was subtracted from the corresponding first rating *T*-score. Using this approach, identical scores on the two ratings would result in a value of 0. Table 3.6 for the DESSA 2 mini forms and Table 3.7 for the DESSA 2 provides the test–retest mean scores and standard deviations received by the students in the test–retest reliability study.

As shown in Table 3.6, on average, the mean value of the test–retest difference on the four forms of the DESSA 2 mini was less than 1 *T*-score point (–0.3), ranging from –0.7 (Form B) to 0.2 (Form A) *T*-score units. As shown in Table 3.7, the mean score difference on the DESSA 2 SEC was less than 1 *T*-score point (–0.1). On average, the mean value of the test–retest difference on the six social and emotional competence scales was also less than 1 *T*-score point (–0.2), ranging from –1.2 (Optimistic Thinking) to 0.6 (Self-Awareness). Paired samples *t*-tests conducted for each mean score comparison on the DESSA 2 measures yielded no significant differences between the first and second ratings. These results demonstrate that the DESSA 2 mini forms and the DESSA 2 scales and SEC have good stability across a four- to eight-day interval for educator raters. This, in turn, provides increased confidence that, when differences are observed between pretest and posttest scores, they are less likely to be attributed to error variance or the simple passage of time.

Interrater Reliability

The correlation between scores obtained for the same student at 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' perception of the student's behavior. The optimal condition for evaluating the interrater

reliability of an assessment tool is to have two raters observing the same student in the same environment at the same time. Because this optimal condition is not the norm in most educational settings, particularly for middle school grades, we collected a sample of DESSA 2 ratings from two educators who work in the same classroom (e.g., a teacher and teacher aide) or from two educators who see the same student in different classrooms for core academic subjects. We hypothesized some degree of similarity between these raters, but expected some incongruence (i.e., we expected the correlations would be moderate in size). This can occur because the two raters see the student in different environments, where the student's behavior may differ.

The interrater reliability sample included 44 unique pairs of educators rating the same student. Educators included individuals who identified themselves as teachers, teacher aides, or specialists working in the general education classroom. Demographic information on this sample is presented in [Table 3.8](#).

TABLE 3.8
Sample Characteristics for the DESSA 2 mini and DESSA 2 Interrater Reliability Study

	Educator Sample (N = 44)	
	<i>n</i>	%
Grade		
Kindergarten	5	11.4
1st	5	11.4
2nd	5	11.4
3rd	6	13.6
4th	4	9.1
5th	6	13.6
6th	2	4.5
7th	3	6.8
8th	8	18.2
Gender		
Male	25	56.8
Female	19	43.2
Race		
American Indian/Alaskan Native	2	4.5
Asian	0	0
Black/African American	6	13.6
Native Hawaiian/Other Pacific Islander	0	0
White	25	56.8
Two or More	3	6.8
Don't Know/Missing	8	18.2

(continued)

	Educator Sample (N = 44)	
	<i>n</i>	%
Ethnicity		
Hispanic/Latinx	7	15.9
Not Hispanic/Latinx	35	79.5
Don't Know	2	4.5
Region of Residence		
Northeast	9	20.5
Midwest	11	25.0
South	8	18.2
West	16	36.4
Free or Reduced-Price Lunch Eligibility		
Yes	13	29.5
No	15	34.1
Don't Know	16	36.4

The correlations of a set of ratings obtained for the same student by two teachers (or a teacher and teacher aide or specialist) are provided in [Table 3.9](#) for the DESSA 2 mini forms and [Table 3.10](#) for the DESSA 2. These results indicate that pairs of educators who saw the students in the same or a similar educational environment within three days of one another rated the students similarly. All the correlations are significant ($p < .01$) and moderate in magnitude. The coefficients range from $r = .60$ (Forms A and B) to $r = .72$ (Form D) on the DESSA 2 mini. The coefficient for the DESSA 2 SEC score was .65, while the DESSA 2 scale coefficients ranged from $r = .43$ (Self-Awareness) to $r = .73$ (Self-Management).

TABLE 3.9
Interrater Reliability Results for DESSA 2 mini Ratings

Forms	<i>r</i>	Rater 1		Rater 2		<i>T</i> -score Difference
		Mean	<i>SD</i>	Mean	<i>SD</i>	
Form A	.60	49.9	10.2	48.2	9.4	1.7
Form B	.60	47.1	9.2	46.9	9.2	0.2
Form C	.64	49.7	9.6	48.4	8.7	1.3
Form D	.72	49.1	10.1	48.4	8.7	0.7

Note: All correlations are significant at $p < .01$.

TABLE 3.10**Interrater Reliability Results for DESSA 2 Ratings**

Scales	<i>r</i>	Rater 1		Rater 2		<i>T</i> -score Difference
		Mean	<i>SD</i>	Mean	<i>SD</i>	
Social-Emotional Composite	.65	47.8	9.4	46.9	8.7	0.9
Optimistic Thinking	.59	47.4	9.4	45.3	9.9	2.1
Self-Awareness	.43	48.2	8.2	47.7	9.9	0.5
Self-Management	.73	49.5	10.5	48.5	9.7	1.0
Social Awareness	.57	49.4	10.0	50.1	9.5	−0.7
Relationship Skills	.58	50.4	10.0	49.3	9.3	1.1
Responsible Decision Making	.68	49.7	9.5	48.9	9.1	0.8

Note: All correlations are significant at $p < .01$.

Stability of the Interrater Ratings

The coefficients do not indicate the actual similarity in the scores. Table 3.9 for the DESSA 2 mini and Table 3.10 for the DESSA 2 provide the mean scores and standard deviations students received by both raters in the interrater reliability study. As shown in Table 3.9, on average, the mean value of the test–retest difference on the four forms of the DESSA 2 mini was 1 *T*-score point (0.98), ranging from 0.2 (Form B) to 1.7 (Form A) *T*-score units. As shown in Table 3.10, the mean score difference on the DESSA 2 SEC was also 1 *T*-score point (0.9). On average, the mean value of the test–retest difference on the six social and emotional competence scales was less than 1 *T*-score point (0.8), ranging from −0.7 (Social Awareness) to 2.1 (Optimistic Thinking). Paired samples *t*-tests conducted for each mean score comparison on the DESSA 2 measures yielded no significant differences between the two educator raters. These results demonstrate that the scores obtained on the four DESSA 2 mini forms and the DESSA 2 are similar across educator raters rating the same student in similar educational environments. This, in turn, provides increased confidence that a DESSA 2 mini or DESSA 2 score is less likely to be attributable to error variance in the rater than reflective of the student’s actual social and emotional competence.

Reliability Study Summary

The results of several reliability studies of the four forms of the DESSA 2 mini and the DESSA 2 provide evidence that the scores these measures yield are reliable for assessing K–8 students’ social and emotional competence. The results of the internal consistency data demonstrate that the DESSA 2 measures meet or exceed the standards suggested by Bracken (1987). The alternate form study results demonstrated excellent reliability between the four forms of the DESSA 2 mini, providing evidence of the equivalence and interchangeability of the four forms. The test–retest study shows that educator raters rank students’ scores on the DESSA 2 measures similarly over relatively brief periods of time. The stability findings from this study further

indicate that not only the rankings but also the actual mean scores received by the student at different points in time over a relatively brief interval are quite similar. The results of the inter-rater reliability study show that different educator raters rank students' scores similarly. The stability findings from this study further indicate that not only the rankings, but also the actual mean scores received on the DESSA 2 measures are similar. These studies are important because they indicate that when differences are found across time or raters, they are likely to reflect meaningful differences such as responses to strategies or interventions or actual differences between a student's behavior in different environments.

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 evidence and theory support the interpretations of test scores for proposed uses of tests” (AERA, 2014, p. 11). According to the *Standards for Educational and Psychological Testing* (AERA, 2014), the sources of validity evidence can be conceptualized in various ways. We investigated the validity of the DESSA 2 mini and the DESSA 2 regarding *test content* (content validity), *internal structure and relations to other variables* (construct validity), and *use of the DESSA 2 mini as a screener for social and emotional competence*.

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 2, content-related validity addresses how well the 40 items (as well as the 8 items on each form of the DESSA 2 mini) represent the domain of behavioral characteristics related to social and emotional competence in elementary and middle school students.

As detailed in Chapter 2, we based the items comprising the DESSA 2 on a thorough review of the literature on social and emotional competence, positive youth development, and resilience in elementary and middle school students. The items were based, in part, on the first edition of the DESSA (LeBuffe et al., 2009/2014), which has its own research base (for a review, see Hwang et al., 2023 and LeBuffe et al., 2018) and was developed to align to the CASEL Framework (www.casel.org). Prior to conducting national standardization, the items were reviewed by four experts in the field of social and emotional development, including individuals with expertise in the related fields of educational equity, special education, and school psychology and mental health. As described in Chapter 2, the reviewers positively evaluated the item pool from the perspectives of social and emotional competence content coverage, developmental appropriateness for elementary and middle school students, and equity across diverse groups of students.

Construct-Related Validity

This type of validity examines the degree to which the assessment instrument measures the theoretical construct of interest. In the case of the DESSA 2, two types of construct validity were investigated. The first pertains to the DESSA 2's internal scale structure, examined using confirmatory factor analysis and variability of scale scores. These studies are discussed in the Internal Structure section. The second concerns the relationships between DESSA 2 scale scores and scores on other well-developed measures of social and emotional behavioral strengths in students, specifically the SSIS SEL Brief Scale (Elliott et al., 2020) and the first edition of the DESSA (LeBuffe et al., 2009/2014). These studies are discussed in the section entitled Convergent Validity (see page 44).

Internal Structure

One approach to establishing construct validity is to examine the internal structure of an assessment to determine the degree to which relationships among the items conform to the construct(s) on which score interpretations are based. Chapter 2 of this manual described the item- and scale-level analyses completed to guide the organization of the DESSA 2 items into statistically and logically derived scales. We examined this scale structure of the DESSA 2 using confirmatory factor analysis. It should be noted that because our intent was to align the DESSA 2 to the CASEL Framework and the existing suite of DESSA assessments, we did not conduct an exploratory factor analysis before proceeding to the confirmatory factor analysis.

Confirmatory Factor Analysis. To better explore the validity of the DESSA 2's scale structure through factor analysis, confirmatory factor analysis was completed among the standardization sample, excluding cases missing one or more item response(s) ($N = 1,550$). We fit a six-factor model in which each item was assigned to one factor in alignment with its earlier assignment to one of the six DESSA 2 scales (Self-Management, Relationship Skills, etc.). Chapter 2 of this manual provides a discussion of the assignment of items to the six scales.

Confirmatory factor analysis was completed in R using the lavaan package (Rosseel, 2012). Weighted Least Square Mean and Variance Adjusted Estimators (WLSMV) were used, given the ordinal nature of the data (Li, 2016). The six-scale solutions exhibited a good model fit as described by Hu and Bentler (1999), indicated by a Tucker-Lewis Index (TLI) value of 0.995 and a Root Mean Square Error of Approximation (RMSEA) value of .038.

This evidence suggests that the six-factor DESSA 2 model fits the standardization data well. For comparison, two alternative models were explored, representing other popular conceptualizations of social and emotional competencies:

1. A three-factor model that assigned items to three factors: Intra-Personal (comprised of the DESSA 2 scales of Optimistic Thinking, Self-Awareness, and Self-Management); Inter-Personal (comprised of the DESSA 2 scales of Social Awareness and Relationship Skills); and Decision Making (comprised of the DESSA 2 scale of Responsible Decision Making).
2. A one-factor model that assigned all items to a single factor.

Fit indices for the six-scale model and the two additional models are presented in [Table 3.11](#). Each model tested exhibits a high TLI value (ranging from 0.993 for the one-scale model to 0.995 for the three- and six-scale models) and a low RMSEA value (ranging from 0.038 for the six-scale model to 0.045 for the one-scale model), indicating a good fit to the data.

The model fit indices suggest that all tested models fit the data well. To evaluate the fit of the proposed DESSA 2 model relative to the alternative models, the proposed DESSA 2 model was compared to the three-scale model and one-scale model, pairwise, via a series of scaled chi-square difference tests. The results of the pairwise comparisons are included in [Table 3.12](#).

These results indicate that the proposed DESSA 2 six scale model fits the data significantly better than the tested three-scale model and the one-scale model. Marginal improvements in TLI and RMSEA values suggest that the model that assigns DESSA 2 items to scales as described in Chapter 2 fits the data *as well as*, if not *slightly better than*, the alternatives tested.

