Review of Submitted Prekindergarten Assessments

Methods and Recommendations

Final Report

June 2023

Report Authors

Kristi L. Santi, Ph.D.

Professor

Special Populations

Educational Leadership and Policy Studies

Coleen Carlson, Ph.D.

Associate Research Professor

Associate Director of the Texas Institute

for Measurement, Evaluation, and Statistics

Table of Contents

List of Tables

List of Figures

Executive Summary

The University of Houston (UH) was awarded through the Texas Education Agency (TEA) Early Childhood Assessment Support (ECAS) grant to provide an independent evaluation and recommendation of prekindergarten assessments for monitoring student progress for potential inclusion on the Texas Education Agency (TEA) Commissioner's List of Approved Prekindergarten Assessment Instruments. This project was a collaborative effort between the TEA, the UH, and the Region 4 Education Service Center (ESC). The Request for Information (RFI) to solicit information from Publishers on the criteria developed by the TEA, the UH, and the Region 4 ESC was posted twice on the TEA website. The UH team of panelists provided reviews of all submitted assessments. Drs. Carlson and Santi provided fidelity checks on the information reviewed. The panel discussed the instruments in detail to determine which assessments met the criteria listed in the RFI. The UH, the Region 4 ESC, and the TEA teams met three times to review the process, the required documentation, and finally, the final working drafts of the documents and the data behind the recommendations.

Method

The Panel

The University of Houston convened an expert panel for the review process. The panel consisted of experienced psychometricians, prekindergarten teachers who are specialists in the content areas of Emergent Literacy, Social and Emotional and Physical Development (a.k.a. Health and Wellness), and Mathematics. Over half the panelists (8 of the 11) are multilingual, and seven are fluent in Spanish (see Appendix A – Panelists).

The UH faculty leads (Drs. Carlson and Santi) convened the panel and developed the training materials, the matrix and glossary, and the scoring system. The UH team organized the three meetings with the TEA and Region 4 ESC. Finally, the leads developed the master scoring sheet, the recommendations for the prekindergarten report, and the final report.

The Process

In November 2022, the UH team started working with the Region 4 ESC team to delineate a timeline for the tasks required to execute the project. All documents (see Table 1) were to be formatted similarly to the previous review of prekindergarten assessments for consistency. The UH team and Region 4 ESC team met once a week during this planning period to review and update timelines, document details, and set future meeting dates. The first RFI was posted in the December 30, 2022 issue of the Texas Register (47 TexReg 9057), with a deadline for submission on January 31, 2023. The notice of the extension for the submission of Prekindergarten Progress Monitoring Instruments was published in the March 17, 2023 issue of the Texas Register (48 TexReg 1583). The review process started after all submissions were checked for the inclusion of required materials.

Table 1: List of Deliverables for the Project

- Request for Prekindergarten Progress Monitoring Instrument Information (RFI)
- Commissioner's List of Approved Prekindergarten Assessment Instruments - Recommendation
- Panelist's Biographies
- Matrix and Glossary
- Final Report
- Master Scoring Sheet
- Publisher/Vendor Questions

Questions from Publishers

After the RFI was published, the UH team fielded a few questions. The emails are directly copied in the next paragraph. The second email was submitted by three Publishers, and therefore only one email is included as the wording is almost identical between the three emails.

First emailed question:
> *Thank you for the email.*
> *1. For clarification, per the TEA PreK Assessment document there will be two separate submissions for the Prekindergarten Submission form.*
> > *1. The first will be an email submission with the Prekindergarten Submission form with all attachments embedded sent to klsanti@uh.edu.*
> > *2. The second submission will require Two thumb drives containing the Prekindergarten Submission Form and technical documents, the electronic versions of the submission form and all supporting documentation, and three paper copies of the form and attachments, in addition to access*

information or software for online/electronic instruments. Is this information correct regarding the submission?

RESPONSE #1 *– There is one submission process, and the submission deadline is January 31, 2023, at 3:00 PM CST. The required documents are 1) the completed PDF of the submission form, 2) the documents requested on that PDF, and 3) the copies of the assessment (or a link to the website where a review of all assessment materials may be reviewed by the panelist).*

2. Per the TEA PreK Assessment document Section 2.1 could you please specify what is meant by electronic versions of the submission form? Does this mean a PDF version or Word document for the thumb drives?

RESPONSE #2 *- If all required and necessary documents can be submitted via the electronic submission form, you do not need a hard copy submission or the two extra thumb drives.*

3. Per the TEA PreK Assessment Instrument Submission Form Section B. Content Social & Emotional Development (a.k.a. health & wellness), could you please elaborate on the skills assessed under the Self-Regulation- including attention?

