



Wisconsin Early Child Care Study

Wisconsin Early Child Care Study Findings on the Validity of YoungStar's Rating Scale

Executive Summary

Katherine Magnuson, PhD & Ying-Chun Lin, MSW

UW–Madison, School of Social Work and Institute for Research on Poverty

May 9, 2016

Wisconsin's child care quality rating and improvement system, YoungStar, was created to improve the overall quality of child care through two strategies: increasing parents' knowledge about the quality of early care and education providers and supporting providers' efforts to deliver high quality care. The YoungStar rating system assigns participating child care providers a star level from 1 to 5 based on objective indicators of quality in four domains: education and professional training, curriculum and learning environment, business and professional practices, and child health and well-being. YoungStar is administered by Wisconsin's Department of Children and Families, who, at the time of the study, had contracted the operation of the program's regional offices, technical assistance, and rating implementation to a consortium of three organizations (Celebrate Children's Foundation, Supporting Families Together Association, and Wisconsin Early Childhood Association).

In Wisconsin, the process of criterion indicator development and implementation was informed by other states' efforts and input from both experts and practitioners. An important goal for the Department of Children and Families has been to use empirical evidence to investigate whether the rating scale and the rating process work as intended to differentiate programs with respect to classroom quality and participating children's school readiness gains. The Wisconsin Early Child Care Study (WECCS) was undertaken to provide such an examination of the validity of YoungStar's rating scale. The study was designed to examine whether the rating scale is able to differentiate programs according to their levels of independently observed quality, and whether children who attend more highly rated programs gain more in terms of school readiness over the course of a school year than children attending programs rated at lower levels. This Executive Summary describes the findings



from two research reports that investigate these questions.

Study Design

WECCS staff recruited a sample of family and group child care providers participating in the YoungStar program in May of 2013 from the Northeast and Milwaukee YoungStar regions. Sampling was designed to facilitate comparisons across quality levels and ensure representation across communities. Programs were asked to participate in the study if they met basic eligibility requirements related to the age of children served and languages spoken. If the program administrator agreed to participate and at least four children between ages 3 and 5 had completed study parental consent forms, the program was considered enrolled in the study. Reflecting the distribution of programs in the state at that time, most programs in this study were in the 2 Star and 3 Star rating categories. In addition, most 4 and 5 Star programs participating in the study were rated through YoungStar's automated rating option for accredited programs.

In the fall of 2013, 887 children attending 157 Early Care and Education (ECE) programs were given parental consent to participate in direct child assessments. Starting in September, a battery of standardized developmentally appropriate assessments was administered by trained UW Survey Center field staff in participating ECE programs to collect the baseline information about children's school readiness. Additionally, surveys were administered to children's parents, teachers, and program administrators. Children's behaviors were assessed by teacher-report questionnaires during the same data collection period.

In the winter of 2013 after this first wave of data collection was completed, a subset of skilled field workers who had been conducting child assessments were trained to observe classroom quality using the Early Childhood Environment Rating Scale-Revised and the Family Child Care Environment Rating Scale-Revised (ECERS-R and FCCERS-R, jointly referred to as the Environment Rating Scales, ERS). Of the 157 sites enrolled in the study, 155 programs were part of the observational ratings, with valid observations of 239 classrooms and family provider settings being completed by April of 2014.

Starting in April of 2014, another round of direct child assessments was conducted in the same participating child care programs. At this time, 725 previously assessed children (82%) remained in the study and the same direct child assessments were administered by UW Survey Center field staff in order to evaluate children's gains in academic skills and learning related behaviors. Teachers were also given questionnaires to collect information about their ongoing experiences and children's behavior. Of 725 children having both fall and spring assessments on academic skills in 151 programs, about 87% were rated by teachers on their behaviors (656 in the fall and 644 in the spring).

