

Strengthening Texas Rising Star Implementation Study

Strengthening Texas Rising Star Implementation Study Summary

TWC funded a study to strengthen implementation of TRS to ensure consistency and accuracy in ratings across the state:

1. Analysis of the validity and reliability of the TRS assessment tool
2. Development of assessor/mentor certification program and ongoing reliability monitoring framework
3. Delivery of specialized TRS technical assistance (remote coaching) to inform quality improvement plans

Preliminary Data Collection Summary

- Recruited a balance of sites across socio-economic status in the greater Houston and Dallas areas
- Classrooms assessed:
 - Infant classrooms: 136
 - Toddler classrooms: 204
 - Preschool classrooms: 207
 - School Aged classrooms: 100

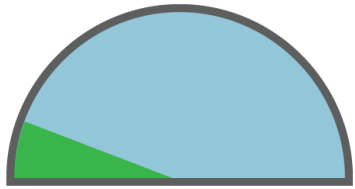
Recommendations Summary

Adjusting category weights

Improving the accuracy of scores

Strengthening quality improvement plans

Recommendation 1: Adjusting Category Weights

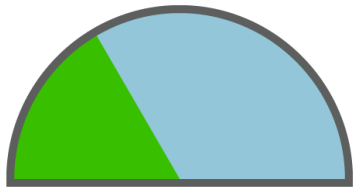


Less Related to Outcomes

Category 1

Category 5

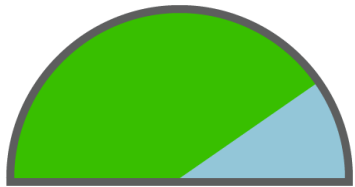
(Measures related to qualifications and written program policies)



Category 3

Category 4

(Measures related to physical environments, materials, and curriculum)



More Related to Outcomes

Category 2

Category 3, IFAL

(Measures capturing interactions between caregivers and children)

Outcomes

High quality experiences for children receiving public support

- Social and emotional development
- Oral language skills
- Early literacy skills
- Mathematics

Recommendation 2: Improving the Accuracy of Scores

The 3 indicators of reliability:

- Internal consistency (do the items within the same categories and sub-categories measure the same concepts)
- Inter-rater agreement (two people scoring side-by-side)
- Generalizability coefficient (does the team of raters draw similar conclusions, accounting for differences across the raters and sites assessed)

Key Findings that Inform Recommendations: Category 1

- Many of the key elements required for Category 1 could be scored using TECPDS individual profile reports of staff qualifications and training
- Several indicators within Category 1 items are difficult to consistently capture based on typical personnel files:
 - Years of experience within a TRS or TRS-recognized nationally accredited center
 - Years of experience within licensed or registered child care facility
 - Current job status (e.g., difficult to track transitions between full time, part time, substitute, volunteer)

Improving the Accuracy of Scores

Category 1

Category would be strengthened by:

- Removal of measure-level criteria that are inconsistently captured and available for review
- Increasing integrity of Category 1 scores by relying on TECPDS individual profile reports to reduce scoring errors and ensure authenticity of documents related to staff qualifications and training

Key Findings that Inform Recommendations: Category 2

- With rigorous training, assessment team was able to reach reliability on all items
- Items are strong in all three indicators of reliability
- The study was able to identify alternate scoring that results in greater reliability for frequency-based items (i.e., counts of behaviors)
- The study was able to identify alternate scoring for classroom ratio that is a more meaningful reflection of quality

Improving the Accuracy of Scores

Category 2

Category would be strengthened by changing how some of the items are scored:

- The study was able to identify alternate scoring that results in greater reliability for frequency-based items (i.e., counts of behaviors):
 - For example, instead of counting the number of times a caregiver does X, base their score on the caregiver's style across different settings (e.g., meal time, structured or unstructured activities).
- The study was able to identify alternate scoring for classroom ratio (e.g., using actual ratio rather than current points system based on enrolled children)

Key Findings that Inform Recommendations: Category 3

- Instructional formats and approaches to learning items are strong across all three indicators of reliability
- Lesson planning items as currently written are not providing a strong measure of curriculum (with the exception of preschool)
- The ratings system for lesson planning and curriculum does not capture a wide enough range of quality
- The special needs and respecting diversity items are too often excluded by assessors to reflect quality in these areas

Improving the Accuracy of Scores

Category 3

- Category 3 items are not providing a strong measure of curriculum as currently written, with the exception of *Instructional Formats and Approaches to Learning*
- Category would be strengthened by removing or rewriting items:
 - Recommend to remove lesson plans items for infant, toddler, and afterschool. Preschool items could be retained for further analysis, if desired.
 - Recommend to remove planning for special needs and respecting diversity items. The special needs and respecting diversity items are too often excluded to reflect quality in these areas.
- *Lesson Plans and Planning for Special Needs and Diversity* may be best addressed through quality improvement activities and monitoring

Key Findings that Inform Recommendations: Category 4

- The ratings system for nutrition does not capture a wide enough range of quality and contains too few items to be able to fully assess reliability
- Indoor learning environment items (across all ages) shows acceptable reliability
- Outdoor learning environment items shows acceptable reliability for all ages except infants

Improving the Accuracy of Scores

Category 4

- Category would be strengthened by removing or rewriting items related to nutrition
- Recommend minor adjustments to *Outdoor Learning Environment* items

Key Findings that Inform Recommendations: Category 5

- Measures are typically scored at the extreme high or low score
- Several of the indicators do not involve objective review of evidence, such as documents or observed behavior, and rely too heavily on self-report

Improving the Accuracy of Scores

Category 5

Category would be strengthened by adjusting the weight of this category within the overall star rating

Key Findings that Inform Recommendations: System-Level

Study data does not support automatic 4-star ratings for nationally accredited programs

- Of the certified providers assessed through the study, none were scored at a 4-star level

Improving the Accuracy of Scores System-Level

We recommend revising procedures for automatic certification of nationally accredited providers:

- Observe caregiver-child interactions and instructional formats in all classrooms and for all ages for nationally accredited providers prior to certification

Improving the Accuracy of Scores System-Level continued

Recommend a rigorous training and reliability monitoring process to ensure accurate star rating across the state:

- Train assessors and mentors to reliability using research-supported standards and procedures prior to initial data collection
- Require a monthly minimum of classroom observations per assessor to maintain reliability
- Require assessors to be accountable to a central body that certifies reliability and conducts routine reliability monitoring
- Require specific notetaking and documentation procedures to strengthen ratings and communication about scores with TRS staff and providers

Recommendation 3:

Strengthening Quality Improvement Plans

- Recommend establishing a quality improvement framework that ensures providers receive targeted technical assistance and professional development in alignment with their current star ratings:
 - Provide public-facing overviews, video exemplars, and sample documents of TRS standards on the TRS website to familiarize providers with the guidelines for certification
 - Automate CLI Engage course and lesson recommendations within TRS provider reports
 - Train and support TRS mentors to implement the quality improvement framework using a continuous improvement approach

Next Steps