RESPONSE #3 *– The skills and examples of the skills can be found on the TEA website. The document to review is the 2022 Texas Prekindergarten Guidelines PK4 Outcomes.*

There were three Publishers who asked the same question:

We need to email you the submission form but also mail 2 thumb drives with electronic versions of the submission form and all supporting documentation. You do not require any paper version of the documents at all.

RESPONSE: *If all required and necessary documents can be submitted via the electronic submission form, then you do not need a hard copy submission or the two extra thumb drives.*

Background on Prekindergarten Assessments

A review of the literature and other state approaches to assessment in Prekindergarten was conducted. *The Center for Standards, Assessment, and Accountability (CSAA) State of the States, PreK Assessment Report* (WestEd, 2023) was resourced to assist with this review. CSAA updates this report approximately every three years, and it is a comprehensive review of the data available on the design, development, and implementation of pre-k assessments. In addition, the Educational Testing Service (Ackerman & Coley: ETS, 2012) also issued the *State Pre-K Assessment Policies: Issues and Status Report*, which detailed the policies and approaches to assessment for prekindergarten students. The ETS report is helpful for understanding the preferences for observation protocols, direct assessments, and work samples for determining growth over time.

While some states have changed some assessments/requirements since the 2017-2022 Texas Education Agency Commissioner's List of Approved Prekindergarten Assessment Instruments, the

changes have been to improve the assessments. For example, the 2022 Texas Prekindergarten Guidelines are aligned with the current research in the area of early childhood learning and development, and other states are aligned with Texas in scope and sequence of what is being assessed and at time points the assessments occur. It is also important to note that most states, including Texas, embed the Office of Special Education Programs Early Childhood Outcomes in their prekindergarten guidelines.

Background on Content

Preventing Reading Difficulties in Young Children (Snow et al., 1998) and the Report of the National Reading Panel: Teaching Children to Read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NRP, 2000) were two seminal research publications consulted for the review. In addition to seminal works, the U.S. Department of Education, Institute of Education Sciences, published *Preparing Young Children for School: Educator's Practice Guide* (Burchinal et al., 2022), which details recommendations for preschool education in the areas of social-emotional and executive skills, mathematics, and early literacy skills including vocabulary, letter knowledge, print recognition, and comprehension, was consulted.

The literature review focused on the domains and required skills in the Request for Prekindergarten Progress Monitoring Instruments filing document (see Table 2).

Table 2: Required Domains and Skills

Domain	Required Skills (*must be assessed*)	Additional Skills (*may be assessed*)
Social and Emotional Development (a.k.a. Health and Wellness)	• Gross and Fine Motor • Self-Regulation (behavioral and emotional, includes attention)	• Personal Health and Safety • Self-Concept • Relationships with Others • Social Awareness
Emergent Literacy - Language and Communication	• Listening Comprehension • Vocabulary	• Speaking (conversation) • Articulation • Sentences and Structure
Emergent Literacy - Reading	• Phonological Awareness • Alphabet Knowledge	• Comprehension of Text • Concepts of Print • Motivation to Read
Emergent Literacy - Writing	• Conventions in Writing	• Motivation to Write • Writing as a Process
Mathematics	• Number Sense • Classifications and Patterns	• Joining and Separating • Geometry and Spatial Sense • Measurement

Background on Psychometrics

Two seminal publications were consulted for this review, The Handbook of Research Synthesis (Cooper & Hedges, 1994) and the Standards for Educational and Psychological Testing (2014). In addition, three main categories were reviewed by the panelists.

Reliability. Reliability is the extent to which the measure of a construct is consistent or dependable. A construct is a trait, an ability, or a behavior that cannot be seen. The trait, ability, or behavior is thought to be responsible for a student's response to a test question. Internal consistency reliability is a measure of consistency between items intended to measure the same construct. Test-retest reliability is a measure of reliability obtained by administering the same test twice over a period of time to a group of individuals to evaluate the test for stability over time. Inter-rater reliability is a measure of consistency used to evaluate the extent to which different judges or raters agree in their assessment decisions. Inter-rater reliability is important as different individuals will not necessarily interpret answers or representative behaviors in the same way, and raters may disagree on how well certain responses, behaviors, or materials demonstrate the knowledge of the construct or skill being assessed.

<u>Note</u>: None of the submitted instruments had alternate versions, so alternate form reliability was removed from the ratings and calculations.

Validity. Validity refers to how well the components of an assessment measure what they are intended to measure. This review focuses on construct validity (i.e., concurrent, convergent, discriminative) and predictive validity.

Generalizability. Generalizability is the degree to which the results can be applied to a broader context. Thus, for this to happen, the sample being tested should accurately represent the broader population.

