



# DESSA Screener for Externalizing and Internalizing Risk (DESSA SEIR)

User Guide and Technical Manual

The DESSA Screener for Externalizing and Internalizing Risk (SEIR) scales and the contents of this manual are reprinted with permission through a licensing agreement with SAIL CoLab, INC.

Copyright © 2025 by Riverside Assessments, LLC with permission. All rights reserved. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval system without the prior written permission of Riverside Assessments, LLC. Requests for permission to reproduce any part of the work should be sent via email to [permissions@riversideinsights.com](mailto:permissions@riversideinsights.com) or addressed via mail to Riverside Insights, Attention: Permissions, One Pierce Place, Suite 101C, Itasca, Illinois 60143.

These tests contain questions that are to be used solely for testing purposes by qualified organizations. No test items may be disclosed or used for any other reason. By accepting delivery of or using these tests, the recipient acknowledges responsibility for maintaining test security that is required by professional standards and applicable federal, state, and local laws, regulations, and policies governing proper use and administration of tests.

# TABLE OF CONTENTS

LIST OF TABLES AND FIGURES.....	III
ACKNOWLEDGMENTS.....	V
USER QUALIFICATIONS.....	VII
CHAPTER 1: Introduction to the DESSA SEIR .....	1
Expanding Social and Emotional Skill Assessment for Mental Health Screening .....	2
The DESSA Screener for Externalizing and Internalizing Risk (DESSA SEIR) Background. ....	3
Overview of the DESSA SEIR .....	3
Administration Guidelines .....	5
Completing the DESSA SEIR.....	6
Scoring and Scores.....	7
Raw Scores With Confidence Intervals.....	7
Criterion-Referenced Concern Levels .....	8
Validity Indices for Student Self-Ratings.....	11
Reporting Results .....	11

<b>CHAPTER 2: Reliability, Validity, and Fairness of DESSA SEIR Scores.</b>	<b>13</b>
Samples Used to Identify Items and Create an Interpretation Approach.	14
Identification of Items	16
Developing an Informative and Fair Framework to Interpret Concern Levels for Students With Internalizing and Externalizing Risks	17
Criterion-Referenced Emotional Behavior Concerns Framework (CREBCF)	17
Fairness of Criterion-Referenced Emotional Behavior Concern Levels.	18
Evidence for the Reliability and Validity of Scores From the DESSA SEIR	20
Reliability Evidence	20
Validity Evidence	21
Summary Evidence for the Reliability and Validity of the DESSA SEIR Scales.	22
Conclusions About Evidence for the Reliability, Validity, and Fairness of the DESSA SEIR Scales' Scores and Uses.	22
Future Directions	24
<b>REFERENCES</b>	<b>25</b>
<b>APPENDIX A: DESSA SEIR Raw Scores and Criterion- Referenced Concern Levels</b>	<b>29</b>

# LIST OF TABLES AND FIGURES

TABLE 1.1. DESSA SEIR Student Form (Grades 6–12) Items . . . . .	4
FIGURE 1.1. The DESSA and DESSA SEIR Assessment Framework. . . . .	5
FIGURE 1.2. DESSA SEIR Student Form Presented in the DESSA System Student Portal . . . . .	7
FIGURE 1.3. The Criterion-Referenced Internalizing Risks and Externalizing Risks Concern Levels . . . . .	9
TABLE 1.2. Cut Score Ranges for Concern Levels on the Internalizing and Externalizing Risks Scales of the DESSA SEIR Student Form . . . . .	11
TABLE 2.1. Demographic Characteristics of K–12 Students in Development and Performance Cut Score Samples (Percentages) . . . . .	15
TABLE 2.2. DESSA SEIR Scales (Student Form): Local Dependence, Sex DIF, and Race DIF Indices. . . . .	17
TABLE 2.3. Cut Scores Demarcating Criterion-Referenced Concern Levels for Each DESSA SEIR Student Form Scale . . . . .	18
TABLE 2.4. Impact of Concern Level Cut Scores on the Percentage of Students on the Internalizing Risks and Externalizing Risks Scales of the DESSA SEIR Student Form. . . . .	19
TABLE 2.5. Criterion-Referenced Emotional Behavior Concern Level Comparisons by Sex and Race/Ethnicity Subgroups on the Internalizing Risks and Externalizing Risks Scales of the DESSA SEIR Student Form . . . . .	19
TABLE 2.6. Reliability Statistics for DESSA SEIR Student Form Scales. . . . .	20
FIGURE 2.1. Test Information Functions (TIFs) for DESSA SEIR Student Form Scales. . . . .	21

TABLE 2.7. Correlations Between the DESSA SEIR Student Form Scales and the DESSA Student Self-Report (SSR).....	21
TABLE 2.8. Correlations Between the DESSA SEIR Student Form Scales and the BASC-2 Student Form Scales .....	22
TABLE A.1. DESSA SEIR Student Form (Grades 6–8) Internalizing Risks and Externalizing Risks Raw Scores .....	30
TABLE A.2. DESSA SEIR Student Form (Grades 9–12) Internalizing Risks and Externalizing Risks Raw Scores .....	31



# ACKNOWLEDGMENTS

This manual provides technical guidance for the administration, interpretation, and application of the DESSA Screener for Externalizing and Internalizing Risk (SEIR) scales to support a whole-child approach to student well-being. The DESSA SEIR scales and the contents of this manual are reprinted with permission through a licensing agreement with SAIL CoLab, INC. The items included in the DESSA SEIR are the Internalizing and Externalizing items of the SSIS SEL Brief + Mental Health Scales developed by Elliott, S. N., Anthony, C. J., DiPerna, J. C., Lei, P., & Gresham, F. M. (2020) and published by SAIL CoLab, INC. Tables, figures, and information in this manual were developed with the Internalizing and Externalizing items from the SSIS SEL Brief + Mental Health Scales (Elliott et al., 2020) and studies were completed using the Social Skills Improvement System (SSIS) Rating Scales (Gresham & Elliott, 2008) and are reprinted with permission.







# USER QUALIFICATIONS

The DESSA Screener for Externalizing and Internalizing Risk (DESSA SEIR) is a screener used for identifying adolescents in Grades 6–12 who may be experiencing some emotional behavior concerns.

It is relatively brief and highly accessible. The directions for completing it are clear and concise so that students with basic reading skills and an understanding of concepts of truth can easily respond to each item. It is completed by students in the online DESSA System Student Portal.

DESSA SEIR results are provided to administrators and educators in the online DESSA System. Interpretation of the DESSA SEIR requires some formal academic training or professional development in the basic principles and limitations of educational and psychological testing and in the communication of test results. Ultimately, DESSA SEIR users should administer the assessment according to the recommended guidelines and should be able to demonstrate their competency with interpreting and communicating results. All users should be knowledgeable about the benefits and limitations of general assessment and intervention techniques.