TABLE 3.11
Fit Indices for the DESSA 2 Six-Scale Model and Two Alternative Models

Model	Test Statistic (Standard) // <i>p</i> -Value (Chi-Square)	Test Statistic (Robust) // <i>p</i> -Value (Chi-Square)	Degrees of Freedom	Tucker-Lewis Index (TLI)	Root Mean Square Error of Approximation (RMSEA)
Six-Scale Model	2341.74 // <i>p</i> < .001	4791.57 // <i>p</i> < .001	725	.995	.038
Three-Scale Model	2594.09 // <i>p</i> < .001	5118.44 // <i>p</i> < .001	737	.995	.040
One-Scale Model	3072.25 // <i>p</i> < .001	5764.07 // <i>p</i> < .001	740	.993	.045

TABLE 3.12
Comparisons Between the DESSA 2 Six-Scale Model and Two Alternative Models

Comparison	Chi-Square of Six-Scale Model	Chi-Square of Comparison Model	Chi-Square Difference	df Difference	<i>p</i>
Six-Scale Model vs. Three-Scale Model	2341.7	2594.1	276.78	12	<i>p</i> < .001

Variability of DESSA 2 Scale Scores. Evidence for the construct validity of DESSA 2 scales was also explored through an examination of the variability of scale scores. For each student in the standardization sample ($N = 1,550$), the student's highest scale *T*-score and lowest scale *T*-score were identified. We calculated the difference between maximum and minimum *T*-scores and expressed these results as a frequency distribution and descriptive statistics of the *T*-score difference. These results are presented in [Table 3.13](#).

TABLE 3.13

Cumulative Frequencies of the *T*-score Difference Between the Highest and Lowest DESSA 2 Scale Scores

Scale Difference	Cumulative Percent
0	0.3
1	0.8
2	1.9
3	7.5
4	14.6
5	24.3
6	34.8
7	46.0
8	56.3
9	65.3
10	71.2
11	77.9
12	83.0
13	87.2
14	90.3
15	92.6
16	94.3
17	95.5
18	96.4
19	97.0
20	97.7
21	98.3
22	98.8
23	99.0
24	99.2
25	99.3
26	99.5
27	99.7
28	99.7
29	99.8
30	99.9
31	99.9
32	100.0
Mean	8.72
<i>SD</i>	4.48
<i>N</i>	1,550

There are several important points to consider when examining the variability of DESSA 2 scale scores. First, the mean difference between students' highest and lowest *T*-scores is 8.7 ($SD = 4.5$). This means that the typical K–8 student will show a difference of about 8 *T*-score points between the highest and lowest of the six DESSA 2 scales. Second, the cumulative percentages of DESSA 2 scale *T*-score differences reported in Table 3.13 tells us that few students (24.3%) rated by an educator had minimal or no variation (defined as five or fewer points) between their highest and lowest DESSA 2 scale *T*-score. Similarly, very few students (7.4%) had a difference of 15 points or more. This, along with the mean difference reported at the bottom of Table 3.13, indicated that typically, the six DESSA 2 scales do differ from one another and are measuring different social and emotional domains.

As Chapter 5 of this manual will explain, using the numerical scale score provides important information about the degree to which the student is similar to, or not similar to, the normative group. However, scale scores can also be examined within each student to consider whether the student is showing an expected or unusual amount of intra-scale variability on the DESSA 2 and to identify their relative strengths or needs for instruction as an individual.

Convergent Validity

One common approach to establishing the construct validity of an assessment tool is to demonstrate that scores on the measure in question correlate positively with scores of similar constructs on other well-developed measures. This is referred to as convergent validity. To provide evidence of convergent validity, we correlated *T*-scores on the DESSA 2 with raw scores from the SSIS SEL Brief Scale (Elliott et al., 2020), a behavior rating scale that assesses the social and emotional learning skills of children and adolescents. Specifically, correlations were made between (a) the DESSA 2 SEC score with the SSIS SEL scale and composite scores, (b) the SSIS SEL Composite score with the DESSA 2 SEC and scale scores, and (c) DESSA 2 and SSIS SEL scale scores based on theoretical similarities in the constructs being measured (e.g., the Self-Management scales on both measures). We hypothesized scores across these comparisons would yield moderate correlations.

Educators ($N = 60$) completed the DESSA 2 and the SSIS SEL in one session. The demographic characteristics of the students involved in this study are presented in Table 3.14. These data indicate that this sample was diverse in terms of demographics.

Descriptive statistics for the DESSA 2 and the SSIS SEL ratings are presented in Table 3.15. Pearson product-moment correlations between the DESSA 2 and the SSIS SEL scale and total scores are displayed in Table 3.16. The DESSA 2 SEC correlated significantly ($r = .92$, $p < .01$) with the SSIS Composite score and its five scale scores, including Self-Awareness ($r = .80$, $p < .01$), Self-Management ($r = .79$, $p < .01$), Social Awareness ($r = .84$, $p < .01$), Relationship Skills ($r = .84$, $p < .01$), and Responsible Decision Making ($r = .87$, $p < .01$). Comparisons at the scale level across the two measures also yielded significant moderate correlations, as can be seen in Table 3.16. Taken together, the results suggest that the DESSA 2 corresponds closely to another psychometrically sound instrument that measures similar constructs, yet the correlations were not so high as to raise concerns about the measures being exact replications of each other.

TABLE 3.14**Sample Characteristics for the DESSA 2 Construct Validity Sample**

	Educator Sample (N = 60)	
	<i>n</i>	%
Grade		
Kindergarten	10	16.7
1st	6	10.0
2nd	9	15.0
3rd	7	11.7
4th	6	10.0
5th	5	8.3
6th	3	5.0
7th	6	10.0
8th	8	13.3
Gender		
Male	31	51.7
Female	29	48.3
Race		
American Indian/Alaskan Native	3	5.0
Asian	0	0
Black/African American	5	8.3
Native Hawaiian/Other Pacific Islander	0	0
White	40	66.7
Two or More	2	3.3
Don't Know/Missing	10	16.7
Ethnicity		
Hispanic/Latinx	8	13.3
Not Hispanic/Latinx	50	83.3
Don't Know/Missing	2	3.3
Region of Residence		
Northeast	7	11.7
Midwest	22	36.7
South	12	20.0
West	19	31.7
Free or Reduced-Price Lunch Eligibility		
Yes	11	18.6
No	21	35.6
Don't Know/Missing	27	45.8

TABLE 3.15

Results of the DESSA 2 Construct Validity Study ($N = 60$): Means and Standard Deviations of the DESSA 2 and the SSIS SEL Scales and Composite Scores

	Mean	SD
SSIS SEL		
SSIS Self-Awareness Scale Score	6.7	2.4
SSIS Self-Management Scale Score	7.3	2.7
SSIS Social Awareness Scale Score	7.7	2.7
SSIS Relationship Skills Scale Score	7.8	2.4
SSIS Responsible Decision Making Scale Score	8.2	2.7
SSIS Composite SEL Raw Scores	37.6	11.7
DESSA 2		
Self-Awareness <i>T</i> -score	47.5	9.0
Self-Management <i>T</i> -score	48.8	10.0
Social Awareness <i>T</i> -score	49.0	10.3
Relationship Skills <i>T</i> -score	48.1	10.0
Responsible Decision Making <i>T</i> -score	48.8	10.2
Optimistic Thinking <i>T</i> -score	47.0	9.4
Social-Emotional Composite <i>T</i> -score	46.8	9.7

TABLE 3.16

Results of the DESSA 2 Construct Validity Study ($N = 60$): Correlation of the DESSA 2 with the SSIS SEL

SSIS SEL	DESSA 2						
	SA	SM	SO	RS	RDM	OT	SEC
	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
SSIS Self-Awareness Scale Score	.78	—	—	—	—	—	.80
SSIS Self-Management Scale Score	—	.83	—	—	—	—	.79
SSIS Social Awareness Scale Score	—	—	.83	—	—	—	.84
SSIS Relationship Skills Scale Score	—	—	—	.85	—	—	.84
SSIS Responsible Decision Making Scale Score	—	—	—	—	.88	—	.87
SSIS Composite SEL Raw Scores	.71	.90	.88	.89	.91	.83	.92

Note: All correlations are significant at $p < .01$.

SA = Self-Awareness; SM = Self-Management; SO = Social Awareness; RS = Relationship Skills; RDM = Responsible Decision Making; OT = Optimistic Thinking; SEC = Social-Emotional Composite.

Comparability of the First Edition DESSA and the DESSA 2

As described in Chapter 1, the DESSA 2 differs from its first edition counterpart in a number of important ways. First, about 65% of the items on the DESSA 2 are different from the original DESSA. Second, the DESSA 2 scale structure has been updated to measure six social and emotional competency domains, while the first edition DESSA measured eight domains. Third, new norms with a contemporary standardization sample have been developed. These changes suggest that the comparability of the scale *T*-scores obtained with the original DESSA versus those obtained with the new items, scales, and norms should be examined. In other words, would a student who was rated on the DESSA and received a SEC *T*-score of, for example, 50 expect to receive a similar score on the DESSA 2?

To examine this important question, we recruited a sample of educators ($N = 63$) to complete both the first edition DESSA items as well as the DESSA 2 items on the same student during the same session. The two assessments were scored using their respective normative samples (i.e., the 2005–2006 norms for the original DESSA and the 2023–2024 norms for the DESSA 2). We hypothesized moderate to high correlations between the two assessments at both the SEC and the scale level for theoretically similar constructs. To account for the differences between the six constructs assessed on the DESSA 2 and the eight constructs assessed on the DESSA, we compared the DESSA 2 Self-Management score with both the Self-Management and Goal-Directed Behavior scales on the DESSA. Similarly, the DESSA 2 Responsible Decision Making scale was compared to both the DESSA Decision Making and Personal Responsibility scales. The demographic characteristics of the students involved in this study are presented in [Table 3.17](#).

TABLE 3.17
Sample Characteristics for the DESSA and DESSA 2 Comparability Study

	Educator Sample ($N = 63$)	
	<i>n</i>	%
Grade		
Kindergarten	10	15.9
1st	6	9.5
2nd	11	17.5
3rd	7	11.1
4th	7	11.1
5th	5	7.9
6th	3	4.8
7th	6	9.5
8th	8	12.7

(continued)

	Educator Sample (<i>N</i> = 63)	
	<i>n</i>	%
Gender		
Male	34	54.8
Female	28	45.2
Race		
American Indian/Alaskan Native	2	3.2
Asian	1	1.6
Black/African American	6	9.5
Native Hawaiian/Other Pacific Islander	0	0
White	33	52.4
Two or More	4	6.3
Don't Know/Missing	17	27.0
Ethnicity		
Hispanic/Latinx	7	11.1
Not Hispanic/Latinx	51	81.0
Don't Know	5	7.9
Region of Residence		
Northeast	8	12.7
Midwest	23	36.5
South	11	17.5
West	21	33.3
Free or Reduced-Price Lunch Eligibility		
Yes	10	15.9
No	16	25.4
Don't Know/Missing	37	58.7

The results of the study are presented in [Table 3.18](#). As can be seen, scores on the two assessments correlated highly with one another. Specifically, the DESSA 2 SEC score correlated significantly ($r = .98, p < .01$) with the DESSA SEC score. Similarly, moderate to high significant correlations were also observed between the DESSA and DESSA 2 corresponding scale scores.

Although the correlations were high, it could still be the case that the actual *T*-scores received by the students on comparable scales differed. The mean *T*-score difference between the SEC score on the two assessments was found to be less than 1 *T*-score point (0.8). The average *T*-score difference across the scale comparisons was also less than 1 *T*-score point (–0.3). The scale comparisons ranged from –3.3 (DESSA Goal-Directed Behavior vs. DESSA 2 Self-Management) to 1.7 (DESSA Decision Making to DESSA 2 Responsible Decision Making). Paired samples *t*-tests conducted for each mean score comparison yielded significant differences between the DESSA and DESSA 2 SEC ($p = .003; d = .39$), the DESSA Goal-Directed Behavior

scale and the DESSA 2 Self-Management scale ($p < .001$; $d = .60$), and the DESSA Decision Making scale and the DESSA 2 Responsible Decision Making scale ($p = .002$; $d = .40$), with effect size estimates considered to be small for the SEC and Decision Making/Responsible Decision Making comparison and moderate for the Goal-Directed Behavior/Self-Management comparison. All other comparisons yielded no significant differences.

Overall, these results indicate that scores on the 2009 DESSA and the 2024 DESSA 2 are quite similar. At the total score level, students' scores on the DESSA and DESSA 2 were very highly correlated ($r = .98$). Mean DESSA 2 SEC scores were found to be 0.8 *T*-score points lower than the corresponding DESSA SEC scores. However, the magnitude of this difference is small. At the scale level, scales that have remained theoretically similar on the two assessments (e.g., Self-Awareness, Social Awareness, Relationship Skills, and Optimistic Thinking) are highly correlated with one another and show a nonsignificant mean *T*-score difference of about 1 *T*-score point. Significant differences at the scale level were only observed for the scales that have undergone substantial updates as part of the item content revision for the DESSA 2. Taken together, these results suggest that in those cases where the two different forms are being compared (e.g., pretest and posttest comparisons), one can expect that, in general, the score on the DESSA 2 SEC may be slightly lower than the score that would have been received had the DESSA been used for both ratings. When possible, we recommend comparisons between the two forms at the total score level or between theoretically similar scale scores. In addition to practice guidance implications, these findings provide evidence for generalizing research done on the DESSA to the DESSA 2.