Growth/Improvement: Growth or Improvement (a.k.a. responsiveness or sensitivity to change) refers to the ability instruments have to measure meaningful differences in the constructs of interest. An instrument is said to be sensitive to growth when it measures increases and decreases in the construct measured such that scores increase with skill improvement and decrease when skills decline.

Receipt of Assessments

Round one submission was due on January 31st by 3:00 PM, and round two was due on April 14th by 3:00 PM. All publishers provided the materials using a secure website link on their site. The materials (except for the actual online assessments) were downloaded and housed in the UH IT-approved secure website OneDrive. The panelists had access to this folder during the entire process. The panelists also used this website to upload their completed reviews securely. The list of submitted assessments for the review is found in Table 3.

Table 3: List of Assessments Reviewed by Panel

Assessments Reviewed (in alphabetic order and as stated by Publisher on RFI)
CIRCLE Progress Monitoring System
Cognitive ToyBox for Schools
COR Advantage
CPALLS+STEM (CIRCLE Progress Monitoring administered by Tango)
Developmental Indicators for the Assessment of Learning™, 4th Edition (DIAL™ -4)
Frog Street Assessment – Criterion Referenced Assessment
GOLD®
LION for Prekindergarten
myIGDIs Early Literacy + myIGDIs Profile of Preschool Learning and Development Readiness (ProLADR)
Ready, Set, K!
The Work Sampling System®, Fifth Edition (WSS)

Review Process

Training

The trainings were held in two phases, one for the initial review and one for the psychometrics and key features review. All trainings were conducted via Zoom.

First Training. The initial training was conducted with the panelists to discuss the background and purpose of the review, the process (see Figure 1) and deadlines, how to document the administrative features of the instrument (see Table 4), and how to complete the initial review tool (see Table 5). The training also covered how to view the materials in the OneDrive folder.

Figure 1: The Process

Second Training. The second training focused on completing the full instrument review (see Figure 2). The first step was a review of the content features and scoring (see Table 4), focusing specifically on the alignment of required skills found in the *2022 Texas Prekindergarten Guidelines: PK3 and PK4 Comprehensive Guide*. The non-required additional skills listed in Table 2 were also rated. However, those scores are not included in recommendation calculations. The ratings of non-required skills are used only for presentation purposes to provide readers with additional context regarding the instrument. In addition, the training also focused on scoring the instrument psychometrics, including evidence of reliability, validity, generalizability, and growth. The training also covered the materials needed to complete the review, where to find them, and how to use them.

Figure 2: Full Review

Overview

The reviews covered three primary areas: Administrative features, Content features, and Psychometric features (see Table 4). More specific information regarding the review process is provided below. Additional documents related to the review can be found in the Appendices.

Appendix	Title
A	Review Panel for the Evaluation of Prekindergarten Assessment Instruments
B	Request for Prekindergarten Progress Monitoring Instrument Information (RFI)
C	Commissioner's List of Approved Prekindergarten Assessment Instruments: Matrix and Glossary
D	Commissioner's List of Approved Prekindergarten Assessment Instruments: Master Scoring Sheet
E	Commissioner's List of Approved Prekindergarten Assessment Instruments: Recommendation

Table 4: Review Areas

Administrative Features	Content Features	Psychometric Features
• Title • Publisher • Languages • Age Levels • Time points • Format: Physical • Format: Administrative • Training Requirements • Scoring Method • Available Scores • Score Reports • Time per student • Price per student	• Depth of 2022 Prekindergarten Outcome Coverage **<u>Scored for Review Calculations:</u>** • Gross and Fine Motor Skills • Self-Regulation (behavior, emotion, and attention) • Listening Comprehension • Vocabulary • Phonological Awareness • Alphabet Knowledge • Conventions in Writing • Number Sense • Classifications and Patterns **<u>Rated for Presentation Purposes Only:</u>** • Personal Health and Safety • Self-Concept, Relationships with Others, Social Awareness • Speaking (Conversation), Articulation, Sentences and Structure • Comprehension of Text, Concepts of Print, Motivation to Read • Motivation to Write, Writing as a Process • Joining and Separating, Geometry and Spatial Sense, Measurement	• Reliability • Validity • Growth/Improvement • Generalizability • Feasibility

Administrative Features

For the initial review, panelists recorded general information about the assessment, including:

- Title
- Publisher
- Languages
- Age Levels
- Time points
- Format: Physical
- Format: Administrative

- Training Requirements
- Scoring Method
- Available Scores
- Score Reports
- Time per student
- Price per student

Following this, raters examined the submissions for the required features in Table 5. The information presented on the RFI was crossed referenced with the assessment materials and the websites provided to complete these ratings. Multiple raters were assigned to each assessment, and upon completion of this initial review, information, and ratings were reviewed for fidelity.