## Chapter 1

# **INTRODUCTION TO THE DESSA SEIR**

# Expanding Social and Emotional Skill Assessment for Mental Health Screening

---

Parents and educators want children to be happy, healthy, and academically successful and want to help them avoid having negative experiences and exhibiting problem behaviors. Of course, it is normal that all children have some difficult days socially and emotionally as they grow and learn. Unfortunately, however, some children experience prolonged social and emotional challenges that can interfere with their successful functioning and well-being at school and home. These children, perhaps 10%–15% in many schools, display undesirable or negative emotional behaviors (e.g., Forness et al., 2012). Common negative emotional behaviors that often persist and interfere with development involve internalizing concerns (e.g., anxiousness, depression) and externalizing concerns (e.g., aggressiveness, bullying) (Vidair et al., 2014). We refer to these undesirable behaviors as **internalizing and externalizing risks**. If identified early, virtually all of these risks can be significantly reduced or eliminated. However, in many schools today, identification of these risks has not been formally addressed in large-scale student support programs under the banner of multi-tiered student support (MTSS), positive behavior support, or even whole-child education initiatives. Rather, in many schools, teachers have been entrusted to nominate/refer students who exhibit behaviors indicative of social and emotional problems (e.g., Bruhn et al., 2014). Although teachers often identify students with externalizing difficulties for further assessment and/or intervention, they routinely have been found to under-identify students experiencing internalizing problems (e.g., Dowdy et al., 2013). Also, teachers in middle and high schools have more difficulty identifying students with internalizing concerns than do teachers in elementary schools because they generally interact far less with their students.

Identifying students with internalizing and externalizing risks as early as possible is a sound practice that is often overlooked in most social and emotional skill development programs. Many social and emotional learning leaders and scholars have advocated for a stronger focus on positive student strengths in today's schools. Although such a perspective has unquestionable merit, when taken to an extreme, there is potential to neglect students' needs in other areas, such as internalizing and externalizing risks, which can be present in students with all levels of social and emotional skills. Similarly, current models conceptualize mental health as being composed of two distinct dimensions: one dimension involves *psychosocial well-being/positive experiences*, and the other dimension involves *psychosocial distress/negative experiences*.

In summary, although one may be primarily interested in improving students' social and emotional competencies, it is wise to consider screening and monitoring for indicators of psychological distress as part of multi-tiered assessment and intervention efforts. To address this need, we provide two brief student-completed internalizing and externalizing risks scales (collectively referred to as the DESSA SEIR) that can be administered in conjunction with the DESSA screeners and assessments (LeBuffe et al., 2018; Robitaille et al., 2023, 2024, 2025) to provide an efficient screening of students' mental health.

# The DESSA Screener for Externalizing and Internalizing Risk (DESSA SEIR) Background

---

The DESSA SEIR items were selected and researched by Elliott and colleagues (2020). These externalizing and internalizing items are published as part of the SSIS SEL Brief + Mental Health Scales (Elliott et al., 2020) and are available in the DESSA System through a licensing agreement. The DESSA SEIR scales have their origins in the research and scale development work of Gresham and Elliott (1990, 2008). Over a thousand studies have been published in the past two decades using their scales (including the Social Skills Rating System [SSRS; Gresham & Elliott, 1990]) and the Social Skills Improvement System Rating Scales (SSIS; Gresham & Elliott, 2008) to document both the prosocial and problem behaviors of children from ages 3 to 18 (Elliott & Gresham, 2008). In general, Gresham and Elliott's research, as well as the research of many others, has documented that domains of problem behaviors have a moderate negative relationship with socially desired behavior domains. However, knowing this does not replace the need to directly assess both desired and undesired social and emotional behaviors of children—especially when efficient measures with strong psychometric evidence are available.

## Overview of the DESSA SEIR

---

The **DESSA SEIR** scales include 10 items that collectively assess a broad range of negative social and emotional behaviors. Specifically, the scales measure two broad risk domains:

- **Internalizing Risks:** Negative emotions and behaviors mostly directed inwardly involving feeling worried, anxious, sad, or lonely; exhibiting poor self-esteem; showing a lack of interest or limited engagement with others.
- **Externalizing Risks:** Negative emotions and behaviors mostly directed toward others involving verbal or psychological aggression, threatening and bullying others, poor control of temper, arguing with others, actively excluding others from activities.

Using an Item Response Theory (Hambleton et al., 1991) approach to item efficiency maximization, Elliott and colleagues (2020) created the DESSA SEIR by selecting the five most informative items from the student form of the SSIS Rating Scales (Gresham & Elliott, 2008) for each emotional behavior risk domain. Table 1.1 provides the items included on the DESSA SEIR Student Form. Scores from each DESSA SEIR scale are interpreted via a 3-level, concerns-referenced framework (No Concern, Possible Concern, Concern).

**Table 1.1. DESSA SEIR Student Form (Grades 6–12) Items**

DESSA SEIR Scale	Student Form Items
Internalizing Risks	<ul style="list-style-type: none"><li>■ I think bad things will happen to me.</li><li>■ I feel lonely.</li><li>■ I think no one cares about me.</li><li>■ I feel nervous with my classmates.</li><li>■ I feel sad.</li></ul>
Externalizing Risks	<ul style="list-style-type: none"><li>■ I often do things without thinking.</li><li>■ I make people do what I want them to do.</li><li>■ I do not let others join my group of friends.</li><li>■ I fight with others.</li><li>■ I break the rules.</li></ul>

The DESSA SEIR scales and the DESSA are used together to provide a comprehensive picture of a student’s overall social and emotional well-being. While the DESSA assesses resilience factors that highlight a student’s strengths, the DESSA SEIR identifies risk factors that may indicate the need for additional support. By integrating these measures, educators gain a comprehensive understanding of both protective factors and potential concerns (Elliott et al., 2021; Johnson & Parks, 2024).

The DESSA SEIR scales complement the DESSA screeners and assessments by supporting a whole-student approach to social and emotional assessment. Together, these tools help:

- Screen for risk by highlighting emotional and behavioral concerns,
- Identify strengths that contribute to resilience and positive development,
- Provide insights into areas that can be leveraged for growth and those that may require targeted intervention or support.

By combining the DESSA’s strength-based insights with the DESSA SEIR’s risk identification, these tools help educators make data-driven decisions about instruction, intervention, and support.

Figure 1.1 illustrates how the DESSA and DESSA SEIR together define key dimensions of students’ mental health and well-being, offering a clear framework for assessing and fostering their overall social and emotional development.

**Figure 1.1. The DESSA and DESSA SEIR Assessment Framework**

	Intra-Personal Dimension	Inter-Personal Dimension
<b>Resilience Factors (Social and Emotional Skills)</b>	Self-Awareness Self-Management Optimistic Thinking Responsible Decision Making	Social Awareness Relationship Skills
<b>Risk Factors (Emotional Behavior Concerns)</b>	Internalizing Risks	Externalizing Risks

Collectively, the results from the DESSA and DESSA SEIR provide valuable insights that can be used to:

- **Design targeted interventions** that build students’ social and emotional skills (resilience),
- **Document students’ progress over time** to ensure that timely and effective supports are provided, and
- **Assess the impact of mental health programs** by evaluating changes in students’ risk and resilience factors.

## Administration Guidelines

The DESSA and DESSA SEIR can be used by educators, school psychologists, behavior specialists, social workers, or other qualified education professionals to efficiently assess an individual student’s or an entire school of students’ social and emotional competencies and emotional behavior risks. Prior to administering these assessments, it is important for users to be familiar with the administration guidelines. DESSA SEIR administration guidelines are presented in this section; DESSA administration guidelines are presented in Chapter 4 of the respective DESSA technical manuals. In addition, professionals who interpret scores resulting from the completion of the DESSA SEIR must meet the minimum qualification requirements described in the **User Qualifications** section on page vii.