TABLE 3.18

Results of the Comparability Study ($N = 63$): Means, Standard Deviations, and Correlations of the DESSA and DESSA 2 Scales and Composite Scores

DESSA Scales	DESSA		DESSA 2		Mean <i>T</i> -score Difference	<i>r</i>
	Mean	<i>SD</i>	Mean	<i>SD</i>		
Self-Awareness	45.3	9.5	45.9	9.9	−0.6	.90
Self-Management	47.3	9.0	47.7*	9.2	−0.4	.88
Goal-Directed Behavior	44.4	9.6			−3.3	.82
Social Awareness	46.9	9.7	47.9	9.0	−1.0	.85
Relationship Skills	47.8	9.4	47.4	8.9	0.4	.83
Personal Responsibility	46.2	9.3	45.6*	9.0	0.6	.79
Decision Making	47.3	9.4			1.7	.90
Optimistic Thinking	45.0	10.3	44.4	10.3	0.6	.88
Social-Emotional Composite	45.8	9.1	45.0	8.7	0.8	.98

* The DESSA 2 construct of Self-Management includes components of the DESSA scales Self-Management and Goal-Directed Behavior. Therefore, the DESSA 2 Self-Management score is being compared to these two DESSA scales. The DESSA 2 construct of Responsible Decision Making includes components of the DESSA scales Personal Responsibility and Decision Making. All correlations are significant at $p < .01$.

Use of the DESSA 2 mini as a Screener for Social and Emotional Competence

The primary use of the DESSA 2 mini is as a universal screener for social and emotional competence in elementary and middle school students. As such, we investigated the validity of the DESSA 2 mini in relation to its ability to meet the stated goals of (a) high correlation with the DESSA 2 Social-Emotional Composite (SEC) *T*-score; and (b) consistency of identification of individual students with the DESSA 2 and DESSA 2 mini. Goal (a) was evaluated by correlating the DESSA 2 mini SET *T*-scores with the full DESSA 2 SEC *T*-scores. Goal (b) was examined by determining the percentages of students identified by both the DESSA 2 and DESSA 2 mini as in need of further instruction in the social and emotional domain.

DESSA 2 mini Correlations with the DESSA 2

As a screener, SET scores on the DESSA 2 mini should correlate strongly with the SEC scores of the full DESSA 2. The means, standard deviations, and correlations of the DESSA 2 SEC *T*-scores with each DESSA 2 mini SET *T*-score are provided in Table 3.19. All of the correlations are significant ($p < .01$) and high in magnitude. These results suggest that the correlations between the four mini forms and the DESSA 2 are quite strong. The values should be considered somewhat inflated because the items in each DESSA 2 mini form are also included in the DESSA 2. Table 3.20 provides the results when the four DESSA 2 mini *T*-scores were correlated with the DESSA 2 total item raw scores excluding the items from each respective DESSA 2 mini form. Again, all of the correlations were significant at $p < .01$ and high in magnitude. The findings again indicate that each of the four DESSA 2 mini forms produced scores that are strongly correlated with DESSA 2 scores.

TABLE 3.19

Means, *SD*s, *N*s, and Correlations of the DESSA 2 with the Four DESSA 2 mini Forms

DESSA 2 mini SET <i>T</i> -scores	DESSA 2 SEC <i>T</i> -scores			
	<i>r</i>	Mean	<i>SD</i>	<i>N</i>
Form A	.98	50.6	10.1	1,550
Form B	.97	50.6	10.0	1,550
Form C	.97	50.6	10.0	1,550
Form D	.97	50.6	10.0	1,550
DESSA 2 SEC	—	49.7	10.0	1,550

Note: All correlations are significant at $p < .01$.

TABLE 3.20

Means, *SDs*, *Ns*, and Correlations of the DESSA 2 Raw Scores with the Four DESSA 2 mini *T*-scores Corrected for Item Overlap

DESSA 2 mini <i>T</i> -scores	DESSA 2 Raw Scores					
	DESSA 2 Raw Score	DESSA 2 Raw Score Minus Mini A Items	DESSA 2 Raw Score Minus Mini B Items	DESSA 2 Raw Score Minus Mini C Items	DESSA 2 Raw Score Minus Mini D Items	<i>N</i>
Form A	.98	.96	.98	.98	.98	1,550
Form B	.97	.97	.96	.97	.97	1,550
Form C	.97	.97	.97	.95	.97	1,550
Form D	.97	.97	.97	.97	.95	1,550
Mean	102.4	81.7	81.8	81.8	81.7	1,550
<i>SD</i>	31.0	24.6	24.7	24.8	24.8	1,550

Note: All correlations are significant at $p < .01$.

Consistency of Identification of Students in Need of Instruction

As a measure of behaviors related to social and emotional competence, scores on the DESSA 2 mini should predict social and emotional functioning of school-aged students. Given that one of the main purposes of the DESSA 2 mini is to identify students whose score indicates a need for instruction, one source of validity is the extent to which decisions made on the basis of the DESSA 2 mini SET *T*-score are consistent with decisions made on the basis of scores on the DESSA 2 SEC *T*-score. To answer this question, we examined how often each student scored 40 or less on both the DESSA 2 SEC and each DESSA 2 mini as well as those who obtained scores above 40 on both measures. Table 3.21 provides evidence that there is considerable agreement between each of the four DESSA 2 mini forms and the DESSA 2 SEC using the cutoff score of 40 to determine if the student is in need of instruction. That is, the DESSA 2 mini was accurate (e.g., the student needs instruction or not) 94.1% to 95.7% of the time. These findings suggest that when the DESSA 2 mini is used, a high degree of confidence can be had about the selection of students for intervention or instruction.

TABLE 3.21

Rates of Agreement Between Each DESSA 2 mini and the Social-Emotional Composite *T*-score From the DESSA 2

	Form A	Form B	Form C	Form D
Number of Cases	1,550	1,550	1,550	1,550
Number of Agreements	1,483	1,478	1,459	1,463
Percent Agreement	95.7%	95.4%	94.1%	94.4%

Note: Percent agreement was based on the number of cases where the DESSA 2 mini and the DESSA 2 Social-Emotional Composite yielded the same conclusion (need or not need instruction) over the total number of cases.

Validity Study Summary

The content-related validity evidence provided in this chapter associated the DESSA 2 items with both the research and practice literature on social and emotional competence in students and was supported through experts' review of the content. The construct-related validity studies provide evidence in support of the six-scale model structure of the DESSA 2 and demonstrate that the DESSA 2 scales show convergent validity with a similar strength-based measure. The first edition DESSA and DESSA 2 comparability study provide evidence for the similarity of scores across the two editions. The studies of the DESSA 2 mini indicate that it can be used with confidence as a screener for social and emotional competence. Specifically, DESSA 2 mini Social-Emotional Total scores are strongly correlated with the Social-Emotional Composite scores on the full DESSA 2. In the large majority of cases, both assessments identify the same students as being in need of instruction.

The authors of the DESSA 2 welcome any opportunities to assist other researchers in further exploring the validity and utility of the measures in assessing and ultimately helping to promote the social and emotional competence of students. The authors can be reached through www.ApertureEd.com.



Chapter 4

ADMINISTRATION AND SCORING

CHAPTER 4

Administration and Scoring



General Administration Guidelines

The DESSA 2 mini and DESSA 2 can be completed by an educator. This typically includes teachers, teacher aides, assistant teachers, instructional assistants, out-of-school time (OST) program staff, and other professionals who interact directly with the student on a regular basis. For simplicity, these individuals are referred to as “educators” in the online system for administering the DESSA 2 and in associated resources. The person who completes and provides the ratings is referred to as the “rater.” The person who interprets and uses the ratings is referred to as the “user” and is often the same person as the rater. However, student support personnel such as school counselors, psychologists, and social workers as well as staff from social service, mental health, or child welfare agencies may also serve as users. The qualifications of raters and users were described in Chapter 1.

To implement the DESSA 2 mini and DESSA 2 effectively, educators need to be prepared to complete their ratings and a plan is needed for educators, student support personnel, and building leaders to review and respond to the data. This plan should also include ongoing support to educators throughout the year, as well as the logistical and technological aspects of implementation. It is imperative that building leaders plan for and communicate information about these implementation activities prior to the beginning of the school year. This coordinated plan will help ensure that schools get the most out of the data provided by these assessments. A detailed description of these activities is beyond the scope of this chapter; however, we recommend users review the guide titled *The Aperture Education Guide to Data-Driven SEL: A Multi-Tiered Systems of Support Approach to Develop Students’ Social and Emotional Competence*. This resource is available for download in the online system’s Support Portal (https://selcompass.zendesk.com/hc/en-us/article_attachments/28683129510285).

It is recommended that raters have the opportunity to observe the student for at least four weeks before completing the rating. Before completing their ratings, raters should attend the recommended professional learning courses offered by Aperture Education and familiarize

themselves with the items. In addition to these courses, the following general guidelines for completing the DESSA 2 mini and DESSA 2 should be reviewed with the rater:

- The rater should complete the DESSA 2 mini and DESSA 2 during a quiet time when there are few distractions.
- The rater should base the ratings on direct observations of the student, considering only behaviors that the rater has seen. The rater should not consider behaviors that were reported to occur in other classrooms or settings. A student's behavior, including the demonstration of social and emotional competencies, may vary in different environments and with different adults. Capturing and understanding these differences by comparing ratings provided by different raters can provide a more complete, nuanced, and accurate picture of the student's social and emotional competencies. For similar reasons, the use of group ratings in which two or more educators collaborate to provide a single rating is discouraged because it obfuscates these important contextual differences.
- The rater should consider only those behaviors that have occurred in the past four weeks. If the rater did not observe the student engaging in a behavior during that time, they should select the "Never" response.
- The rating should be based solely on the number of times the student being rated exhibited the behaviors, not how frequently the student exhibits the behavior in comparison to other students in the classroom.
- The rater is requested to answer every item. An inability to complete the items indicates that the rater has had insufficient opportunity to observe that student, and another rater should be used. If a behavior is not observed, the rater should be encouraged to answer "Never." A response to every item is required on the DESSA 2 mini. The rater may leave up to two items blank on the DESSA 2 as long as those items are on different scales.

Specific Directions for Completing the DESSA 2 Measures

The DESSA 2 mini and DESSA 2 are available through the online system; there are no print versions available. PDFs of the rating forms can be downloaded as needed to collect pencil and paper responses for entry into the online system. There are four separate, parallel** DESSA 2 mini rating forms (forms A, B, C, and D), which allows for repeated administrations and progress monitoring throughout the year. The four versions of the DESSA 2 mini do not have to be administered in sequence. The same four forms are used for all students in kindergarten through 8th grade. There is one DESSA 2 form used for all students in kindergarten through 8th grade. In non-graded programs, the DESSA 2 mini and DESSA 2 can be used with students ages 5–14 years. Specific directions for completing the ratings are provided in the next section. This information can also be found in DESSA professional learning sessions and other resources available in the online system's Support Portal.

** We have chosen to use the term "parallel" because of the similarities in means, standard deviations, standard errors of measurement, alpha reliabilities, and correlations between each mini and the full DESSA 2, as documented in Chapters 2 and 3.

Completing the Ratings

The online DESSA 2 mini and DESSA 2 forms contain the following directions to the rater:

This form describes a number of behaviors seen in some students. Read the statements that follow the phrase: ***During the past 4 weeks, how often did the student...*** and click the button 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, just click on the button for your new choice.

A “Never” response can mean that you have never observed the student engaging in that behavior or that the student does not engage in that behavior.

The 8 items that comprise the DESSA 2 mini or the 40 items that comprise the DESSA 2 are presented in a scrolling list (see [Figure 4.1](#)). The rater responds to each item by clicking the appropriate button underneath the words “Never,” “Rarely,” “Sometimes,” “Often,” or “Almost Always.” When all items have been completed, the rater is directed to submit their rating, and the rating is immediately scored by the system. To ensure the security of the system and to protect sensitive student information, ratings must be completed in one session. Partially completed ratings are not stored.

FIGURE 4.1
DESSA 2 Educator Form Presented in the Online System

Malayah Abbott		DESSA 2				
Student ID: 3464254 Minimize Instructions						
This form describes a number of behaviors seen in some students. Read the statements that follow the phrase: During the past 4 weeks, how often did the student... and click on the button 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, just click on the button for your new choice.						
A “Never” response can mean that you have not observed the student engaging in that behavior or that the student does not engage in that behavior.						
During the past 4 weeks, how often did the student...	Never	Rarely	Sometimes	Often	Almost Always	
1. show an awareness of their personal strengths?	☐	☐	☐	☒	☐	
2. prepare for school, activities, or upcoming events?	☐	☐	☐	☐	☒	
3. accept that making mistakes is part of learning?	☐	☐	☒	☐	☐	
4. get along well with different types of people?	☐	☐	☐	☒	☐	

Use of the DESSA 2 Measures with Raters Who Have Limited English Proficiency

If the rater has difficulty reading and completing the DESSA 2 mini or the DESSA 2, the items may be read to them. The person reading the items for the rater should try not to influence the ratings. The items should be read in an even, neutral tone and explanations of the items or examples should not be given. The person reading the items should also not provide any feedback or react in any way to the rater's responses.

As of the date of publication, the DESSA 2 mini and DESSA 2 are available in English and Spanish. For more detailed and updated information about translations, please contact Aperture Education.