Table 5: Initial Review Criteria

Initial Review Criteria	Yes or No
Intended for progress monitoring use in Prekindergarten	
Offered in English and Spanish	
Administered three times a year (beginning, middle, and end)	
Age levels appropriate for Prekindergarten	
Administration time is ≤ 20 min./student per domain	
Administration time for cumulative test is ≤ 100 min./student	
Individually administered to each student	
English/Spanish assess the same domains and subdomains	
Normative/technical data no more than 15 years old (i.e., 2008+)	
Intended for progress monitoring use in Prekindergarten	
Required skills within each 2022 Texas Prekindergarten domain are assessed	
Health and Wellness	
Emergent Literacy - Language & Communication	
Emergent Literacy - Reading	
Emergent Literacy - Writing Skill	
Mathematics	
Preferred but Not Required	
Parent Reports	
Instructional resources for teachers	
Instructional resources for parents	

Content Features

Each required skill from the *2022 Texas Prekindergarten Guidelines: PK3 and PK4 Comprehensive Guide* (henceforth 2022 Texas PK Outcomes) was scored based on the degree to which the instrument addresses the main "gist" of the skills described in the 2022 Texas PK Outcomes. Raters used the 2022 Texas PK Outcomes statements and took into consideration the child behavior examples provided.

Skills within each domain required to be considered for inclusion in the 2024-2027 Commissioner's List of Approved Prekindergarten Assessment Instruments are presented in **bold** in Table 6. All other skills were also rated, but those scores were not included in the scoring for inclusion in the Commissioner's List. Scores for non-required skills are presented for informational purposes only.

Content Features: Depth of Coverage Scoring Rubric

4	Very strongly addresses key aspects of the skill
3	Strongly addresses key aspects of the skill
2	Moderately addresses key aspects of the skill
1	Minimally addresses key aspects of the skill
0	Does not address key aspects of the skill

Table 6: Content Features: Prekindergarten Skills Coverage

2022 Prekindergarten Skills (required in bold)	Score	Included in Calculations
Health and Wellness		
Gross and Fine Motor Development	**(0-4)**	**Yes**
Personal Health and Safety	(0-4)	No
Self-Regulation	**(0-4)**	**Yes**
Self-Concept	(0-4)	No
Relationships with Others	(0-4)	No
Social Awareness	(0-4)	No
Emergent Literacy: Language and Communication		
Listening Comprehension	**(0-4)**	**Yes**
Vocabulary	**(0-4)**	**Yes**
Speaking (Conversation)	(0-4)	No
Articulation	(0-4)	No
Sentences and Structure	(0-4)	No
Emergent Literacy: Reading		
Phonological Awareness	**(0-4)**	**Yes**
Alphabet Knowledge	**(0-4)**	**Yes**
Comprehension of Text	(0-4)	No
Concepts of Print	(0-4)	No
Motivation to Read	(0-4)	No
Emergent Literacy: Writing		
Conventions in Writing	**(0-4)**	**Yes**
Motivation to Write	(0-4)	No
Writing as a Process	(0-4)	No
Mathematics		
Number Sense	**(0-4)**	**Yes**
Classification and Patterns	**(0-4)**	**Yes**
Joining and Separating	(0-4)	No
Geometry and Spatial Sense	(0-4)	No
Measurement	(0-4)	No

All content coverage scores for required skills were totaled, and the sum was divided by 4 (highest possible rating; see Table 7). The resultant average score represents the score each instrument received for content coverage.

Table 7: Content Features: Score Calculations

2022 Prekindergarten Domains and Skills	Highest Possible Score
Health and Wellness	
Gross and Fine Motor Development	4
Self-Regulation	4
Emergent Literacy: Language and Communication	
Listening Comprehension	4
Vocabulary	4
Emergent Literacy: Reading	
Phonological Awareness	4
Alphabet Knowledge	4
Emergent Literacy: Writing	
Conventions in Writing	4
Mathematics	
Number Sense	4
Classification and Patterns	4
Possible Content Coverage Score	36
Possible Content Coverage Average Score	4.00

Psychometrics

Each instrument was reviewed based on its psychometric properties. Psychometric properties reviewed fell into four broad categories: (1) reliability, (2) validity, (3) generalizability, and (4) evidence of sensitivity to growth/improvement.