Obtaining students’ self-ratings is critical for mental health screening, given that several emotions and beliefs cannot be observed by adults. The DESSA SEIR has a second-grade readability level, so most students in Grades 6–12 should be able to read and comprehend the items without accommodation; however, reading support can be provided without affecting the validity of scores because reading is not the target behavior of concern.

Ideally, students should complete the form in a controlled setting at a school to facilitate monitoring the administration. Some students may lack motivation to complete the form because they don’t have experience with this assessment or because they think the task is part of a referral for special services. These students may require prompting and support to complete the assessment, and users should

examine their scores to ensure validity. The DESSA SEIR is completed by students in the DESSA System Student Portal. Users are usually classroom educators or other school staff who work with students to help them access and start the assessment.

It is always important for users to establish good rapport with students when administering an assessment. Parents, educators, and students may often be concerned about the eventual use of information obtained from the rating forms, and some students may not appreciate the importance of accurate ratings. Users should be open and honest with students regarding such issues as the intended use of results, reasons for the evaluations, limits of confidentiality, and legal issues like policies/laws regarding privileged communication and consent to participate in an assessment. Users should emphasize to students the importance of responding to all items on the assessment form, even if students are uncertain about how to answer a given item. Users should be sure to stress to students the value of honest responses and should emphasize that there are no right or wrong answers.

## Completing the DESSA SEIR

---

The DESSA SEIR is administered online using the DESSA System Student Portal. It is presented to the student immediately following the DESSA Student Self-Report (DESSA SSR) and includes the instructions followed by each of the 10 items. Instructions for responding to the DESSA SEIR items are brief. Each item is a sentence that describes a social-emotional behavior, and students are expected to decide how true each item is for them. A response is required for each item, and the directions indicate that there are no right or wrong answers. The instructions, items, and response format were developed in accordance with the fairness guidelines from the *Standards for Educational and Psychological Testing* (AERA et al., 2014). Figure 1.2 provides an illustration of how the DESSA SEIR instructions, items, and response options are displayed for students within the DESSA System Student Portal.



**Figure 1.2.** DESSA SEIR Student Form Presented in the DESSA System Student Portal

Directions: You will be presented with different things that students your age may do. Please read each sentence and think about yourself. Decide how true each sentence is for you.

If you think it is **not true** for you, select **Not True**.

If you think it is **a little true** of you, select **A Little True**.

If you think it is **a lot true** of you, select **A Lot True**.

If you think it is **very true** of you, select **Very True**.

Please answer all questions with the best response for you for each sentence, even if it is hard for you to make up your mind. There are no right or wrong answers. Please ask questions if you do not know what to do.

**These questions should take about one minute to complete.**

**BEGIN**

**DESSA**

Decide how true each sentence is for you.

**I often do things without thinking.**

☒ Not True

☐ A Little True

☐ A Lot True

☐ Very True

[< GO BACK](#)

## Scoring and Scores

The DESSA SEIR yields two types of scores for describing and evaluating students' emotional and behavioral risks: raw scores and criterion-referenced concern levels.

### Raw Scores With Confidence Intervals

Each of the two DESSA SEIR scales has 5 items. Each item has four possible responses—*Not True*, *A Little True*, *A Lot True*, *Very True*—worth 0, 1, 2, and 3 points, respectively. Summing the points for the 5 items on each scale creates a total raw score ranging from 0–15. There is not a combined or composite score for the total set of 10 items as the two scales (Internalizing Risks and Externalizing Risks) represent distinct constructs that demonstrate a moderate relationship with

each other. Scores for the individual scales represent undesired levels of emotional and behavioral concerns, with higher scores indicating more concerns and potentially a more negative mental health status.

Many assessments transform raw scores to standard scores to enable normative interpretation and put all scores on the same scale (e.g., the DESSA assessments); however, such a transformation of scores is unnecessary with the DESSA SEIR scales because there are an equal number of items on both the Internalizing Risks and Externalizing Risks scales.

It should be noted that psychological measurement is imperfect. The calculation of confidence intervals establishes “reasonable limits” for interpretation of test scores, considering measurement error, and are imperative for ethical use of psychoeducational measures. The application of confidence intervals helps ensure that a proper level of precision is used when interpreting a performance level. This approach also discourages overemphasis on small, unreliable differences between the scores of groups of individuals or within an individual’s own profile of scores. As such, 95% confidence intervals are provided in **Appendix A** for each DESSA SEIR scale.

## Criterion-Referenced Concern Levels

To facilitate the meaning of the resulting total raw score for both the DESSA SEIR Internalizing Risks and Externalizing Risks scales, Elliott and colleagues (2020) developed a three-level, criterion-referenced framework to characterize the relative level of mental health concern. The three levels for the Criterion-Referenced Emotional Behavior Concerns Framework (CREBCF) are *No Concern*, *Possible Concern*, and *Concern*. This approach is cautious and appropriate given the primary purpose of the DESSA SEIR scales is screening. The DESSA SEIR scales are intended to identify students who might be experiencing mental health concerns. As explained further in Chapter 2, most students’ scores will fall within the *No Concern* level, and no further action is necessary unless their status changes in future screenings. Some students’ scores will fall within the *Possible Concern* level, indicating that, when considering measurement error, their true functioning may be at the level of *Concern*. For students with scores in this range, we recommend monitoring their functioning between screenings to ensure that they are not experiencing mental health difficulties. Finally, a small percentage (< 10%) of students will score within the *Concern* level. For students at this level, a mental health professional (school psychologist, counselor, social worker, etc.) should promptly follow up with the student about their current mental health status, and based on the outcome of that interaction, should conduct a more comprehensive assessment of any social, emotional, and/or mental health concerns. Figure 1.3 documents the description of each of the three levels of concern for the Internalizing Risks and Externalizing Risks scales.

**Figure 1.3. The Criterion-Referenced Internalizing Risks and Externalizing Risks Concern Levels**

<b>Internalizing Risks Concern Levels</b>
<p><b>No Concern Level</b></p> <p>Students at this level rarely (<b>never or seldom</b>)...</p> <ul style="list-style-type: none"><li>■ Comment about negative events in life,</li><li>■ Act worried,</li><li>■ Feel down,</li><li>■ Act lonely,</li><li>■ State others don't like them.</li></ul> <p>These students do not require additional attention or support at this time.</p>
<p><b>Possible Concern Level</b></p> <p>Students at this level sometimes (<b>seldom or often</b>)...</p> <ul style="list-style-type: none"><li>■ Comment about bad things happening,</li><li>■ Act nervous interacting with others,</li><li>■ Feel sad,</li><li>■ Act lonely,</li><li>■ State others don't care about them.</li></ul> <p>These students may need additional attention and some focused support, along with periodic monitoring of their potential internalizing concern(s), to ensure the situation does not worsen.</p>
<p><b>Concern Level</b></p> <p>Students at this level frequently (<b>often or almost always</b>)...</p> <ul style="list-style-type: none"><li>■ Talk about bad things happening to them,</li><li>■ Act anxious interacting with others,</li><li>■ Feel depressed,</li><li>■ Act lonely,</li><li>■ State few people care about them.</li></ul> <p>These students likely need ongoing support and monitoring from school professionals until a comprehensive assessment of their internalizing concern(s) can be completed and appropriate support services determined.</p>

**Figure 1.3. The Criterion-Referenced Internalizing Risks and Externalizing Risks Concern Levels (*continued*)**

### **Externalizing Risks Concern Levels**

#### **No Concern Level**

Students at this level rarely (**never or seldom**)...