Treatment of Missing or Blank Items

DESSA 2 mini

A response to every item on the DESSA 2 mini is required for scoring. If one or more items are left blank and the rater attempts to submit the rating, a message will appear in the online system that alerts the rater they have left too many questions blank and directs them to complete the missing items and submit again. The rater can either go back and provide the missing ratings or, if the rater truly cannot answer the items, they may click the button that indicates they are unable to rate that student and the rating will be discarded.

DESSA 2

The online system limits the number of items that can be left blank on the DESSA 2 to ensure the validity and reliability of the ratings. If one or more items are left blank and the rater attempts to submit the rating, a message will appear that alerts the rater that they have left questions blank and directs them to complete the missing items and submit again. The rater can either go back and provide the missing ratings, or, if the rater truly cannot answer the items, they may click the button that indicates they are unable to rate that student and the rating will be discarded.

The DESSA 2 will be scored if the following conditions are met:

1. There are no more than two items left blank on the entire DESSA 2.
2. There is no more than one item left blank on any individual scale (e.g., Self-Awareness, Self-Management, etc.).

If the above two conditions are met, the system will compute the mean score for the other items assigned to the same scale as the missing item, round the mean to the nearest whole number, and substitute that value for the missing item. The rater will not see the substituted value on the rating form, but it will be used in calculating the rating results.

Scoring the DESSA 2 mini

The online system automatically scores and saves the DESSA 2 mini administration as soon as the rater submits their rating. DESSA 2 mini scores are computed in the following way.

Calculating the DESSA 2 mini Sum of Item Scores

A response to every item is required to score the DESSA 2 mini. The sum of item scores is obtained by adding the raw scores for all 8 items that comprise the form using the following item raw score values: Never = 0, Rarely = 1, Sometimes = 2, Often = 3, and Almost Always = 4.

Determining the DESSA 2 mini T-Score and Percentile Rank

The sum of item scores is the DESSA 2 mini raw score. Each possible raw score is converted to a *T*-score and a corresponding percentile rank. This conversion of raw scores to derived scores is accomplished using the norms table based on the national standardization sample (see Chapter 2 for details on the standardization sample and norms creation). There is a separate norms table for each of the four DESSA 2 mini forms, which can be found in Appendix A. These four norms tables can be used for all ratings; the same norms are used for grades kindergarten through 8 and for all genders. Each DESSA 2 mini form provides one score, referred to as the Social-Emotional Total (SET) score.

Determining the DESSA 2 mini Descriptive Range for the SET

Each possible *T*-score falls into a descriptive range. High scores (*T*-scores of 60 and above) are referred to as *strengths*. This range of scores is indicated by green shading on reports in the online system. *T*-scores that fall between 41 and 59 inclusive are described as *typical* and are indicated by blue shading on reports. Low scores (*T*-scores of 40 and below) are described as a *need for instruction*. This range of scores is indicated by red shading on the reports. The interpretation and use of these scores for universal screening, monitoring progress, and evaluating program outcomes is described in the next chapter.

Scoring the DESSA 2

The online system automatically scores and saves the DESSA 2 administration as soon as the rater submits their rating. DESSA 2 scores are computed in the following way.

Calculating the DESSA 2 Scale Raw Scores

Scale raw scores for the six scales (Self-Awareness, Self-Management, Social Awareness, Relationship Skills, Responsible Decision Making, and Optimistic Thinking) are obtained by adding the raw scores for all of the items that comprise each scale using the following raw score values: Never = 0, Rarely = 1, Sometimes = 2, Often = 3, and Almost Always = 4.

Determining the DESSA 2 Scale T-Scores and Percentile Ranks

The scale raw scores are converted to *T*-scores and percentile ranks for each scale using a norms table based on the national standardization sample (see Chapter 2 for details on the standardization sample and norms creation). There is only one DESSA 2 norms table for educator raters; the same norms are used for grades kindergarten through 8 and for all genders. The DESSA 2 norms table can be found in Appendix B.

Determining the DESSA 2 T-Score and Percentile Rank for the Social-Emotional Composite

The *T*-score and percentile rank for the Social-Emotional Composite (SEC) are based on the sum of the *T*-scores of the six DESSA 2 scales. That is, the sum of the scale *T*-scores is treated as a raw score for calculating the corresponding *T*-score and percentile rank based on the national norms. This method is used to determine the standard scores for the SEC because it gives equal weight to each of the six DESSA 2 scales.

Determining the DESSA 2 Descriptive Range for Each Scale and SEC

For each scale and the SEC, high scores (*T*-scores of 60 and above) are referred to as *strengths*. This range of scores is indicated by green shading on reports. *T*-scores that fall between 41 and 59 inclusive are described as *typical* and are indicated by blue shading on reports. Low scores (*T*-scores of 40 and below) are described as a *need for instruction*. This range of scores is indicated by red shading on the reports. The interpretation and use of these scores for providing data-driven instruction, monitoring progress, and evaluating program outcomes is described in the next chapter.

Note for Researchers: Aperture Education encourages the use of the DESSA suite of assessments, including the DESSA 2 mini and DESSA 2, in research. Please contact our team at Aperture Education regarding research policies, licensing agreements, and availability of syntax for scoring DESSA research protocols.



Chapter 5

INTERPRETATION

CHAPTER 5

Interpretation



Effective interpretation of any rating scale demands that the user be familiar with what is being measured, the scores that are provided, and how these scores should be interpreted and used to improve outcomes for students.

General Interpretation Guidelines

When interpreting scores from the DESSA 2 measures, the user should always consider the following general guidelines.

First, the DESSA 2 user should have a thorough understanding of the meanings and appropriate uses of the various standard scores and descriptive ranges. Although the DESSA 2 measures meet or exceed accepted professional standards for reliability, the user needs to realize that all assessments contain some degree of measurement error that should always be considered in interpreting results and making data-based decisions.

Second, always consider the student's and family's cultural heritage and family background when interpreting DESSA 2 findings. Although we took many steps during development to avoid items that might elicit different responses from various racial and ethnic groups, cultural differences in the prevalence and meaning of specific DESSA 2 items might exist, as they would with any assessment. Therefore, the user should be sensitive to cultural differences when interpreting DESSA 2 results.

The Center for Mental Health Services of the federal Substance Abuse and Mental Health Services Administration (SAMHSA) has published Cultural Competence Standards (2000). Although these standards are more than 20 years old, they remain pertinent and useful, and are consistent with more recent standards for similar disciplines (for example, SAMHSA, 2014). Among the provider competencies, the following are particularly relevant to DESSA 2 users:

- An understanding of psychosocial stressors and traumas such as the COVID-19 pandemic, 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 student’s and family’s culture will result in more sensitive interpretations of DESSA 2 findings, and more useful recommendations to students, parents, and educators.

Third, users should appreciate that the DESSA 2 is one source of information about the social and emotional competence of students. Each set of DESSA 2 scores is based on the ratings provided by a single adult. Therefore, the scores reflect the unique interactions between the student and that adult in a particular context, often the classroom. A different rater who sees the student in a different context may well provide somewhat different ratings. Therefore, we recommend that DESSA 2 users interpret scores in light of other information (e.g., observations, discussions with the student, developmental and social histories, and results from other assessment instruments) related to the student. We also strongly recommend the evaluation of the consistency of the student’s behavior across environments, using multiple raters, both to enhance understanding and to facilitate conversation with students.

Considerations Regarding the Use of the DESSA 2 with Students with Special Needs

Although the DESSA 2 is not intended to be used as part of a special education eligibility determination, knowledge of a student’s social and emotional strengths and needs can be very helpful in informing an individual education plan (IEP) or other support plans. The DESSA 2 can provide critical information about how the student’s disability is impacting their social and emotional functioning. By identifying specific social and emotional skills that were rated in the strength range, the DESSA 2 assists IEP teams in meeting the requirements of section 300.324 of the Individuals with Disabilities Education Act (IDEA), which requires educators to consider the strengths of the student when creating the IEP. Similarly, items that were rated in the need for instruction range can be incorporated into the IEP as functional goals. Used in this way, the DESSA 2 can inform the IEP, resulting in student-specific, empirically grounded, data-driven strength and goal statements.

More specific issues regarding the interpretation of the DESSA 2 are provided in the remainder of this chapter. This will include a summary of the types of scores the measures yield, the mechanics of how these scores should be examined, and methods for their interpretation.

Types of Scores Given

Note Regarding Raw Scores

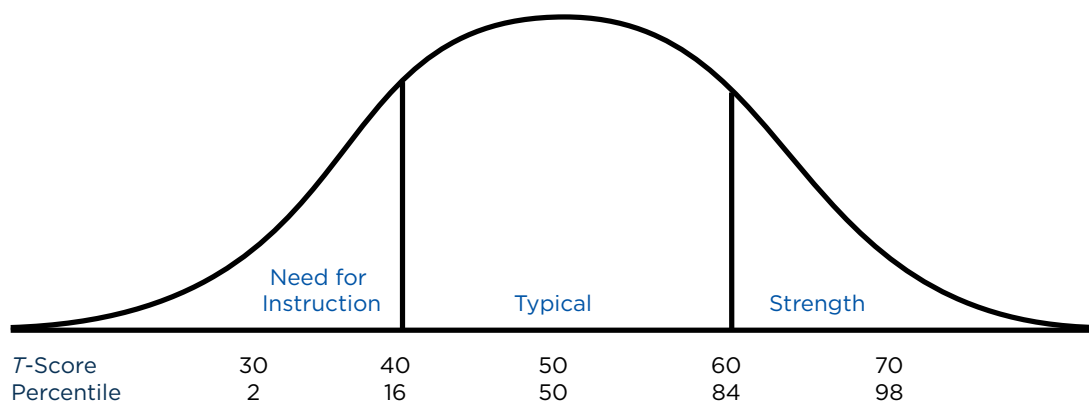
Although the online system that supports the DESSA 2 measures does not display raw scores, they are discussed here because they are the basis for determining the standard scores that are provided. Scale raw scores are determined by adding the item raw score values (Never = 0; Rarely = 1; Sometimes = 2; Often = 3; and Almost Always = 4) for all the items comprising a form (DESSA 2 mini) or a scale (DESSA 2). Because the items comprising the various DESSA 2 scales differ, raw scores cannot be directly compared and provide little information about the overall level of the student's social and emotional competencies. For instance, the Social Awareness scale has 6 items. Therefore, an average rating of "Sometimes," which has an item raw score value of 2, would result in a Scale Raw Score of 12. In contrast, an average rating of "Sometimes" on the 7-item Optimistic Thinking scale would result in a Scale Raw Score of 14.

Standard Scores

The DESSA 2 measures provide standard scores derived from the national standardization sample. On the DESSA 2, this enables scores on the six separate scales of the DESSA 2 to be directly compared. Standard scores also enable the comparison of a given student's behavior to that of the students in the standardization sample. The DESSA 2 measures provide two standard scores: *T*-scores and their corresponding percentile ranks. [Figure 5.1](#) shows the relationships between the *T*-scores, percentile ranks, the normal distribution, and the *T*-score range descriptions for the DESSA 2 measures. These standard scores and range descriptions are described below.

FIGURE 5.1

Relationship of DESSA 2 *T*-scores, Percentile Ranks, and the Normal Curve



T-Scores

Each DESSA 2 *T*-score is a standard score set to have a mean of 50 and a standard deviation of 10. Like the percentile ranks, *T*-scores are based on the raw score ratings received by the students in the standardization sample. In contrast to percentile ranks, however, DESSA 2 *T*-scores have the same meaning throughout their range. That is, the 5-point difference between the *T*-scores of 50 and 55 is equivalent to the 5-point difference between the *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 DESSA 2 results and when comparing scores earned on the various scales. On the DESSA 2 measures, *T*-scores can range from 28 to 72.

Percentile Ranks

Percentile ranks compare the student's behavior to that of other students who have been rated using the DESSA 2. The percentile rank indicates the percentage of students in the standardization sample who earned the same or lower raw score. For example, if a student earns a percentile rank of 65, that means that 65% of the students in the standardization sample earned the same or a lower raw score. DESSA 2 percentile ranks range from a minimum of 1 to a maximum of 99.

Percentile ranks 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 ranks is that the differences between the ranks do not have the same meaning across the 1–99 scale. This means that comparing two DESSA 2 scales using percentile ranks will likely mislead the practitioner to conclude that a significant difference exists when it does not. Consequently, although percentile ranks are useful for describing the relative standing of a student versus other students in the standardization sample, they should not be used to compare students' scores across the DESSA 2 scales because their meaning changes at different points on the normal distribution. It is important to remember that these ranks should **never** be averaged or used in mathematical computations. Only DESSA 2 *T*-scores should be used for that purpose.

***T*-Score Range Descriptions**

The DESSA 2 measures are strength-based assessment tools in which the items reflect positively valued social and emotional competencies; therefore, high scores are desirable. For example, when rating how often a student “keep(s) trying when unsuccessful” or “show(s) appreciation of others,” 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 provide descriptions for the *T*-score ranges, which are presented in [Table 5.1](#). This recommended language is used for both the DESSA 2 mini and DESSA 2 scores and is designed to align with the *T*-score ranges and corresponding descriptions used when interpreting results for each of the assessment tools in the DESSA assessment suite.