Ratings of psychometrics focused on the assessment scores that were relevant to the required skills. Because different assessments create and apply scores in different ways, the exact number of assessment "scores" reviewed for the ratings differed across instruments. For each instrument, psychometric data were expected for scores representing required skills (e.g., specific skills scores (i.e., vocabulary, etc.) and scores incorporating the specific skill (i.e., language domain score, etc.). Because the number of scores varies across instruments, ratings for internal consistency, test-retest reliability, inter-rater reliability, construct validity, and predictive validity were averaged (separately), so all instruments received scores for each of these five aspects of reliability and validity on the same scale (0-3).

Reliability

Reliability is the consistency with which scores on an instrument measure an underlying construct. A construct is a trait, an ability, or a behavior that cannot be seen. The trait, ability, or behavior is thought to be responsible for a student's response to a test question.

The following expectations were outlined for raters:

- Not all aspects of reliability will be applicable to all assessments.

- All instruments should report some type of internal consistency.

- All instruments should report some type of test-test reliability.

- If instrument administrators make some type of determination in order to record a "score" for student responses or abilities (e.g., making ratings or indicating correct and incorrect), then some form of inter-rater reliability should be provided.

The rubric below (see Table 8) was used to rate internal consistency, test-retest, and inter-rater reliability separately.

Table 8: Internal Consistency, Test-retest, and Inter-rater Reliability Scoring Rubric

Score Value	Evidence
3 = Strong	Majority of estimates are greater than .80
2 = Moderate	Majority of estimates are between .70 and .79
1 = Minimal	Majority of estimates are below .70
0 = No evidence	Estimates are not provided
NA = Not Applicable	Estimates are not applicable to this assessment

All internal consistency, test-retest, and inter-rater reliability ratings were summed (separately) across components and averaged based on each instrument's total number of applicable components. The resulting average scores for internal consistency, test-retest, and inter-rater reliability were then summed and averaged (divided by 3), resulting in one overall score for reliability (see Table 9).

Table 9: Psychometric Features: Reliability Score Calculations

Overall Reliability Score	Possible Score
Internal consistency average score	0 – 3.00
Test-retest average score	0 – 3.00
Inter-rater average score	0 – 3.00
Possible Overall Reliability Total Score	0 – 9.00
Possible Overall Reliability Average Score	0 – 3.00

Validity

Validity refers to how well the components of an assessment measure what they are intended to measure. This review focused on construct validity (i.e., concurrent, convergent, discriminative) and predictive validity.

The following expectations were outlined for raters:

- All instruments should report some type(s) of construct validity.

- All instruments should report predictive validity.

The rubric below (see Table 10) was used to separately rate construct validity and predictive validity.

Table 10: Construct and Predictive Validity Scoring Rubric

Score Value	Evidence
3 = Strong	estimates with other relevant outcome measures are typically above .70
2 = Moderate	estimates with other relevant outcome measures are typically between .50 and .70
1 = Minimal	estimates with other relevant outcome measures are inconsistent and include correlations below .50
0 = No evidence	estimates are not provided

All construct and predictive validity ratings were summed (separately) across instrument components and averaged based on the total number of applicable components for each instrument. The resulting average scores for construct and predictive validity were then summed and averaged (divided by 2), resulting in one overall score for validity (see Table 11).

Table 11: Psychometric Features: Validity Score Calculations

Overall Validity Score	Possible Score
Construct validity average score	0 – 3.00
Predictive validity average score	0 – 3.00
Possible Overall Validity Total Score	0 – 6.00
Possible Overall Validity Average Score	0 – 3.00

Generalizability

Generalizability is the degree to which the sample(s) of students used to develop the assessment and establish psychometric properties is sufficiently large and demographically similar to the Texas student population.

For each norming or technical sample employed, raters scored the following:

Sample Size
3 = Large
2 = Moderate
1 = Limited
0 = Not provided

Representativeness
3 = Representative
2 = Relatively representative
1 = Not Very representative
0 = Not provided

Across all samples, the sample size and representativeness scores were summed and then averaged (divided by the total number of ratings). The resulting **Generalizability** score is interpreted on the following scale:

Generalizability Score

3 = Strong

2 = Moderate

1 = Minimal

0 = No evidence

Growth/Improvement

An instrument's ability to detect student growth or improvement over time is an important component of any progress monitoring assessment. Psychometric examination of growth (or improvement) can take many different forms. As such, one specific type of psychometric approach was not required over another. Instruments were rated based on the psychometrics resulting from valid psychometric approaches indicating the instrument is sensitive and detects change or improvement over time (e.g., ROC curve analysis, repeated measures analyses with student or group factors, t-tests of sample means of change scores, correlational analyses, etc.). Each instrument received a score of 0-3 for Growth/Improvement (see Table 12).