- Persuade others to do something they dislike doing,
- Verbally argue with others,
- Act without thinking things through,
- Break agreed-upon rules,
- Keep others out of social groups.

These students do not require additional attention or support at this time.

#### **Possible Concern Level**

Students at this level sometimes (**seldom or often**)...

- Pressure others to do things against their will,
- Fight physically with others,
- Act impulsively,
- Break rules intentionally,
- Exclude others from social groups.

These students may need additional attention and some focused support, along with periodic monitoring of their potential externalizing concern(s) to ensure the situation does not worsen.

#### **Concern Level**

Students at this level frequently (**often or almost always**)...

- Force others to do things against their will,
- Fight physically with others,
- Act impulsively,
- Break rules intentionally,
- Isolate others from groups.

These students likely need ongoing support and monitoring from school professionals until a comprehensive assessment of their externalizing concern(s) can be completed and appropriate support services determined.

The DESSA SEIR scale raw scores and their translation to a criterion-referenced concern level are listed in **Appendix A**. Table 1.2 documents the cut score ranges for each of the concern levels.

**Table 1.2. Cut Score Ranges for Concern Levels on the Internalizing and Externalizing Risks Scales of the DESSA SEIR Student Form**

Concern Level	Internalizing Risks	Externalizing Risks
No Concern	0–6	0–4
Possible Concern	7–8	5–6
Concern	9–15	7–15

*Note.* The DESSA SEIR Internalizing Risks scale and Externalizing Risks scale both have 5 items that are each scored from 0–3; thus, the possible score range is 0–15, with a higher score indicative of a higher frequency of concerning behavior.

## Validity Indices for Student Self-Ratings

Although rating scales are an efficient and effective assessment tool, they are subject to inconsistencies in reporting, especially for self-report scales. It is important to note that scores above 13 on either scale were extremely rare or nonexistent in the Performance Cut Score Sample (see Chapter 2 for more information about this sample). As such, if a student receives a score above these thresholds (specified in **Appendix A**), either the student is demonstrating substantial mental health concerns or the student’s responses are not valid. In either case, it is the user’s responsibility to follow up promptly to determine which interpretation is accurate.

## Reporting Results

Several types of reports can be generated in the DESSA System for the DESSA SEIR, including individual student and group reports. Administrators and educators may select from the available filters in the DESSA System to design reports that provide score breakdowns for different groups of students (e.g., classroom, grade, school) and for a variety of demographic subgroups (e.g., gender, race and ethnicity, special populations). Within the Individual Student Report, users are provided with a snapshot of both DESSA SEIR and DESSA results to provide a profile of students’ social and emotional competencies and emotional behavior risks together.

More information about available reporting is provided within the DESSA System Support Portal. An examination of the reports illustrates two key points:

- Scores on the Internalizing Risks and Externalizing Risks scales are indicators of the frequency with which students report exhibiting behaviors commonly indicative of psychosocial distress, and
- Three basic criterion-referenced concern levels provide an effective means for characterizing concern and the need for follow-up action regarding a student’s mental health support needs.

In summary, a DESSA SEIR report is an important outcome of the assessment process. When the assessment process is driven by a purpose and focuses on social and emotional concerns with known relevance to a student’s mental health support needs, the time-efficient DESSA SEIR is a useful tool.

Substantial evidence supports the utility and technical soundness of the DESSA SEIR scales and their resulting scores. **Chapter 2** provides an account of this evidence regarding the reliability, validity, and fairness of the scores.



## Chapter 2

# **RELIABILITY, VALIDITY, AND FAIRNESS OF DESSA SEIR SCORES**

Together, the DESSA and DESSA SEIR provide time-efficient screening assessments that can be used within multi-tiered programs to support the social and emotional well-being of all students. This chapter will focus on the development and psychometric properties of the DESSA SEIR.<sup>1</sup>

As explained in the previous chapter, the DESSA SEIR was developed using the original SSIS Rating Scales pool of problem behavior items (Gresham & Elliott, 2008), which have substantial validity evidence for a large, diverse sample of students. Item response theory (IRT) provided the method by which the most informative items were selected to develop brief scales for screening the emotional behavior risks of students in Grades 6–12 (i.e., the DESSA SEIR). This process is discussed in this chapter, followed by a description of the development of the criterion-referenced framework for interpreting student's scores on the Internalizing and Externalizing Risks scales.

Underlying all the development steps for the DESSA SEIR was the design imperative of fairness. That is, the DESSA SEIR scales were designed to be responsive “to individual characteristics and testing contexts so that test scores will yield valid interpretations for intended uses” (AERA et al, 2014, p. 50).

## Samples Used to Identify Items and Create an Interpretation Approach

---

Two samples of students were used at different stages of the development and initial validation of the DESSA SEIR scales. These student samples were each derived from the original national SSIS sample of 4,000 children ages 5–18 from 36 states as rated by their teachers, parents, and themselves in 2006–2007 (Gresham & Elliott, 2008). The first sample, the *Development Sample*, consisted of the standardization samples of the SSIS SEL (Gresham & Elliott, 2017). Next, the Performance Cut Score Sample consisted of a sample drawn from the broader sample of original SSIS cases. These samples were drawn to be representative of demographic targets for sex and race/ethnicity projected for the U.S. in 2025 (Snyder et al., 2018). Table 2.1 provides the demographic characteristics of these samples. For a full explanation of these samples, see the technical manual for the SSIS SEL Brief + Mental Health Scales (Elliott et al., 2020). Note that Table 2.1 presents the demographic characteristics for the full original samples containing student, parent, and teacher ratings; the DESSA SEIR was developed specifically using ratings from students in Grades 6–12.

For this chapter, all IRT analyses related to the development of the DESSA SEIR used the Development Samples, and all validation work (i.e., Cronbach's  $\alpha$ , interscale correlations, fairness analyses) was conducted using the Performance Cut Score Samples. Further validation analyses (e.g., test-retest reliability analyses, concurrent validity analyses) were conducted with various subsamples stratified by grade clusters of the total sample appropriate for each analysis and noted where appropriate in this chapter.

---

<sup>1</sup>Information about the development and psychometric properties of the DESSA screeners and assessments can be found in Chapters 2 and 3 of each DESSA technical manual.



**Table 2.1. Demographic Characteristics of K–12 Students in Development and Performance Cut Score Samples (Percentages)**

Characteristic	Development Samples			Performance Cut Score Samples			U.S. Population (2025 Projected) <sup>b</sup>
	Teacher (n = 750)	Student (n = 800)	Parent (n = 2,400)	Teacher (n = 611)	Student (n = 772)	Parent (n = 1,507)	
Female	50	50	50	51	51	51	49 <sup>c</sup>
Race							
White	59	59	59	50	50	50	48 (45)
Black	16	15	16	15	15	15	15 (15)
Hispanic	19	19	20	25	25	25	26 (28)
Other	6	7	6	10	10	10	10 (12)
Grade							
K–3	42	17 <sup>a</sup>	44	42	10 <sup>a</sup>	28	30
4–6	26	38	33	30	39	33	23
7–8	16	21	13	16	25	20	15
9–12	16	23	11	12	25	19	30
Region							
Northeast	18	18	18	17	15	18	16
Midwest	22	22	22	22	28	23	21
South	36	36	37	33	33	36	39
West	25	24	23	28	24	23	24
Parent's Education Level							
Grade 11 or Less	14	13	13	14	11	12	11
Grade 12 or GED	29	29	29	32	32	30	19
1–3 Years of College	31	30	30	33	32	33	26
4+ Years of College	27	27	28	20	25	26	44
Educational Status							
General Education	92	91	92	90	92	90	87
Special Education	8	9	8	10	8	10	13

*Note.* Some percentages do not sum to 100 due to rounding.