TABLE 5.1**Descriptive Categories and Interpretations of the DESSA 2 *T*-Scores**

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

The term “need for instruction” (or “need” for short) is used to describe DESSA 2 *T*-scores of 28 to 40 inclusive. In the reports provided by the online system, scores in the need for instruction range are color-coded as red. *T*-scores of 40 or less mean that the student was rated as showing few behaviors associated with social and emotional competence on the DESSA 2 mini, or in the particular social and emotional competency on the DESSA 2 (e.g., Self-Management). Students with scores in this range can be considered at risk for exhibiting or developing social and emotional problems (Shapiro et al., 2017). Similarly, they can be considered at promise for developing social and emotional competency in this area (LeBuffe et al., 2021). Approximately 16% of the students in the standardization sample received scores in the need for instruction range. If a student receives a score in the need for instruction range on the DESSA 2 mini, a follow-up assessment of their specific social and emotional strengths and needs should be conducted using the full DESSA 2. The assessment findings should then be used to develop an individualized plan to assist the student in developing these important skills. Within a multi-tiered system of support (MTSS) framework, these students might receive Tier 2 or Tier 3 social and emotional supports in addition to Tier 1 programming. The aligned DESSA strategies and Tier 2 intervention programs can be used for these purposes. These instructional resources can be found in the Strategies section of the online system.

Scale *T*-scores of 41 to 59 inclusive should be described as “typical” and are color-coded as blue in reports in the online system. Approximately 68% of students in the standardization sample received scores in this range. Students who receive scores in the typical range will likely benefit from universal/Tier 1 strategies designed to promote the social and emotional competence of all students, such as the DESSA strategies.

DESSA 2 scale *T*-scores of 60 to 72 inclusive should be described as “strengths” and are color-coded as green in the online system. Approximately 16% of the students in the standardization sample received scale scores in the strength range. Educators should consider and implement strategies to support, sustain, and broaden social and emotional competencies that are rated in the strength range.

The various descriptions and their relationship to DESSA 2 *T*-scores are summarized in Table 5.1. The DESSA 2 user should keep in mind that these are guidelines for the categorization and interpretation of DESSA 2 scores and should not be rigidly applied, over-interpreted, or reified. Although the DESSA 2 mini and DESSA 2 scales have high internal reliability, and consequently minimal standard errors of measurement (see Table 3.1), DESSA 2 users should take measurement error into account when interpreting DESSA 2 scores. This is particularly important when the student’s *T*-scores are close to the thresholds presented above.

The Meaning and Interpretation of the DESSA 2 Scales

The DESSA 2 mini Social-Emotional Total

This scale gives an overall indication of the student's social and emotional competence. Because administering the DESSA 2 mini is reliable and efficient, and because it characterizes the student's social and emotional competence with a single number, the DESSA 2 mini is typically administered to all students in a school or district community as a universal screener of social and emotional competence. For similar reasons, the DESSA 2 mini Social-Emotional Total (SET) score is also useful in outcome measurement and program evaluation.

The DESSA 2 Scales

The following brief descriptions are to aid in the interpretation of the DESSA 2 scales. More thorough information on the content and meaning of these scales is presented in Chapter 1.

- ***Optimistic Thinking:*** Optimistic Thinking is the belief and demonstration of confidence, hopefulness, and positive thinking regarding oneself, others, and one's life situations in the past, present, and future.
- ***Self-Awareness:*** Self-Awareness is the ability to understand emotions, thoughts, and values and how they influence one's behavior; recognize strengths and limitations; and develop healthy identities and a sense of purpose.
- ***Self-Management:*** Self-Management is the ability to manage emotions and behaviors across different situations and environments and to demonstrate agency as one works to set and achieve personal and collective goals.
- ***Social Awareness:*** Social Awareness is the understanding of social norms for behavior; the ability to empathize with, respect, and take the perspectives of others; and the feeling of connection and belonging with family, peers, schools, and community groups.
- ***Relationship Skills:*** Relationship Skills are the abilities to establish and maintain healthy and positive relationships, including effective communication, collaborative problem solving, negotiating conflict, and demonstrating helpful and supportive behaviors.
- ***Responsible Decision Making:*** Responsible Decision Making is the ability to make careful, reliable, and constructive choices about personal and social behavior that are appropriate across diverse situations; to consider the personal, social, and collective impact of one's actions; and to demonstrate curiosity and open-mindedness to learning.

The DESSA 2 Social-Emotional Composite

This scale gives an overall indication of the student's social and emotional competence. It is the most reliable and valid overall indicator within the DESSA 2. Because it characterizes the student's social and emotional competence with a single number, the Social-Emotional Composite (SEC) is particularly useful in outcome measurement and program evaluation.

Basic Interpretation of the DESSA 2 mini

Examine the SET *T*-score and note the corresponding range description (i.e., strength, typical, need for instruction). This score provides an indication of a student's overall social and emotional functioning. If a student receives a score in the need for instruction range on the DESSA 2 mini, a follow-up assessment of their specific social and emotional strengths and needs should be conducted using the full DESSA 2.

Basic Interpretation of the DESSA 2

Interpretation of the DESSA 2 results proceeds in a stepwise fashion from the most general indicator of the student's social and emotional status to increasingly more specific information. This process should include the following three steps:

Step 1: The Social-Emotional Composite

First, examine the SEC *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 student's social and emotional competence. The SEC *T*-score is a highly reliable indicator of the student's overall social and emotional functioning and serves as the starting point in interpreting the DESSA 2. The score a student receives on the SEC also provides a frame of reference for the remaining interpretative steps.

Step 2: Examining Scale Scores

Next, examine the six separate DESSA 2 scales, and note the *T*-scores and corresponding strength, typical, and need for instruction ranges. Examination of the separate DESSA 2 scale *T*-scores provides useful information about the student's specific social and emotional competencies. For instance, the scores can suggest whether a student's strengths or needs are primarily intrapersonal (as evidenced by high or low scores on the Optimistic Thinking, Self-Awareness, and Self-Management scales) or interpersonal (as shown by high or low scores on Social Awareness and Relationship Skills). Examination of the DESSA 2 Individual Student Rating Report is particularly useful at this step, as the visual depiction of the scale scores can make patterns easier to discern. [Figure 5.2](#) provides a sample Individual Student Rating Report as presented in the online system.

Step 3: Identifying Specific Strengths and Need for Instruction Items

Each of the six DESSA 2 scales represents a group of items that relate to a common social and emotional competency (e.g., Self-Management). However, these competencies are broad categories that encompass varying and more specific social and emotional skill sets. For example, a student with a need for instruction on the Self-Management scale may have difficulties

managing their emotions and behaviors across different situations (e.g., item # 5, stay focused despite a distraction; item # 37, stay calm when faced with a challenge) or in setting and achieving a goal (e.g., item #30, persist to achieve a goal; item #39, contribute to group or team goals).

Step 3 enables the DESSA 2 user to move beyond scale scores to gain an understanding of the specific behaviors that are strengths (i.e., in the student’s behavioral repertoire) or needs for instruction (i.e., not yet acquired) for the student.

Identification of specific behavioral strengths and needs for instruction involves a method called Individual Item Analysis. Any item can represent a need for instruction if the rating the student received is substantially lower than the rating given to students who have typical scores. That is, an individual item is considered to indicate a need for instruction if the score the student received is at least one standard deviation below the mean for that item in the national standardization sample. Less than 16% of the students in the standardization sample received scores in the need for instruction range on each item on the DESSA 2. Such a score on an individual item indicates that the rater has reported that the student is not yet demonstrating this behavior in the rater’s presence to the extent considered typical in other students. Individual items rated in the need for instruction range should be considered as targets for social and emotional instruction.

Similarly, any item can represent a strength if the rating is substantially higher (at least one standard deviation above the national mean) than that given to students with typical scores. For each item, no more than 16% of the students in the national standardization sample received ratings in the strength range. DESSA 2 users should consider how these focal strengths can be leveraged or built upon in a support plan. Students should be given many opportunities to demonstrate and reinforce their strengths. The item score values associated with the need and strength ranges are found in [Table 5.2](#).

FIGURE 5.2
A Sample DESSA 2 Individual Student Rating Report

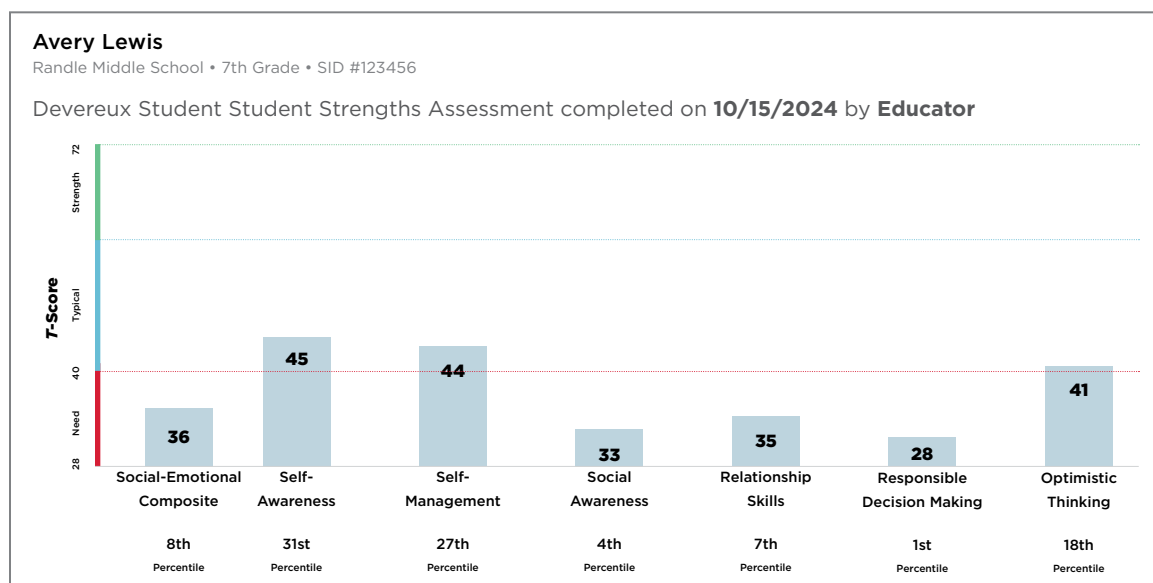


TABLE 5.2**Individual Item Analysis Values for the DESSA 2**

Item Number	Need for Instruction	Typical	Strength
1	0, 1	2, 3	4
2	0, 1	2, 3	4
3	0, 1	2, 3	4
4	0, 1	2, 3	4
5	0, 1	2, 3	4
6	0, 1	2, 3	4
7	0, 1	2, 3	4
8	0, 1	2, 3	4
9	0, 1	2, 3	4
10	0, 1	2, 3	4
11	0, 1	2, 3	4
12	0, 1	2, 3	4
13	0, 1	2, 3	4
14	0, 1	2, 3	4
15	0	1, 2, 3	4
16	0, 1	2, 3	4
17	0, 1	2, 3	4
18	0, 1	2, 3	4
19	0, 1	2, 3	4
20	0, 1	2, 3	4
21	0, 1	2, 3	4
22	0, 1	2, 3	4
23	0, 1	2, 3	4
24	0, 1	2, 3	4
25	0, 1	2, 3	4
26	0, 1	2, 3	4
27	0, 1	2, 3	4
28	0, 1	2, 3	4
29	0, 1	2, 3	4

(continued)

Item Number	Need for Instruction	Typical	Strength
30	0, 1	2, 3	4
31	0, 1	2, 3	4
32	0, 1	2, 3	4
33	0, 1	2, 3	4
34	0, 1	2, 3	4
35	0, 1	2, 3	4
36	0, 1	2, 3	4
37	0, 1	2, 3	4
38	0, 1	2, 3	4
39	0, 1	2, 3	4
40	0, 1	2, 3	4

The primary advantage of this method is that it allows for the identification of specific behaviors that can be leveraged (strengths) or acquired (needs for instruction) by specific interventions. Individual item identification facilitates the development of support plans that are individualized and behaviorally grounded. For instance, if the student’s rating on item #2, “prepare for school, activities, or upcoming events,” was in the need for instruction range, then developing or improving planning skills can become a goal, and each component skill (e.g., creating a calendar, task analyzing larger activities) can become an objective on the support plan. Conversely, if item #17, “encourage others,” is a strength for the student, then involving this individual as a leader in a peer group would be an appropriate way of supporting and further developing this desired behavior. The identification of specific strengths and needs is an important step in linking DESSA 2 assessment results to social and emotional strategies and tiered interventions.

Another advantage of the Individual Item Analysis method is that it allows the DESSA 2 user to identify specific needs for instruction even if the student’s scale scores are not in the need for instruction range. That is, even though a scale score may be in the typical or even strength range, examination of the individual items may identify specific behaviors that were rated in the need for instruction range. These specific skills can then be taught resulting in a more complete repertoire of social and emotional skills. This approach is particularly important for schools and programs that are committed to thriving; that is, maximizing the social and emotional competence of each student.

In the online system, the results of the individual item analysis are available on the Individual Student Rating Report. The DESSA 2 user has the option of viewing the item-level results for an individual competency or all six competencies. Within each competency, the item-level results are sorted by their descriptive range so that all the strengths, typical ratings, and needs for instruction are presented together. [Figure 5.3](#) provides an example of this functionality.