Table 12: Growth/Improvement Scoring Rubric

Score Value	Evidence
3 = Strong	Provides strong evidence of ability to detect growth/improvement over time.
2 = Moderate	Provides moderate evidence of ability to detect growth/improvement over time.
1 = Minimal	Provides minimal evidence of ability to detect growth/improvement over time.
0 = No evidence	Provides no evidence of ability to detect growth/improvement over time.

Administrative Feasibility

An additional important feature of an instrument is the feasibility of instrument administration and use. Feasibility was rated holistically in regard to teacher feasibility, student feasibility, and administrator feasibility. Reviewers rated each of these three areas of feasibility based on the "Components to consider" portion of Table 13 below.

Table 13: Psychometric Features: Feasibility Scoring Rubric

Feasibility Feature	Components to Consider	Rating Description	Rating
Teacher Friendly	• Administration time is manageable. • Administration training requirements are minimal. • Minimal additional materials are required for administration. • Scores and score reports are immediately available. • Scores are easy to interpret. • Score reports for parents are easily generated. • Aggregated score reports are easily generated (e.g., groups, skills, whole class).	3 = Strong 2 = Moderate 1 = Minimal 0 = No evidence	
Student Friendly	• Time requirement is manageable. • Directions and tasks are easy to understand. • Assessment is visually appealing. • Assessment is engaging.	3 = Strong 2 = Moderate 1 = Minimal 0 = No evidence	
Administrator Friendly	• Administration training requirements are minimal. • Scoring requires minimal time. • Scores are easy to interpret. • Score reports for parents are easily generated. • Aggregated score reports are easily generated (e.g., whole class, whole school).	3 = Strong 2 = Moderate 1 = Minimal 0 = No evidence	

The three feasibility ratings were averaged to create one overall feasibility score to be included in recommendation calculations. The three separate feasibility scores (before averaging) are presented on the individual instrument reports for information purposes (see Table 14).

Table 14: Psychometric Features: Feasibility Score Calculations

Overall Feasibility Score	Possible Score
Teacher Feasibility	0 – 3
Student Feasibility	0 – 3
Administrator Feasibility	0 – 3
Possible Overall Feasibility Total Score	0 – 9.00
Highest Possible Overall Feasibility Average Score	0 – 3.00

Inclusion for Recommendation

Overall Instrument Score

An overall score was calculated for each instrument using the overall scores for the different features described above. Specifically, all overall scores (see Table 15) were summed and then divided by 19 (total possible points). Each instrument's score was converted into a percentage score representing the percentage of total possible points.

The overall percentage for an instrument was 77%, and the lowest was 55%. The average percentage was 64%, and the standard deviation was equivalent to 8%. The cut point was set at 1 standard deviation below the mean or higher. As such, instruments scoring 55% or higher were included on the recommended list of assessments for prekindergarten progress monitoring (Appendix E). Given the scoring process described above, the specific manner by which specific instruments' overall scores were achieved differs, and each instrument has differing strengths and limitations.

A summary sheet for each instrument is presented in tandem with the overall recommended matrix in Appendix E to provide context for the reader. These one-page summaries present information about 1) the instrument in general (i.e., most of the administrative features listed above); 2) the depth 2022 Texas PK Outcomes coverage (required and non-required skills); 3) psychometric ratings with notes regarding psychometric strengths and limitations; 4) feasibility ratings with notes regarding feasibility strengths and limitations; and 5) comments on notable resources provided for teachers and families.

Table 15: Total Overall Instrument Score Calculations

Overall Instrument Score	Score
Overall Content Coverage (required skills)	0 – 4.00
Overall Reliability	0 – 3.00
Overall Validity	0 – 3.00
Generalizability	0 – 3.00
Growth/Improvement	0 – 3.00
Overall Feasibility	0 – 3.00
Total Sum	0 – 19.00
Average Overall Score (Total Sum/19)	0 – 3.00
Average (Total Sum/19)	0 – 100.00%

Practical Guidelines for Assessments

General

Assessments should be administered according to the directions in the teacher's guide, operation manual, and technical reports. The reliability and validity of assessments are based on how the test developer intended for the assessment to be delivered. Therefore, it is important to read the directions in the same manner to each child as it is the first time they hear the directions, even though the test administrator may be saying the directions many times. For observational instruments that rely on the collection of work samples, notes, and evidence of student performance, etc., it is important that test administrators are trained on authentic assessment pedagogy and the procedure outlined in the instrument. Establishing inter-rater reliability is especially important to ensure that test administrators are consistent in their determinations regarding indicators of performance and application of scoring rubrics.