<sup>a</sup> Only includes 3rd graders.

<sup>b</sup> Except where noted, estimates and projections are from the *Digest of Educational Statistics 2016* (Snyder, de Brey, & Dillow, 2018).

<sup>c</sup> Derived from the most recent estimates from the 2013-14 Civil Rights Data Collection (CRDC; U.S. Department of Education, Office for Civil Rights, 2016).

## Identification of Items

---

To develop the DESSA SEIR, Elliott and colleagues (2020) utilized Item Response Theory (IRT; see Chapter 4 of Elliott et al., 2020 for more information). Based on the content needs of the DESSA SEIR scales, two pools of original SSIS Problem Behaviors items were utilized for these analyses. First, the set of items making up the SSIS Internalizing scale was used for the development of the DESSA SEIR Internalizing Risks scale. Next, although they were overlapping but technically distinct on the original SSIS, the Externalizing and Bullying scales of that measure were combined for the purposes of selecting items for the DESSA SEIR Externalizing Risks scale. This was done to address the critical importance of identifying bullying-related concerns in school (Elliott et al., 2019). Including the extra items from the SSIS Bullying scale generally resulted in an item pool that augmented the original SSIS Externalizing scale item pool with two to three additional SSIS Bullying scale items. The unidimensionality criterion (ratio of first to second eigenvalues  $> 4$ ) was met in all cases with a first to second eigenvalue ratio of 6.07 for the Internalizing Risks scale and an equivalent ratio of 7.72 for the Externalizing Risks scale. Thus, this completed the item analysis and selection on these pools of items. As a result of this process, the DESSA SEIR Student Form scales were identified for Grades 6–12.

Beyond these standard IRT analyses, Elliott and colleagues (2020) also computed Differential Item Functioning (DIF) analyses to examine possible item bias. Specifically, a two-step procedure (Tay et al., 2015) was utilized to evaluate DIF based on sex (male vs. female) and race/ethnicity (White vs. non-white). These groupings were chosen to ensure that sample sizes were sufficient for DIF analyses. For items flagged as having empirical evidence of DIF, we also computed an effect size measure, the Expected Score Standardized Difference (ESSD; Meade, 2010) to facilitate consideration of items with DIF. This index is standardized and thus Cohen's (1988) criteria for small, medium, and large ESSD values (0.20, 0.50, and 0.80 for small, medium, and large effects, respectively) was utilized. These indices were considered during item selection.

Table 2.2 reports local dependence and sex and race DIF for the DESSA SEIR items. It was impossible to identify DESSA SEIR scales composed entirely of items that met all of the a priori criteria (e.g., no local dependence, no statistical evidence of DIF). Nevertheless, relatively few items were retained that violated a priori criteria, and all violations are considered to represent minor rather than major violations.

**Table 2.2. DESSA SEIR Scales (Student Form): Local Dependence, Sex DIF, and Race DIF Indices**

Scale/Items	Local Dependence	Bias Criteria (ESSD)	
		Sex	Race
Internalizing Risks			
I think bad things will happen to me.	ns	ns	ns
I think no one cares about me.	ns	ns	ns
I feel lonely.	ns	ns	✓(−0.07)
I feel sad.	ns	ns	ns
I feel nervous with my classmates.	ns	ns	ns
Externalizing Risks			
I often do things without thinking.	ns	✓(0.21)	ns
I make people do what I want them to do.	ns	ns	✓(0.02)
I fight with others.	ns	ns	ns
I break the rules.	ns	ns	ns
I do not let others join my group of friends.	ns	ns	ns

Note. *ns* indicates the difference was not significant.

## Developing an Informative and Fair Framework to Interpret Concern Levels for Students With Internalizing and Externalizing Risks

Many, if not most, social behavior rating scales utilize a normative framework for score interpretation. While there is much merit to this approach, there are also advantages to criterion-referenced interpretive frameworks, especially because the distributions for the Internalizing Risks and Externalizing Risks scales were highly skewed (as are other similar measures of student's problem behaviors). Thus, a criterion-referenced approach was developed for the DESSA SEIR.

### Criterion-Referenced Emotional Behavior Concerns Framework (CREBCF)

In the development of the CREBCF for Internalizing and Externalizing Risks scales, Elliott and colleagues (2020) initially sought to identify two levels: *Concern*, a level indicating that the reported behaviors very probably constitute a concern that should be further evaluated and assessed, and *No Concern*, a level indicating that no further action is likely necessary. To determine the scores demarcating each level, three criteria were considered. First, they used anchor scale logic; that is, they considered the plain meaning of scores with reference to the anchors defining each item's rating scale point. For example, on the Externalizing Risks scale, a score of 8

would mean that the student’s average item score on the 5 items constituting this scale was 1.6. Such a score indicates that the behaviors indicative of Externalizing problems were between *Seldom* and *Often* but were closer to *Often* than *Seldom*. Second, they considered the overall score distributions from the Performance Cut Score Samples as well as distributions across race/ethnicity (White, Black, and Hispanic) and sex (Male, Female). These considerations were intended to promote fairness by ensuring that the chosen cut scores would (a) identify no more than 10% of students overall as having Internalizing or Externalizing scores in the Concern Level and (b) result in similar impact regarding percentages of identified students across race/ethnicity and sex. Considering the higher stakes of identifying negative emotional behavior concerns relative to positive social and emotional skills and competencies, Elliott and colleagues were open to having different cut scores across forms/raters rather than a universal cut score. Third and finally, because it is important to consider measurement error when making decisions about individual students, they identified a third level of concern called *Possible Concern* Level. This level indicates that a student’s score does not effectively rule out the possibility of being at the *Concern* Level in the assessed domain, but neither does the score clearly indicate a level of concern that should be followed up with further assessment or intervention. This level was generated by subtracting 2 raw score points from each concern level cut score. This was done because the approximate margin of error for each scale was 2 raw score points. Based on these considerations, they generated cut scores for each DESSA SEIR scale as shown in Table 2.3.

**Table 2.3. Cut Scores Demarcating Criterion-Referenced Concern Levels for Each DESSA SEIR Student Form Scale**

Concern Levels	Internalizing Risks	Externalizing Risks
No Concern	0–6	0–4
Possible Concern	7–8	5–6
Concern	9–15	7–15

## Fairness of Criterion-Referenced Emotional Behavior Concern Levels

A fair interpretation approach to scores should not systematically favor one group over another. However, as highlighted in the *Standards for Educational and Psychological Testing* (AERA et al., 2014), “group differences in outcomes do not in themselves indicate that a testing application is biased or unfair.... For example, some racial and ethnic subgroups have lower mean scores on some standardized tests than do other subgroups. Some of the factors that contribute to these differences are understood..., but even where serious efforts have been made to eliminate possible sources of bias in test content and formats, the potential for some score bias cannot be completely ruled out” (p. 54). As an initial examination of fairness, and to evaluate the success of the cut-score generation process, Elliott et al., 2020 evaluated the overall percentages of a national sample (the Performance Cut Score Samples) falling within each Concern Level (Table 2.4).