FIGURE 5.3**Item Level Identification as Shown on the Individual Student Rating Report**

Individual Item Analysis		Self-Management ▾	
Competency	Item	Response	Category
Self-Management	persist to achieve a goal?	Almost Always	Strength
Self-Management	stay calm when faced with a challenge?	Sometimes	Typical
Self-Management	contribute to group or team goals?	Sometimes	Typical
Self-Management	stay focused despite a distraction?	Rarely	Need
Self-Management	keep trying when unsuccessful?	Rarely	Need
Self-Management	think before they acted?	Rarely	Need
Self-Management	work carefully on projects or schoolwork?	Never	Need

Advanced Interpretation of the DESSA 2 Measures

Progress Monitoring

Progress monitoring is a key component of the response to intervention (RTI) framework. The goal of progress monitoring is to determine if the interventions (in the case of the DESSA 2 measures, social and emotional skill instruction) are being effective in enhancing the student's social and emotional competence by comparing scores on successive assessments. Rather than waiting until the end of the year to determine if growth has occurred, progress monitoring provides opportunities throughout the school year to evaluate growth and make any indicated changes to improve end-of-year outcomes. If the goal is to improve overall social and emotional competence, the use of the eight-item DESSA 2 mini is recommended because of its brevity. However, the DESSA 2 mini yields only one score, the Social-Emotional Total (SET), which is a measure of overall social and emotional competence. Consequently, if the question of interest is improvement in one or more specific social and emotional competencies, the full DESSA 2 should be used.

To evaluate progress, the administrations of the DESSA 2 measures must be separated by at least four weeks so that the second administration is based on a different sample of behaviors. To allow for sufficient time for social and emotional skill instruction, six to eight weeks is recommended between administrations. Many school districts and OST programs have adopted the practice of monitoring progress one or two times during a school year. A typical schedule might be the initial DESSA 2 administration in October. First progress monitoring prior to the holiday break in December. Second progress monitoring in early March, followed by an end-of-year summative assessment in late May or June.

Cohen's *d*-ratio, which was introduced in Chapter 2, is used to evaluate the progress made between successive administrations. Using the *T*-scores on the scale(s) of interest, the pretest or earlier administration scale score is subtracted from the posttest or more recent administration. If the student's score has increased (i.e., shown progress or growth) the resulting difference will be positive. Cohen (1988) suggested that *d*-ratios of 0.2, 0.5, and 0.8 be considered

TABLE 5.3
Interpretation and Guidance for Progress Monitoring

Magnitude of the Difference	Standard Deviation Unit	T-score Units	Guidance
Negligible/None	Less than 0.20	Less than 2	Supports are ineffective; try new supports and strategies. Consult with student assistance personnel.
Small	0.20 to 0.49	2 to 4, inclusive	Supports are minimally effective. Increase frequency, duration, or intensity, or try new strategies. If using only group interventions/supports, consider individualized supports.
Medium	0.50 to 0.70	5 to 7, inclusive	Supports are moderately effective. Consider enhancing if resources, including time and personnel, permit.
Large	Greater than or equal to 0.80	8 or higher	Supports are working well. Continue current plans.

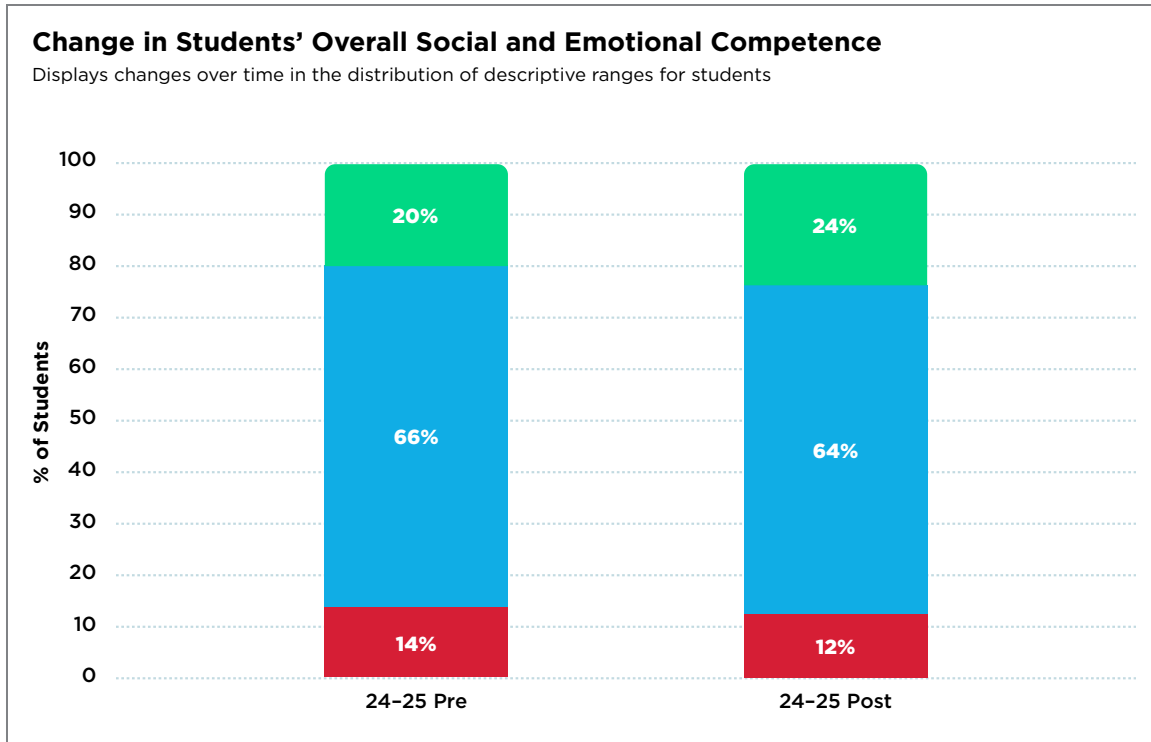
small, medium, and large changes respectively. Because *T*-scores have a standard deviation of 10, these ranges are equivalent to 2–4, 5–7, and 8 or more *T*-score units (changes of 0 or 1 *T*-score unit are considered to be “negligible”). As shown in [Table 5.3](#), DESSA 2 users can modify their social and emotional instruction (e.g., supplementing universal instruction with small group targeted supports) based on the degree of progress shown by the student. The thoughtful use of this progress monitoring technique can result in better end-of-year outcomes. Progress monitoring data and interpretation guidance is provided to educators and administrators in the online system.

Evaluating Programmatic Outcomes and Impact

Whereas the progress monitoring technique previously described is a formative evaluation approach with a goal of improving individual student outcomes, the information in this section describes a summative evaluation approach designed to assess program effectiveness, evaluate impact, and inform continuous quality improvement (CQI) efforts for groups of students. Like progress monitoring, summative evaluation involves comparing changes in scores over time but typically is used to compare the first or beginning-of-year rating with the last or end-of-year rating, with an intervention implemented in between.

The Impact Report in the online system is designed to facilitate outcome evaluation with the DESSA 2 measures. It allows users to compare the progress of students from one rating to the next in the three *T*-score range descriptions of strength, typical, and need for instruction categories (see [Figure 5.4](#)). The Student Movement component of the Impact Report provides specific information on how many students from a given descriptive category (e.g., need for instruction) moved to a different category (e.g., typical, strength) between ratings (see [Figure 5.5](#)). Users may also run the Impact Report by student population (e.g., race/ethnicity, gender, special populations) and review results of disaggregated data across subgroups of students. Data from the Impact Report can be exported from the online system to enable users to conduct statistical analysis and compare to other district- or school-collected data such as academic achievement or behavioral data.

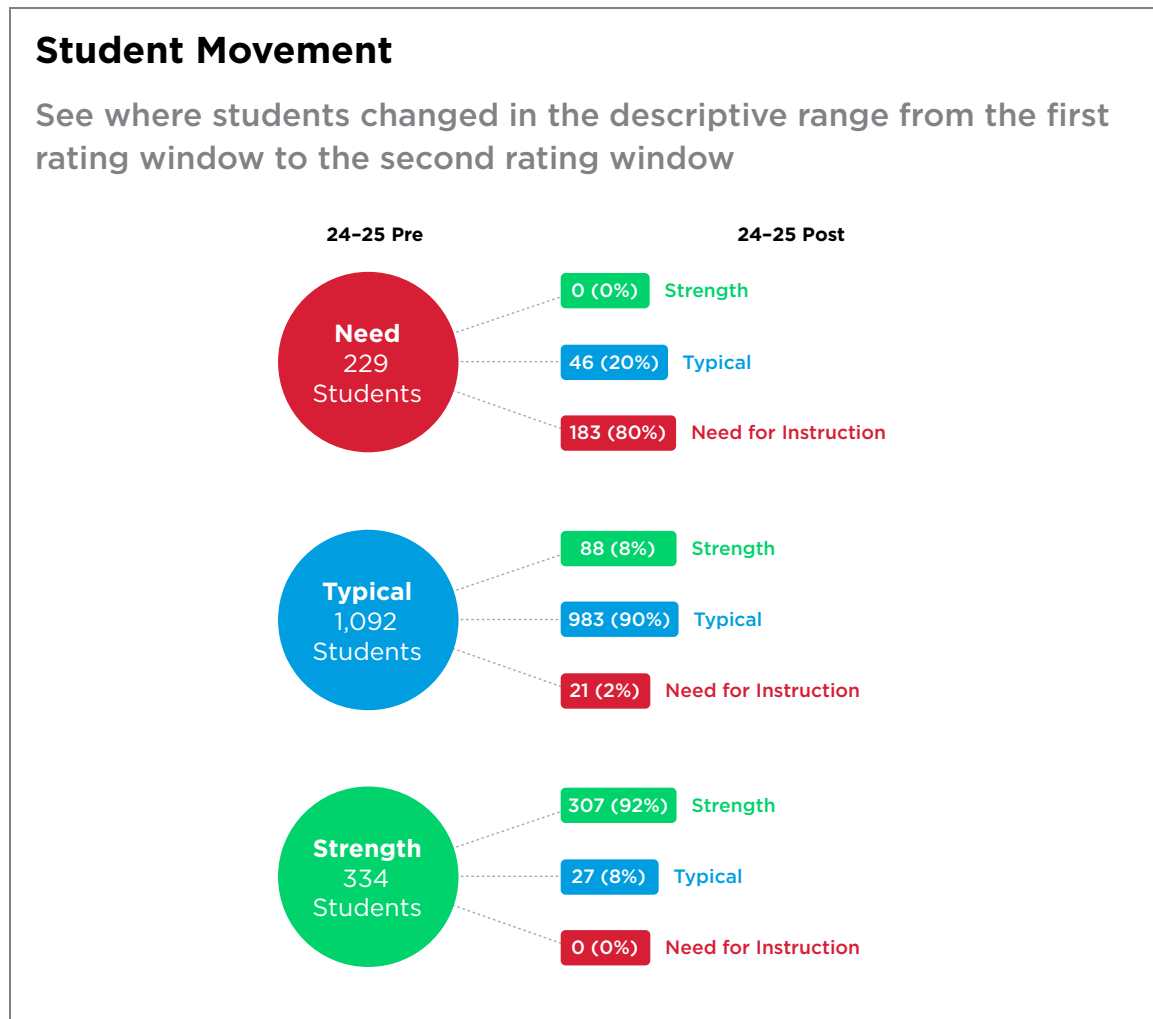
FIGURE 5.4
Sample Impact Report for DESSA 2 Data



Outcome evaluation as applied to helping students develop social and emotional competencies is a flexible and powerful tool. This approach enables the DESSA 2 user to look at the effectiveness of interventions on a scale-by-scale basis and across groups of students. By using this method, we can determine which students benefitted from which interventions in which areas. This student-specific information is especially useful for quality improvement efforts. By aggregating findings across students, classrooms, schools, etc., schools and OST programs can determine the relative impact of their social and emotional instructional efforts on differing social and emotional competencies. For example, aggregated data might show more improvement and better outcomes in the area of self-management as compared to relationship skills. Similarly, this approach can explore different social and emotional outcomes for different groups of students. For example, the data might show that students in the sixth grade are showing more growth than those in the eighth grade. The approach provides valuable data on student outcomes that can inform both program evaluation/CQI efforts as well as efforts to promote educational equity.

FIGURE 5.5

Sample Student Movement Report for DESSA 2 Data



Determining the impact of social and emotional strategies and curricula at the individual student and group levels is essential to continuously improving professional practice, advancing the field, and most importantly, improving outcomes for students. Examining outcomes at the individual student level and using this information to adjust or modify instruction to ensure that each student acquires a full repertoire of social and emotional skills is essential to efforts to promote educational equity; it lies at the heart of data-driven social and emotional skill instruction.

Interpretation Example

The following example illustrates the interpretation of the DESSA 2 and how results facilitate intervention planning. This example concerns a student in the fifth grade, Mia. Mia does well academically, excelling in math and science. The elementary school Mia attends uses the DESSA 2 mini to universally screen students' social and emotional competence at the beginning of the school year. Classroom teachers then do a full DESSA 2 assessment on any student

who receives a score in the need for instruction range (*T*-scores of 40 and below). Mia receives a DESSA 2 mini score of 38, which falls just in the need for instruction range. Her classroom teacher, Ms. Wallace, then completes a DESSA 2 assessment.