Students should receive assessments in the same language (English, Spanish, or both) throughout the year to ensure consistency of measurement and accurate estimates of growth. Administering one language at one time point and another at a different time point (e.g., only Spanish at the beginning and middle of the year and only English at the end of the year) will make the comparison of scores across the year invalid. The importance of consistency of the measure across the year is required for assessment measurement to be valid, reliable, and accurate.

While it is important that students are not over-assessed, there are times when a school may feel the need to supplement the assessments. For example, phonological awareness measures may only ask a child to provide the sound for a few consonants and no vowels. In this case, a teacher may want to check all letter sounds informally. Likewise, alphabet knowledge measures may only assess word reading and no steps that lead up to word reading. In such cases, reviewing the information and resources provided for each assessment is recommended.

The reports aim to help the schools use the data to inform instruction and show growth over the school year. Therefore, the inferences made on the reports can only be made if the assessment was

administered as stated by the test developer. Therefore, teachers must be data detectives and work on targeted instruction to ensure students are ready for kindergarten. The low-stakes testing remains a guide to inform the teachers of modifications they can make to the teaching and learning aspects of education when done consistently throughout the year.

Instructional Implications

Health and Wellness

Gross and Fine Motor. Gross motor skills are those associated with balance, sitting, crawling, and walking. Fine motor skills include smaller body muscles such as the hand, wrist, and fingers. For example, a child's writing ability as part of the developmental stage relies on fine motor skill development. Developmentally, a child is expected to 'write' around three to four years of age, even though the writing may be what adults consider scribbles (Snow et al., 1998). Instructionally, teachers should work to engage students in activities that require the moving of large blocks or toys, stacking small cubes, or drawing pictures (NAEYC, n.d.).

Self-Regulation. According to the National Association for the Education of Young Children, a child's social and emotional well-being directly impacts academic performance in later years (Funk, 2018). This is corroborated by the work presented by Burchinal et al. (2022), which links executive function skills and social-emotional skills. Self-awareness, a part of socio-emotional skills, contributes to students' ability to identify the emotions that assist with executive function skills such as thinking flexibly. The overlap is the ability of a student to represent self-control/self-regulation. Instructionally, teachers should use a variety of interactions with the students that require the students to follow multistep directions, think quickly, or wait patiently. The interactions can become more complex as the academic year progresses (Burchinal et al., 2022).

Emergent Literacy – Language and Communication

Listening Comprehension and Vocabulary. The preschool students' ability to develop language and vocabulary is a crucial developmental step as they move forward in their academic setting. Seminal work presented by Snow et al. (1998) details the development of a child's metalinguistic skills and the

ability to use new vocabulary in their speech and understand stories read aloud. Vocabulary development is closely tied to reading comprehension and thus is an important skill to focus on early in a child's development (Snow et al., 1998; NRP, 2000). Instructionally, teachers should focus on three to five unique words within a specific category to focus on weekly (Burchinal et al., 2022). Instructionally, teachers should use a wide variety of text and start with book and print awareness activities. A read-aloud should be engaging as the teacher frequently stops to point out vocabulary words, ask various comprehension questions, and encourage discussion about the main idea or character.

Emergent Literacy – Reading

Phonological Awareness and Alphabet Knowledge. Phonological awareness (PA) is a critical precursor to successful reading in later grades. Studies have consistently shown that students who do not develop this skill early may not reach optimal reading levels in the later grades (Snow et al., 1998; NRP, 2000). Alphabet knowledge is more than just knowing the names of the letters; it is the child's ability to understand that the 26 letters represent sounds in the language (approx. 44 sounds). Thus, a child's ability to understand the relationships between the letters and the sounds will assist them in building fluency in later grades, which will, in turn, improve their reading comprehension (Snow et al., 1998; NRP, 2000). Instructionally, teachers should introduce the concept of PA by showing how words are made up of individual and different sounds. Rhyming, onset/rime, segmenting, and blending sounds are essential next steps in differentiating sounds. Students should also be taught the link between letter sounds and letter names simultaneously instead of teaching PA in isolation (Burchinal et al., 2022).

Emergent Literacy – Writing

Conventions in Writing. The adage that reading and writing are the different sides of the same coin is an important note when thinking about the importance of writing at an early age. There are many benefits of teaching writing at a young age (Byington & Kim, 2017). First, children who learn to write associate sounds with letters, hence learning the alphabet code (National Early Literacy Panel [NELP], 1998). Second, it is a generative process that assists with expression, a social-emotional link to learning. Third, writing in Prekindergarten is not solely focused on mechanics and composition but

should assist in developing orthographic knowledge. Finally, as noted above, fine motor skills are also required when students learn to write. One of the first ways students learn to write their names is from parents at home. Instructionally, teachers should talk with students and write down their conversations to help promote this portion of the foundational literacy skill set. It is also important to know what level of development each student is at and help them grow from that point (Cabell et al., 2013).