**Table 2.4. Impact of Concern Level Cut Scores on the Percentage of Students on the Internalizing Risks and Externalizing Risks Scales of the DESSA SEIR Student Form**

Grade	Internalizing Risks			Externalizing Risks		
	No Concern	Possible Concern	Concern	No Concern	Possible Concern	Concern
6–8	90%	5%	6%	81%	13%	6%
9–12	87%	9%	4%	75%	14%	11%

As can be seen under each Concern column listed in Table 2.4, the cumulative percentages indicate the authors met their goal of identifying no more than 10% of students overall across grade clusters as having an internalizing or externalizing concern.

In addition to evaluating the overall impact of the CREBCF cut scores, the authors also evaluated whether these cut scores would lead to disparate impact across race/ethnicity and sex. To do so, they calculated Cohen’s *h* statistics and associated statistical significance values (see Chapter 4 of Elliott et al., 2020 for more information). Cohen’s *h* values between .20 and .50 were considered small, those between .51 and .80 were considered medium, and those exceeding .80 were considered large. As can be seen in Table 2.5, all differences were small and not statistically significant, with the exception of a medium difference (–0.54) found for the Internalizing Risks scale for the Black vs. White comparison. Positive *h* values indicate that the first group in the comparison had higher percentages in a particular category and vice versa for negative *h* values.

**Table 2.5. Criterion-Referenced Emotional Behavior Concern Level Comparisons by Sex and Race/Ethnicity Subgroups on the Internalizing Risks and Externalizing Risks Scales of the DESSA SEIR Student Form**

Grade	Girls vs. Boys		Race/Ethnicity					
			Black vs. Hispanic		Black vs. White		Hispanic vs. White	
	I	E	I	E	I	E	I	E
6–8								
No Concern	–0.07	–0.17*	0.05	–0.17	0.11	–0.08	0.06	0.09
Possible Concern	0.01	0.03	0.06	0.07	0.06	0.03	0.00	–0.03
Concern	0.09	0.25	–0.16	0.20	–0.23	0.09	–0.08	–0.11
9–12								
No Concern	0.10	–0.09	0.05	–0.37	0.35*	–0.17	0.31*	0.20
Possible Concern	–0.06	–0.01	0.03	0.45	–0.15	0.22	–0.18	–0.22
Concern	–0.10	0.14	–0.29 <sup>a</sup>	0.00	–0.54 <sup>a</sup>	–0.03	–0.25	–0.03

Note. I = Internalizing Risks scale; E = Externalizing Risks scale.

<sup>a</sup> Due to some zero cells, *p* value was unable to be computed.

\* *p* < .05

In summary, there does not appear to be any discernable pattern indicating disproportionate impact of the CREBCF cut points across race/ethnicity or sex. This provides further evidence that scores from the Internalizing and Externalizing Risks scales within the DESSA SEIR are fair and have substantial consequential validity.

## Evidence for the Reliability and Validity of Scores From the DESSA SEIR

Reliability evidence for the DESSA SEIR scores was evaluated using both Classical Test Theory (CTT) and IRT. Furthermore, validity was evaluated by examining Internalizing and Externalizing Risks scales intercorrelations, as well as the correlations between the scales and the Behavior Assessment System for Children–Second Edition (BASC-2; Reynolds & Kamphaus, 2004).

### Reliability Evidence

Score reliability for the Internalizing and Externalizing Risks scales was evaluated using several forms of estimates. For the CTT portion of these analyses, internal consistency (evaluated through Cronbach’s  $\alpha$ ) and test-retest reliability coefficients were used. As shown in Table 2.6, internal consistency values on the DESSA SEIR Student Form met or were near a .80 criterion for low-stakes screening decisions (Salvia et al., 2010). Test-retest reliability coefficients, also shown in Table 2.6, were moderate for both scales.

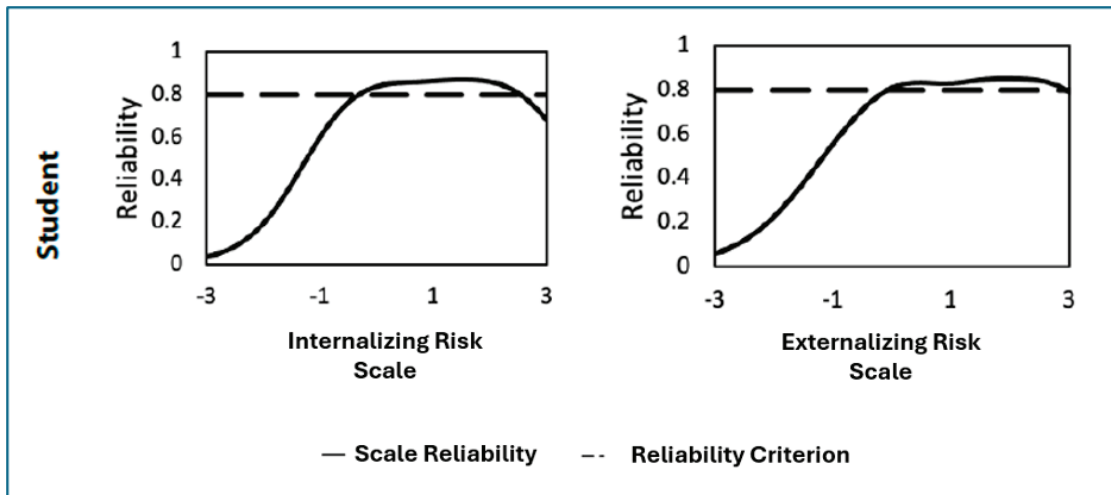
**Table 2.6. Reliability Statistics for DESSA SEIR Student Form Scales**

	Cronbach’s $\alpha$		Test-Retest	
	Internalizing Risks	Externalizing Risks	Internalizing Risks	Externalizing Risks
Student Form	.80	.75	.60	.63

*Note.* All correlations are statistically significant ( $p < .001$ ).

With regard to IRT analyses, Test Information Functions (TIFs) for each DESSA SEIR Student Form scale are presented in Figure 2.1. These TIFs were based on the Development Sample and indicate that precision is likely high enough for low-stakes decisions across a broad spectrum of internalizing and externalizing concerns. Especially encouraging is the high level of precision at the ranges of Internalizing and Externalizing Risks scores indicating concerns in need of further attention.

**Figure 2.1.** Test Information Functions (TIFs) for DESSA SEIR Student Form Scales



Note. Test Information converted to reliability metric through formula from Petrillo et al. (2015).

## Validity Evidence

Validity evidence was primarily evaluated via examination of the intercorrelations between the Internalizing and Externalizing Risks scales, correlations between these scales and the DESSA Student Self-Report (SSR), and correlations between the DESSA SEIR and the BASC-2 (Table 2.7 and Table 2.8). The correlations between the DESSA SSR and the DESSA SEIR scales were conducted as part of a larger study on student well-being (Johnson & Parks, 2024).

**Table 2.7.** Correlations Between the DESSA SEIR Student Form Scales and the DESSA Student Self-Report (SSR)

	DESSA SEIR Student Form	
	Internalizing Risks	Externalizing Risks
<b>DESSA SSR</b>		
Self-Awareness	-.37	-.18*
Self-Management	-.33	-.38
Social Awareness	-.20	-.29
Relationship Skills	-.18	-.32
Responsible Decision Making	-.29	-.29
Optimistic Thinking	-.37	-.19*
Social-Emotional Composite	-.35	-.30
<b>Internalizing–Externalizing Risks Intercorrelations</b>	.39	

Note. Unless otherwise noted, all correlations are statistically significant ( $p < .001$ ).