Step 1: Examination of the Social-Emotional Composite

Ms. Wallace began by examining the SEC score on the DESSA 2 Individual Student Rating Report accessible in the online system. She noted that Mia received a *T*-score of 38, consistent with the SET score on the DESSA 2 mini, and corresponding to a percentile rank of 12, placing her in the higher end of the need for instruction range. This score was not surprising to Ms. Wallace, as she had noticed Mia to be struggling socially in the classroom and had concerns that Mia's social and emotional skills may not be commensurate with her academic performance.

Step 2: Examining Scale Scores

Although the SEC score was in the need for instruction range, an examination of the six scale scores did show variability across the domains. Ms. Wallace began by noting Mia's strength in Responsible Decision Making. She also noted that, consistent with her concerns, Mia was exhibiting a need for instruction in key interpersonal areas including Social Awareness and Relationship Skills. She was surprised, however, to note that Mia was also exhibiting a need for instruction in Optimistic Thinking. The remaining two scales (Self-Awareness and Self-Management) were rated in the typical range.

Step 3: Individual Item Analysis

Although the review of scale scores in step 2 was very helpful in confirming Ms. Wallace's concerns, identifying additional needs for instruction, and making her more aware of Mia's strengths, she was still somewhat at a loss of how to help Mia acquire the critical skills that were not yet in her repertoire. To gain a better understanding of what specific skills Mia would benefit from learning, Ms. Wallace reviewed the individual item analyses presented on the Individual Student Rating Report. Given Mia's lowest score was in Relationship Skills, Ms. Wallace decided to focus her efforts there. A review of the items on this scale that were rated in the need range suggested three behaviors to concentrate her efforts: items #4, "get along well with different types of people," #17, "encourage others," and #35, "cooperate with others to solve a problem."

Wanting to both honor and leverage Mia's strengths, Ms. Wallace next looked at the items on the Responsible Decision Making scale, noting that Mia "prepare(s) for school, activities, or upcoming events" (item #2) and "gather(s) information before making an important decision" (item #18). She then consulted the DESSA strategies in the online system and chose two strategies that would address Mia's needs while leveraging her strengths in the classroom. She will plan to implement these strategies with Mia, in addition to the universal social and emotional skill instruction lessons she implements with all her fifth-grade students, and will plan to review Mia's mid-year progress monitoring results to see whether the instruction has been effective in promoting Mia's relationship skills.

Use of the DESSA 2 Measures within a Multi-Tiered System of Support (MTSS)

The use of the DESSA 2 measures is not limited to the MTSS framework; however, the widespread adoption of MTSS provides a familiar and useful frame of reference for discussing the most common applications of the DESSA 2 measures.^{***} The DESSA assessment suite and their applications at the three tiers of the MTSS framework are presented below.

Use of the DESSA 2 Measures at Tier 1

Tier 1 or *universal* services and supports are provided to all students in a school or OST program. They provide the common foundation for effective social and emotional skill instruction. Programs utilize the educator-completed DESSA 2 mini (or the DESSA-HSE mini for grades 9–12) as a universal screener of social and emotional competence at Tier 1. These screeners consist of four equivalent eight-item forms and take the educator about one minute to complete per student. The DESSA 2 mini has the advantage of brevity, but it yields only one score: the Social-Emotional Total (SET) that provides a measure of overall social and emotional competence. The results are used to identify those students whose overall social and emotional competence is in the need for instruction range and who would benefit from a full educator-completed assessment with the DESSA 2 (or DESSA-HSE). However, some programs have opted to use the full DESSA educator form at the universal level because of the rich information it provides on specific social and emotional competency domains (i.e., Self-Management, Relationship Skills). For these programs, this deeper understanding of each student’s social and emotional strengths and needs across the domains justifies the added time and effort of teachers. At the secondary level, some programs use the student-completed DESSA-HSE SSR or the DESSA-MSE SSR (collectively referred to as the DESSA Student Self-Reports) as a universal assessment. These assessments provide the student an opportunity to report on their own social and emotional competencies.

For programs using either the full DESSA or the DESSA Student Self-Reports, the Competencies Report, available through the online system, is a highly informative and useful report. This report enables the educator to identify the most common strengths and needs for instruction presented by the students in the group. The most commonly occurring needs for instruction can then be addressed through the universal DESSA strategies, which are aligned to the specific social and emotional competency and are available through the online system.

In addition to adult-led planning and instruction, programs using the DESSA Student Self-Reports universally enable all students to identify personal goals and corresponding social and emotional instructional strategies that they can implement on their own. This provides students with a voice and choice in their own social and emotional growth and engages them as active participants in the process.

^{***} Readers who are unfamiliar with the MTSS framework may want to visit the website of the Center on PBIS (Positive Behavioral Interventions and Supports) at <https://www.pbis.org>.

Many schools and programs use the DESSA measures to support their implementation of universal, evidence-based social and emotional curricula, adjusting their delivery of the curriculum based on results. For example, universal strategies can supplement the lesson plans, or the most common needs for instruction can suggest areas that could be emphasized through extension activities or repetition throughout the school year.

Educators may also want to do additional skills checks or knowledge assessments with students demonstrating a need for instruction in a given area to ensure that they are acquiring the skills. Both the Collaborative for Academic, Social, and Emotional Learning (<https://pg.casel.org/review-programs/>) and the Blueprints Program for Healthy Youth Development (<https://www.blueprintsprograms.org/program-search/>) provide searchable listings of evidence-based social and emotional skill development programs.

It is important to recognize that social and emotional skill development occurs in contexts such as a classroom, school, or OST program. This context can influence not only the demonstration of a student's social and emotional skills but also the effectiveness of instruction. Consequently, many programs incorporate school climate and culture surveys as part of their initiatives. Information about school climate and culture can be used in conjunction with the *Foundational Practices*, universal strategies found in the online system that are intended to create a classroom culture and climate that will support student's development and well-being. Whereas the DESSA strategies are aligned to a specific social and emotional competency, the foundational practices are nonspecific and can be implemented immediately at the beginning of the school year. They can also be reinforced and sustained throughout the year.

Use of the DESSA 2 Measures at Tier 2

As mentioned above, most programs use the DESSA Student Self-Reports and/or the educator-completed DESSA 2 mini as universal measures of social and emotional competence. For programs using the DESSA 2 mini, those students whose SET score indicates a need for instruction are then assessed with the full DESSA 2 to identify the specific social and emotional competencies that are not yet being demonstrated to a sufficient degree. These students, as well as students receiving SEC scores in the need for instruction range on the DESSA Student Self-Reports, then may receive Tier 2 or *targeted* supports that supplement the Tier 1 universal social and emotional instruction. Some programs will use the Competencies Report to create small groups of students with similar needs and then utilize the strategies and intervention programs provided in the online system (Adams, 2013). Periodic re-administration of the DESSA2 Student Self-Reports, the DESSA 2, or the DESSA 2 mini is then used to monitor the progress of these students in enhancing their social and emotional competence.

Use of the DESSA 2 Measures at Tier 3

Tier 3 or *indicated* supports and services are provided to those students who have not sufficiently benefited from Tier 1 and Tier 2 services. Tier 3 supports and services are typically intensive and individualized. The Individual Item Analysis technique described above is particularly useful at this stage. The DESSA 2 Individual Student Report identifies those specific items that were rated as strengths for students as well as those rated as indicating a need for

instruction. This information can be used to create highly individualized and data-based plans to reinforce and leverage the student's strengths while addressing their specific needs for instruction. The online system provides strategies and intervention programs that are aligned to the DESSA 2 measures.

It is important to note that at all three tiers, we recommend that the DESSA 2 assessments be used as a formative assessment. That is, assessment data is collected during the school or program year with the goal of better understanding students' strengths and needs so that instruction can be differentiated and improved leading to better outcomes. Our goal is not to categorize or label students based on DESSA 2 scores. Rather, our purpose is to understand better the unique constellation of social and emotional strengths and needs for instruction presented by individual students, classrooms, schools, districts, and OST programs so that social and emotional instruction can be differentiated, progress monitored, and outcomes enhanced. Although the DESSA 2 assessments can also be used as a summative assessment to evaluate programmatic outcomes and inform continuous quality improvement, our primary objective is ensuring that each student has a full complement of social and emotional skills to achieve success in school and in life after graduation.

The authors would like to thank our many colleagues and DESSA clients who have shared their challenges and successes with us since the publication of the DESSA for grades K–8 in 2009. Their feedback has deepened our understanding and led to many improvements in the DESSA assessments and the online system. We hope that you will continue to share thoughts, suggestions, and experiences with us. We can be reached through Aperture Education's website (www.ApertureEd.com).

Appendices



APPENDIX C

List of Data Collection Sites by State

With deep appreciation, we would like to acknowledge the educators and staff from the following schools, out-of-school time programs, and community organizations who participated in the development of the DESSA 2.

CALIFORNIA

California Montessori Project, Carmichael
Clovis Elementary School, Clovis
Dolores Huerta Elementary School, San Francisco

Harriet Tubman Village Charter School,
San Diego

COLORADO

Baker Elementary School, Fort Morgan

Trailside Academy, Denver

CONNECTICUT

Dr. Martin Luther King, Jr. Middle School
(RISE Program), Hartford
East Hartford Middle School (RISE PROGRAM),
East Hartford

Dr. Michael Fox Elementary School (VMLK),
Hartford

DISTRICT OF COLUMBIA

Inspired Teaching Demonstration School,
District of Columbia

FLORIDA

Academy Prep, Tampa

IDAHO

Future Public School, Garden City

ILLINOIS

Jefferson Elementary School, Morton
Lovington Grade School, Lovington

KENTUCKY

Youth Leadership Development, Alexandria

MAINE

Lewiston Middle School, Lewiston

MICHIGAN

Elementary Literacy Program
(Voces in Battle Creek Michigan),
Valley View

MISSOURI

Monroe Elementary School, St. Charles

NEBRASKA

Omaha Nation Public School, Macy

NEW YORK

Westminster Community Charter School, Buffalo

NORTH CAROLINA

D.C. Virgo Preparatory Academy, Wilmington	Topsail Annandale Elementary School,
Maureen Joy Charter School, Durham	Hampstead
The Expedition School, Hillsborough	Voyager Academy Middle School, Durham

OREGON

Luckiamute Valley Charter School, Dallas

PENNSYLVANIA

Pan American Academy Charter School,
Philadelphia

RHODE ISLAND

Exeter West Greenwich Junior High School,
West Greenwich

SOUTH CAROLINA

Butler Academy, Hartsville

TEXAS

Bear Creek Intermediate School, Keller	David E. Smith Elementary in Birdville
Blazier Elementary School, Austin	Independent School District, Haltom City
Blazier Intermediate School, Austin	Enge - Washington Intermediate School,
Caprock Elementary School, Fort Worth	Groesbeck

WASHINGTON

Grand Mound Elementary School, Rochester	Rochester Primary School, Rochester
Rochester Middle School, Rochester	



References

References

- Adams, D. (2013). The application of social-emotional learning principles to a special education environment. *KEDI Journal of Educational Policy*, Special Issue (2013), 108–118.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.
- Atlas, J. A. (2010). Test review of the Devereux student strengths assessment. In R. A. Spies, J. F. Carlson, & K. F. Geisinger (Eds.), *The eighteenth mental measurements yearbook*, (pp. 178–180). Lincoln, NE: Buros Center for Testing.
- Belfield, C., Bowden, A., Klapp, A., Levin, H., Shand, R., & Zander, S. (2015). The economic value of social and emotional learning. *Journal of Benefit-Cost Analysis*, 6(3), 508–544.
<https://doi.org/10.1017/bca.2015.55>
- Bracken, B. A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment*, 5, 313–326.
<https://doi.org/10.1177/073428298700500402>
- Cai, L. (2022). flexMIRT® version 3.65: Flexible multilevel multidimensional item analysis and test scoring [Computer software]. Chapel Hill, NC: Vector Psychometric Group.
- Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (2002). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *Prevention & Treatment*, 5(1), Article 15.
<https://doi.org/10.1037/1522-3736.5.1.515a>
- Centers for Disease Control and Prevention. (2021). *Youth Risk Behavior Survey Data Summary and Trends Report: 2011–2021*.
https://www.cdc.gov/yrbs/dstr/pdf/YRBS_Data-Summary-Trends_Report2023_508.pdf
- Center for Mental Health Services. (2000). *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.