Mathematics

Number Sense and Classification and Patterns. Burchinal et al. (2022) define basic mathematical skills as counting, naming shapes, and creating patterns. The ability of preschool students to understand basic mathematical skills and mathematical language is a precursor for becoming proficient in mathematics in the later grades. Burchinal and colleagues recommend that teachers build skills beyond basic math, including numeracy, geometry, measurement, and patterning (p. 23). Instructionally, teachers should work on basic math skills using manipulatives, pictures, and objects. Engaging students in games that promote basic math skills can help provide an avenue for creating a healthy relationship with numbers.

Progress Monitoring

According to the Center on Multi-Tiered System of Supports (AIR, 2022), progress monitoring has three purposes, assess student performance, indicate improvement in performance, and evaluate the efficacy of instruction with a valid and reliable measure(s). Therefore, the assessments submitted for review were to meet the definition of progress monitoring and be available in English and Spanish.

References

Ackerman, D. J., & Coley, R. J. (2012). State Pre-K Assessment Policies: Issues and Status. Policy

Information Report. Educational Testing Service. Retrieved June 1, 2023, from

https://files.eric.ed.gov/fulltext/ED529449.pdf

American Institutes for Research (2022). Essential Components of MTSS. Retrieved from

https://mtss4success.org/essential-components/progress-monitoring.

Burchinal, M., Krowka, S., Newman-Gonchar, R., Jayanthi, M., Gersten, R., Wavell, S., Lyskawa, J.,

Haymond, K., Bierman, K., Gonzalez, J. E., McClelland, M. M., Nelson, K., Pentimonti, J., Purpura,

D. J., Sachs, J., Sarama, J., Schlesinger-Devlin, E., Washington, J., & Rosen, E. (2022). Preparing

Young Children for School (WWC 2022009). Washington, DC: National Center for Education

Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of

Education. Retrieved from https://whatworks.ed.gov/.

Byington, T. A., & Kim, Y. (2017). Promoting preschoolers' emergent writing. Young Children, 72(5).

Retrieved from https://www.naeyc.org/resources/pubs/yc/nov2017/emergent-writing.

Cabell, S.Q., Tortorelli, L.S., & Gerde, H.K., (May 2013). How do I write…? Scaffolding preschoolers'

early writing skills. The Reading Teacher, 66(8), pp. 650-659.

Center for Standards, Assessment, and Accountability (2023). State of the States, PreK Assessment

Report. Retrieved on June 1, 2023: https://csaa.wested.org/wp-

content/uploads/2023/01/CSAA-SoS_PreK_2023.pdf

Center on Standards and Assessment Implementation (2016). Pre-Kindergarten and Kindergarten Entry

Assessments: What assessments are states and other similar jurisdictions currently

administering at pre-kindergarten and kindergarten entry? WestEd. Retrieved June 1, 2023,

from https://csaa.wested.org/wp-content/uploads/2020/02/CSAI-Report_PreK-

KAssessments.pdf

Cooper, H., & Hedges, L. V. (Eds.). (1994). The handbook of research synthesis. Russell Sage

Foundation.

Funk, J. H. S. (2018). Promoting young children's social and emotional health. Young Children, 73(1). Retrieved from https://www.naeyc.org/resources/pubs/yc/mar2018/promoting-social-and-emotional-health.

National Association for the Education of Young Children (n.d.). https://www.naeyc.org/

National Reading Panel (U.S.) & National Institute of Child Health and Human Development (U.S.). (2000). Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Child Health and Human Development.

Standards for Educational and Psychological Testing (2014). Joint Committee of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education. Retrieved from https://www.testingstandards.net/uploads/7/6/6/4/76643089/9780935302356.pdf

Snow, C. E., Burns, S. M., & Griffin, P. (Eds.). (1998). Preventing Reading Difficulties in Young Children. Washington, DC: National Academy Press.

What Works Clearinghouse (2022). What Works Clearinghouse Procedures and Standards Handbook, Version 5.0. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). This report is available on the What Works Clearinghouse website at https://ies.ed.gov/ncee/wwc/Handbooks.

__Appendices__

A   Review Panel for the Evaluation of Prekindergarten Assessment Instruments

B   Request for Prekindergarten Progress Monitoring Instrument Information (RFI)

C   Commissioner's List of Approved Prekindergarten Assessment Instruments: Matrix and Glossary

D   Commissioner's List of Approved Prekindergarten Assessment Instruments: Master Scoring Sheet

E   Commissioner's List of Approved Prekindergarten Assessment Instruments: Recommendation