\*  $p < .01$

**Table 2.8. Correlations Between the DESSA SEIR Student Form Scales and the BASC-2 Student Form Scales**

BASC-2 Scales	DESSA SEIR Student Form (N = 53)	
	Internalizing Risks	Externalizing Risks
Externalizing/Inattention-Hyperactivity	.31*	.51
Internalizing	.64	.38**
School Problems	.26 <sup>†</sup>	.38**
Adaptive Skills/Personal Adjustment	-.56	-.37**
Behavioral Symptoms/Emotional Symptoms	.69	.43**

*Note.* BASC-2 = Behavior Assessment System for Children–Second Edition. Because BASC-2 scales differ across informant, rows represent content-similar scales, where relevant. All scales occurring after a slash indicate scales on the BASC-2 Student Form but not on the BASC-2 Teacher/Parent Forms. Unless otherwise noted, all correlations are statistically significant ( $p < .001$ ).

\*  $p < .05$ , \*\*  $p < .01$ , <sup>†</sup>  $p < .10$

## Summary Evidence for the Reliability and Validity of the DESSA SEIR Scales

Overall, initial evidence was very positive for the DESSA SEIR scales. Precision/reliability estimates indicated sufficient precision for low-stakes decision making. Furthermore, as indicated by scale TIFs, score reliability was highest precisely in score ranges most likely to indicate concerns that need to be further assessed.

With regard to validity, all correlations were in line with expectations and supported the validity of scores from the DESSA SEIR scales. For example, correlations between the DESSA Optimistic Thinking subscale and the Internalizing Risks scale showed moderate, negative correlations. Additionally, correlations between the Externalizing Risks scale and the BASC-2 Externalizing/Hyperactivity-Inattention scales were higher than the Internalizing Risks scale correlations with these scales. The same pattern held for the Internalizing Risks scale–BASC-2 Internalizing scale correlations. Such findings offer strong evidence of the validity of scores from the efficient Internalizing and Externalizing Risks scales of the DESSA SEIR Student Form.

## Conclusions About Evidence for the Reliability, Validity, and Fairness of the DESSA SEIR Scales' Scores and Uses

A systematic multiphase process was used to identify the best sets of items from the Problem Behavior Scale of the SSIS Rating Scales–Student Version to create efficient scales focused on internalizing and externalizing emotional behavior concerns. These scales had to yield reliable, valid, and fair scores to be used



appropriately in school-based universal screening programs. Although several comprehensive measures of children's and youth's mental health already exist, none of them also measure desired social and emotional competencies while concurrently screening for key indicators of emotional behavior concerns. Thus, the intention was to create a set of scales that could be used to efficiently screen students for possible mental health concerns, rather than create another comprehensive diagnostic assessment.

To achieve this goal, Elliott et al. (2020) selected 5 items that represent common internalizing concerns and 5 items that represent common externalizing concerns of students ages 5 to 18. Assessment efficiency and consistent item content across multiple informants were major considerations in limiting the number of items to 5 for each scale. Shorter scales typically suffer lower reliability estimates but, using an IRT approach with stringent item selection criteria, the authors developed brief scales that still meet or exceed accepted standards.

Evidence for the reliability and precision of scores for the DESSA SEIR scales is very promising. Specifically, internal consistency values were near or at a .80 criterion for low-stakes screening decisions. Finally, TIFs also indicated that precision is high enough for low-stakes decision making across a broad spectrum of internalizing and externalizing concerns. In particular, the precision was greatest in the ranges of DESSA SEIR scale scores indicating concerns in need of further attention. In summary, there is substantial and solid evidence that the DESSA SEIR Student Form provides reliable scores, particularly for students with moderate- to high-frequency internalizing or externalizing risk concerns.

With regard to validity evidence for DESSA SEIR scale scores, all correlations were in line with expectations and supported the validity of scores. The strongest evidence perhaps was the pattern of convergent validity and discriminant validity with the BASC-2 as a criterion measure. Correlations between the Externalizing Risks scale and the BASC-2 Externalizing/Hyperactivity-Inattention scales were consistently higher than the Internalizing Risks scale correlations were with these BASC-2 scales. The same pattern held for the Internalizing Risks scale–BASC-2 Internalizing scale correlations. Such findings offer strong evidence of the validity of the DESSA SEIR scale scores.

Finally, there is also substantial evidence that the DESSA SEIR scales provide scores that are relatively free of bias and translate to concern levels with little or no disparate impact on subgroups of students. Although it was not possible to identify risk scales composed entirely of items that met all of the a priori criteria (e.g., no local dependence, no statistical evidence of DIF), relatively few items were retained that may represent minor degrees of DIF. Most important, the total raw scores for the scales did not result in any discernable pattern that indicates disproportionate impact of the CREBCF cut points across race/ethnicity or sex groups. This provides further evidence that scores from the DESSA SEIR scales are fair and have substantial consequential validity.

In conclusion, there is substantial evidence for the reliability, validity, and fairness of the DESSA SEIR scale scores. Thus, in combination with the DESSA screeners and assessments, these assessment tools offer users a unique combination of positive and negative behavior scales that are efficient and provide actionable scores.

## Future Directions

We believe that the evaluation of reliability and validity evidence for scores from all assessments, including the DESSA SEIR, is an ongoing process. The evidence reported in this chapter was generated using data from the national standardization of the original SSIS scales and from a study using the DESSA SSR and DESSA SEIR (Johnson & Parks, 2024). As such, one important future direction is to further evaluate the reliability and validity of scores from the DESSA SEIR scales relative to their primary intended purpose—screening for emotional or behavioral concerns that impact students’ mental health. These future studies must include diverse samples of students and must examine relationships between scores on the DESSA SEIR scales and those from comprehensive assessments of emotional and behavior concerns. In addition, future studies are necessary to examine the relationships between scores on the DESSA SEIR scales and external criteria such as formal diagnoses of emotional or behavioral disabilities, receipt of intervention services within (and beyond) school settings, school avoidance/dropout, and risk-taking behaviors. Perhaps most important and consistent with our commitment to ensuring fairness, future reliability and validity studies must examine differences in evidence across race/ethnicity, gender, and socioeconomic status. While we are beginning to undertake such future studies, we welcome other researchers’ published data from their own independent studies examining the reliability and validity of scores from the DESSA SEIR scales.



# REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. AERA.
- Bruhn, A. L., Woods-Groves, S., & Huggle, S. (2014). A preliminary investigation of emotional and behavioral screening practices in K–12 schools. *Education and Treatment of Children, 37*, 611–634.
- Cohen, J. (1988). *Statistical power analysis in the behavioral sciences* (2nd ed.). Erlbaum.
- Dowdy, E., Doane, K., Eklund, K., Dever, B. (2013). A comparison of teacher nomination and screening to identify behavioral and emotional risk within a sample of underrepresented students. *Journal of Emotional and Behavioral Disorders, 21*, 127–137.
- Elliott, S. N., Anthony, C. J., DiPerna, J. C., Lei, P., & Gresham, F. M. (2020). *SSIS SEL Brief + Mental Health Scales*. SAIL CoLab, LLC.
- Elliott, S. N., & Gresham, F. M. (2008). *Social Skills Improvement System Performance Screening Guide*. Pearson Assessments.
- Elliott, S. N., Hwang, Y., & Wang, J. (2019). Teachers' ratings of social skills and problem behaviors as concurrent predictors of students' bullying behavior. *Journal of Applied Developmental Psychology, 60*, 119–126.  
<https://doi.org/10.1016/j.appdev.2018.12.005>
- Elliott, S. N., Lei, P. W., Anthony, C. J., & DiPerna, J. C. (2021). Screening the whole social-emotional child: Expanding a brief SEL assessment to include emotional behavior concerns. *School Psychology Review, 52*(1), 15–29.
- Forness, S. R., Kim, J., & Walker, H. M. (2012). Prevalence of students with EBD: Impact on general education. *Beyond Behavior, 21*, 3–9.
- Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System*. AGS.
- Gresham, F. M., & Elliott, S. N. (2008). *Social Skills Improvement System Rating Scales*. Pearson Assessments.
- Gresham, F. M., & Elliott, S. N. (2017). *Social Skills Improvement System Social-Emotional Learning Edition Rating Forms*. Pearson Assessments.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Sage.
- Johnson, E. S., & Parks, T. W. (2024). Promoting student well-being through a risk and resilience framework: Implications for screening. Aperture Education | Riverside Assessments, LLC.
- LeBuffe, P. A., Shapiro, V. B., Robitaille, J. L., & Naglieri, J. A. (2018). *Devereux Student Strengths Assessment High School Edition*. Riverside Assessments, LLC.
- Meade, A. W. (2010). A taxonomy of effect size measures for the differential functioning of items and scales. *Journal of Applied Psychology, 95*(4), 728–743.

- Petrillo, J., Cano, S. J., McLeod, L. D., & Coon, C. D. (2015). Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: A comparison of worked examples. *Value in Health*, 18, 25–34.
- Reynolds, C., & Kamphaus, R. W. (2004). *Behavior assessment system for children* (2nd ed.). NCS Pearson, Inc.
- Robitaille, J. L., LeBuffe, P. A., Johnson, E. S., Naglieri, J. A., & Shapiro, V. B. (2023). *Devereux Student Strengths Assessment Middle School Edition Student Self-Report* (DESSA-MSE SSR). Riverside Assessments, LLC.
- Robitaille, J. L., Johnson, E. S., LeBuffe, P. A., Naglieri, J. A., & Shapiro, V. B. (2024). *Devereux Student Strengths Assessment* (2nd ed., DESSA 2). Riverside Assessments, LLC.
- Robitaille, J. L., Johnson, E. S., LeBuffe, P. A., Naglieri, J. A., & Shapiro, V. B. (2025). *Devereux Student Strengths Assessment High School Edition Student Self-Report* (2nd ed., DESSA 2 HSE SSR). Riverside Assessments, LLC.
- Salvia, J., Ysseldyke, J. E., & Bolt, S. (2010). *Assessment in special and inclusive education* (12th ed.). Cengage Learning.
- Snyder, T. D., de Brey, C., & Dillow, S. A. (2018). *Digest of Education Statistics 2016* (NCES 2017-094). NCES, IES, U.S. Department of Education.
- Tay, L., Meade, A. W., & Cao, M. (2015). An overview and practical guide to IRT measurement equivalence analysis. *Organizational Research Methods*, 18, 3–46.
- United States Department of Education. Office for Civil Rights. (2016). 2013-14 Civil Rights Data Collection (CRDC). Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2017-02-16.  
<https://doi.org/10.3886/E100445V1>
- Vidair, H. B., Sauro, D., Blocher, J. B., Scudellari, L. A., & Hoagwood, K. E. (2014). Empirically supported school-based mental health programs targeting academic and mental health functioning: An update. In H. M. Walker & F. M. Gresham (Eds.), *Handbook of evidence-based practices for emotional and behavioral disorders: Applications in schools*. (pp. 15– 53). Guilford.





## Appendix A

# **DESSA SEIR RAW SCORES AND CRITERION-REFERENCED CONCERN LEVELS**

**Table A.1. DESSA SEIR Student Form (Grades 6–8) Internalizing Risks and Externalizing Risks Raw Scores**

**Student Form Grades 6–8**

95% Confidence Intervals, Percentile Ranks, and  
Corresponding Criterion-Referenced Concern Levels

Internalizing Risks				Externalizing Risks			
Raw Score	Confidence Interval	Percentile Rank	Concern Level	Raw Score	Confidence Interval	Percentile Rank	Concern Level
0	0–3	13	No Concern	0	0–2	13	No Concern
1	0–4	35	No Concern	1	0–3	36	No Concern
2	0–5	52	No Concern	2	0–4	52	No Concern
3	0–6	63	No Concern	3	1–5	64	No Concern
4	1–7	73	No Concern	4	2–6	75	No Concern
5	2–8	81	No Concern	5	3–7	86	Possible Concern
6	3–9	87	No Concern	6	4–8	92	Possible Concern
7	4–10	91	Possible Concern	7	5–9	95	Concern
8	5–11	93	Possible Concern	8	6–10	96	Concern
9	6–12	95	Concern	9	7–11	97	Concern
10	7–13	97	Concern	10	8–12	98	Concern
11	8–14	98	Concern	11	9–13	99	Concern
12	9–15	99	Concern	12	10–14	> 99	Concern
13*	10–15	99	Concern	13*	11–15	> 99	Concern
14*	11–15	> 99	Concern	14*	12–15	> 99	Concern
15*	12–15	> 99	Concern	15*	13–15	> 99	Concern

\* Possibly invalid

\* Possibly invalid



**Table A.2. DESSA SEIR Student Form (Grades 9–12) Internalizing Risks and Externalizing Risks Raw Scores**

**Student Form Grades 9–12**

95% Confidence Intervals, Percentile Ranks, and  
Corresponding Criterion-Referenced Concern Levels

Internalizing Risks				Externalizing Risks			
Raw Score	Confidence Interval	Percentile Rank	Concern Level	Raw Score	Confidence Interval	Percentile Rank	Concern Level
0	0–2	16	No Concern	0	0–2	13	No Concern
1	0–3	39	No Concern	1	0–3	34	No Concern
2	0–4	51	No Concern	2	0–4	49	No Concern
3	1–5	61	No Concern	3	1–5	61	No Concern
4	2–6	70	No Concern	4	2–6	71	No Concern
5	3–7	78	No Concern	5	3–7	79	Possible Concern
6	4–8	84	No Concern	6	4–8	86	Possible Concern
7	5–9	89	Possible Concern	7	5–9	90	Concern
8	6–10	93	Possible Concern	8	6–10	94	Concern
9	7–11	96	Concern	9	7–11	96	Concern
10	8–12	97	Concern	10	8–12	98	Concern
11	9–13	98	Concern	11	9–13	99	Concern
12	10–14	99	Concern	12	10–14	99	Concern
13*	11–15	> 99	Concern	13*	11–15	99	Concern
14*	12–15	> 99	Concern	14*	12–15	> 99	Concern
15*	13–15	> 99	Concern	15*	13–15	> 99	Concern

\* Possibly invalid

\* Possibly invalid

An abstract graphic design featuring a dark blue background with white geometric lines. These lines form a series of interconnected triangles and polygons, creating a sense of depth and structure. The lines vary in thickness and orientation, some running parallel and others intersecting at sharp angles.

## *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.ApertureEd.com](http://www.ApertureEd.com).