- Center for Public Education. (2016, January). *Educational equity: What does it mean? How do we know when we reach it?* National School Boards Association.
<https://www.nsba.org/-/media/NSBA/File/cpe-educational-equity-research-brief-january-2016.pdf>
- Chain, J., Shapiro, V. B., LeBuffe, P. A., Bryson, A. M., & American Indian and Alaska Native Advisory Committee. (2017). Academic achievement of American Indian and Alaska Native students: Does social emotional competence reduce the impact of poverty? *American Indian and Alaska Native Mental Health Research*, 24(1), 1–30.
<https://doi.org/10.5820/aian.2401.2017.1>
- Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48(6), 1–29. doi.org/10.18637/jss.v048.i06
- Chatterjee Singh, N., & Duraiappah, A. K. (Eds.). (2020). *Rethinking learning: A review of social and emotional learning frameworks for education systems*. New Delhi: UNESCO MGIEP.
- Ciarrochi, J., Parker, P., Kashdan, T. B., Heaven, P. C. L., & Barkus, E. (2015). Hope and emotional well-being: A six-year study to distinguish antecedents, correlates, and consequences. *The Journal of Positive Psychology*, 10(6), 520–532.
<https://doi.org/10.1080/17439760.2015.1015154>
- Cohen, J. (1988). *Statistical power analysis in the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- Collaborative for Academic, Social, and Emotional Learning. (2020). *CASEL's SEL framework: What are the core competence areas and where are they promoted?* CASEL.
<https://casel.org/casel-sel-framework-11-2020/>
- Collaborative for Academic, Social, and Emotional Learning. (2021). *SEL policy at the state level*. CASEL <https://casel.org/systemic-implementation/sel-policy-at-the-state-level/>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334. <https://doi.org/10.1007/BF02310555>
- Denham, S. A., Ji, P., & Hamre, B. (2010). *Compendium of preschool through elementary school social emotional learning and associated assessment measures*. Collaborative for Academic, Social, and Emotional Learning and Social and Emotional Learning Research Group, University of Illinois at Chicago.
- Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, 44(4), 237–251. <https://doi.org/10.3102/0013189X15584327>
- Durlak, J. A., Mahoney, J. L., & Boyle, A. E. (2022). What we know, and what we need to find out about universal, school-based social and emotional learning programs for children and adolescents: A review of meta-analyses and directions for future research. *Psychological Bulletin*, 148(11–12), 765–782.
<https://psycnet.apa.org/doiLanding?doi=10.1037%2Fbul0000383>
- Elliott, S. N., Frey, J. R., & Davies, M. (2015). Systems for assessing and improving students' social skills to achieve academic competence. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 301–319). Lynton, U.K.: Guilford Press.

- Garnezy, N. (1985). Stress-resistant children: The search for protective factors. In J. E. Stevenson (Ed.), *Journal of Child Psychology and Psychiatry Book Supplement, No. 4*. (pp. 213–233). Oxford, UK: Pergamon Press.
- Gullotta, T. P. (2015). After-school programming and SEL. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 260–281). New York: Guilford Press.
- Haggerty, K., Elgin, J., & Woolley, A. (2011). *Social-emotional learning assessment measures for middle school youth*. Social Development Research Group, University of Washington: Commissioned by the Raikes Foundation.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Hwang, S. H. J., Shapiro, V. B., Eldeeb, N., Lee, J., Robitaille, J. L., Naglieri, J. A. (2023). Assessing social and emotional competencies in educational settings: Supporting resilience in young people. In S. Goldstein & R. B. Brooks (Eds.) *Handbook of Resilience in children* (pp. 237–249). New York: Springer Science + Business Media.
https://doi.org/10.1007/978-3-031-14728-9_12
- Jagers, R. J., Rivas-Drake, D., & Borowski, T. (2018). *Equity and social-emotional learning: A cultural analysis*. CASEL Assessment Work Group Brief series.
<https://casel.org/uploads/sites/3/2019/02/Equity-Social-and-Emotional-Learning-A-Cultural-Analysis.pdf>
- Kim, B. K. E., Oesterle, S., Catalano, R. F., & Hawkins, J. D. (2015). Change in protective factors across adolescent development. *Journal of Applied Developmental Psychology*, 40, 26–37. <https://doi.org/10.1016/j.appdev.2015.04.006>
- Lebrun-Harris, L. A., Ghandour, R. M., Kogan, M. D., Warren, M. D. (2022). Five-year trends in US children’s health and well-being, 2016–2020. *JAMA Pediatrics*, 176(7): e220056.
doi.org/10.1001/jamapediatrics.2022.0056
- LeBuffe, P. A., Acosta-Price, O., Robitaille, J. L., & Doerr, E. (2021). *To both promote and prevent: The importance and practicality of strength-based screening*. Aperture Education.
- LeBuffe, P. A., & Shapiro, V. B. (2004). Lending “strength” to the assessment of preschool social-emotional health. *The California School Psychologist*, 9, 51–61.
<https://doi.org/10.1007/BF03340907>
- LeBuffe, P. A., Shapiro, V. B., & Naglieri, J. A. (2009/2014). *The Devereux Student Strengths Assessment (DESSA): Assessment, technical manual, and user’s guide*. Aperture Education.
- LeBuffe, P. A., Shapiro, V. B., & Robitaille, J. L. (2018). The Devereux Student Strengths Assessment (DESSA) Comprehensive System: Screening, assessing, planning, and monitoring. *Journal of Applied Developmental Psychology*, 55, 62–70.
<https://doi.org/10.1016/j.appdev.2017.05.002>
- Lee, J., Shapiro, V. B., & Kim, B. K. E., (2023a). Universal school-based social and emotional learning (SEL) for diverse student subgroups: Implications for enhancing equity through SEL. *Prevention Science*, 24, 1011–1022. <https://doi.org/10.1007/s11121-023-01552-y>

- Lee, J., Shapiro, V. B., Robitaille, J. L., & LeBuffe, P. A. (2023b). Measuring the development of social-emotional competence using behavioral rating scales in the context of school-based social and emotional learning. *Social and Emotional Learning: Research, Practice, and Policy*, 2, 100015. <https://doi.org/10.1016/j.sel.2023.100015>
- Lee, J., Shapiro, V. B., Robitaille, J. L., & LeBuffe, P. A. (2024). Gender, racial-ethnic, and socio-economic disparities in the development of social-emotional competence among elementary school students. *Journal of School Psychology*, 104, 101311. <https://doi.org/10.1016/j.jsp.2024.101311>
- Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Mahoney, J. L., Durlak, J. A., & Weissberg, R. P. (2018). An update on social and emotional learning outcome research. *Phi Delta Kappan*, 100, 18–23. <https://kappanonline.org/social-emotional-learning-outcome-research-mahoney-durlak-weissberg/>
- Mahoney, J. L., LeBuffe, P. A., Shapiro, V. B., Robitaille, J. L., Johnson, E. S., & Adamson, J. L. (2022). *What is assessment bias and how is Aperture Education working to reduce it?* Aperture Education.
- Malcomb, K. K. (2010). Test review of the Devereux Student Strengths Assessment. In R. A. Spies, J. F. Carlson, & K. F. Geisinger (Eds.), *The eighteenth mental measurements year-book* (pp. 180–182). Lincoln, NE: Buros Center for Testing.
- Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist*, 56(3), 227–238. <https://doi.org/10.1037/0003-066X.56.3.227>
- Masten, A. S. (2014). *Ordinary magic: Resilience in development*. New York: Guilford Press.
- Masten, A. S. (2018). Resilience theory and research on children and families: Past, present, and promise. *Journal of Family Theory & Review*, 10(1), 12–31.
- Masten, A. S., & Garmezy, N. (1985). Risk, vulnerability, and protective factors in developmental psychopathology. In B. Lahey, & A. Kazdin (Eds.), *Advances in Clinical Child Psychology*. New York: Plenum Publishing.
- Masten, A. S., Lucke, C. M., Nelson, K. M., & Stallworthy, I. C. (2021). Resilience in development and psychopathology: Multisystem perspectives. *Annual Review of Clinical Psychology*, 17, 521–549.
- Merrell, K. W., & Gueldner, B. A. (2010). *Social and emotional learning in the classroom: Promoting mental health and academic success*. New York: Guilford Press.
- Naglieri, J. A., LeBuffe, P. A., & Shapiro, V. B. (2011). *The Devereux Student Strengths Assessment-mini (DESSA-mini): Assessment, technical manual, and user's guide*. Aperture Education.
- Naglieri, J. A., LeBuffe, P. A., & Shapiro, V. B. (2013). Assessment of social-emotional competencies related to resilience. *Handbook of resilience in children* (pp. 261–272). New York: Springer Science + Business Media. https://link.springer.com/chapter/10.1007/978-1-4614-3661-4_15

- National Association of School Psychologists. (2011). Ratio of students per school psychologist by state: Data from the 2009–10 and 2004–05 NASP membership surveys. https://www.nasponline.org/Documents/Research%20and%20Policy/Research%20Center/Ratios_by_State_2005_and_2010.pdf
- National Association for the Education of Young Children. (1988). Standardized testing of young children through 8 years of age. *Young Children*, 4(3), 42–47.
- 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.
- O’Connell, M. R., Boat, T., & Warner, K. E. (Eds.). (2009). *Preventing mental, emotional and behavioral disorders among young people: Progress and possibilities*. Washington, D.C.: The National Academies Press.
- Ozer, E. J., Shapiro, V. B., & Duarte, C. d. P. (2021). *Opportunities to strengthen SEL impact through Youth-led Participatory Action Research (YPAR)*. University Park, PA: Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University.
- 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. <https://doi.org/10.1111/j.1746-1561.2000.tb06468.x>
- Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., & Pachan, M. (2008). *The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews*. Collaborative for Academic, Social, and Emotional Learning. <https://files.eric.ed.gov/fulltext/ED505370.pdf>
- Rizopoulos, D. (2006). ltm: An R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software*, 17(5), 1–25. <https://doi.org/10.18637/jss.v017.i05>
- Rosenthal, R., & Rosnow, R. L. (2008). *Essentials of behavioral research: Methods and data analysis* (3rd ed.). New York: McGraw Hill.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Shapiro, V. B. (2015). Resilience: Have we not gone far enough? A response to Larry Davis. *Social Work Research*, 39, 7–10. <http://dx.doi.org/10.1093/swr/svv001>
- Shapiro, V. B., Accomazzo, S., & Robitaille, J. L. (2017). In the same ballpark or a whole new ball game? Staff as raters of youth behavior. *Journal of Child and Family Studies*, 26, 1051–1055. <http://dx.doi.org/10.1007/s10826-016-0632-1>
- Shapiro, V. B., Kim, B. K. E., Accomazzo, S., & Roscoe, J. N. (2016). Predictors of rater bias in the assessment of social-emotional competence. *International Journal of Emotional Education*, 8(2), 25–44.
- Shapiro, V. B., Kim, B. K. E., Robitaille, J. L., & LeBuffe, P. A. (2017). Protective factor screening for prevention practice: Sensitivity and specificity of the DESSA-mini. *School Psychology Quarterly*, 32(4), 449–464. <https://doi.org/10.1037/spq0000181>

- Shapiro, V. B., Kim, B. K. E., Robitaille, J. L., Mahoney, J. L., Lee, J., & LeBuffe, P. A. (2024). Monitoring the growth of social and emotional competence to guide practice decisions. *Social and Emotional Learning: Research, Practice, and Policy*, 3, 100032. <https://doi.org/10.1016/j.sel.2024.100032>
- Shapiro, V. B., & LeBuffe, P. A. (2006). Using protective factors in practice. *Annals of the New York Academy of Sciences*, 1094, 350–353. <http://dx.doi.org/10.1196/annals.1376.048>
- Substance Abuse and Mental Health Services Administration (2014). *A treatment improvement protocol: Improving cultural competence*. TIP 59. HHS Publication No. (SMA) 144849. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014. <https://pubmed.ncbi.nlm.nih.gov/25356450/>
- Tsang, K. L. V., Wong, P. Y. H., & Lo, S. K. (2012). Assessing psychosocial well-being of adolescents: A systematic review of measuring instruments. *Child: Care, Health and Development*, 38(5), 629–646. <https://doi.org/10.1111/j.1365-2214.2011.01355.x>
- U.S. Census Bureau. (2024). 2018–2022 American Community Survey 5-year estimates. Retrieved June 2024, from <https://data.census.gov/cedsci/>
- U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD). (January, 2024). Public elementary/secondary school universe survey, 2000–01, 2012–13, and 2019–20 through 2022–23. https://nces.ed.gov/programs/digest/d23/tables/dt23_204.10.asp
- U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD). (December, 2023). State nonfiscal survey of public elementary/secondary education, 2022–23. https://nces.ed.gov/programs/digest/d23/tables/dt23_204.70.asp
- Werner, E. E., & Smith, R. S. (1982). *Vulnerable but invincible: A longitudinal study of resilient children and youth*. 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.
- Woods C. M. (2009). Empirical selection of anchors for tests of differential item functioning. *Applied Psychological Measurement*, 33, 4–57. <https://doi.org/10.1177/0146621614561313>



About the DESSA System

The DESSA System from Riverside Insights is a student resilience system that helps educators measure and strengthen social, emotional, and behavioral skills to reduce challenges that interfere with learning. The DESSA universal screening and supplemental assessments are widely used, practical, evidence-based measures of social and emotional competence of students in grades K–12. These powerful data analytic tools enable educators to identify trends, validate curriculum effectiveness, and drive positive academic and well-being outcomes for students. Robust reporting in the DESSA System provides insights for educators and staff to proactively differentiate instruction using DESSA-provided instructional strategies and tiered intervention resources. The DESSA System’s strength-based approach is grounded in resilience theory and designed to bring the whole child into focus. To learn more, visit www.riversideinsights.com.