The *Woodcock-Johnson-III Tests of Achievement Letter-Word Identification* (WJLW) was used to measure children's early reading, specifically their letter and word

identification skills (Woodcock, Schrank, Mather, & McGrew, 2007). The *Woodcock-Johnson-III Tests of Achievement Applied Problems* (WJAP) was used to assess children's early math skills (Woodcock et al., 2007). The *Bracken School Readiness Assessment-Third Edition* (Bracken) was used to measure children's understanding of basic academic concepts in five categories, including colors, letters, numbers/counting,



sizes/comparison, and shapes (Bracken, 2007). The *Test of Preschool Early Literacy* (TOPEL) subset 3: *Phonological Awareness* was used to assess children's early literacy skills specifically on word elision and blending abilities (Lonigan, Wagner, Torgesen, & Rashotte, 2007). The *Head-Toes-Knees-Shoulders* (HTKS) assessment was used to measure children's behavioral self-regulation. The *Preschool Learning Behaviors*

Scale (PLBS) was used to assess teacher reports of children's learning behaviors in the child care setting. The teacher version of the *Social Competence and Behavior Evaluation-Short Form* (SCBE-30) was used to assess children's Social Competence, Anger-Aggression, and Anxiety-Withdrawal with 30 items on a 6-point Likert scale (LaFreniere & Dumas, 1995; LaFreniere & Dumas, 1996).

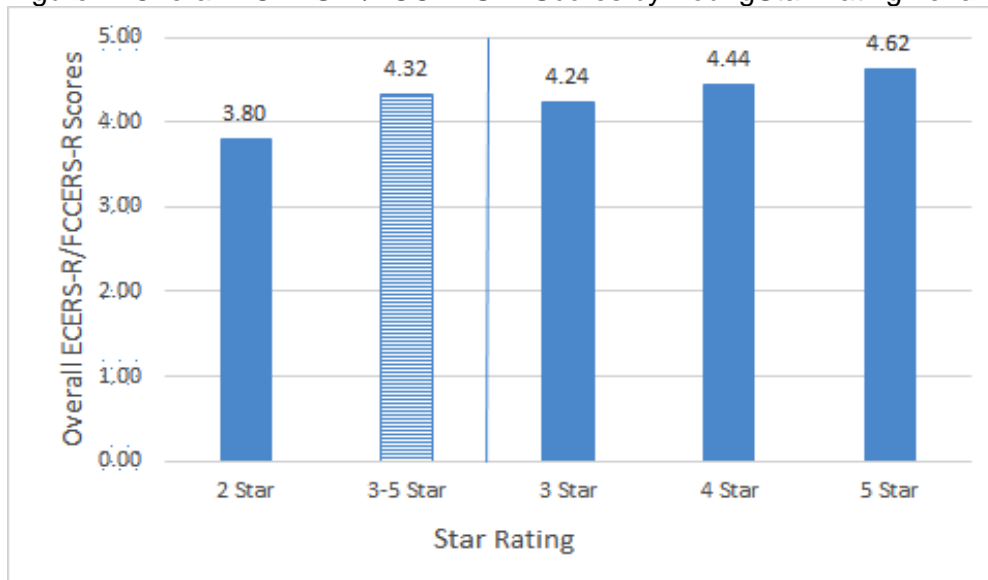
Findings and Discussion

Results from analyses of the data provide answers to two important questions about the validity of the YoungStar rating scale: whether the YoungStar rating predicted observed quality and whether the rating scale predicted children's spring levels of school readiness. With respect to the first question, analyses of data found that the YoungStar star rating level does differentiate among programs of varying observed quality (Figure 1). In particular, programs rated as 2 Star had scores on the global ERS that were about 0.5 points lower than programs rated as 3 Star or above. These differences were statistically significant and meaningful, representing a fairly large proportion of the variation in ERS score ratings (about half of a standard deviation). Yet, it is important to note that differences represent improvement within the range of minimal (ERS=3) to good (ERS=5) quality care.

Additionally, this study examined whether the YoungStar rating points that serve as the basis for the star level categorization also predicted a program's observed classroom quality. As expected, the total number of rating points within each domain were highly correlated with points in other domains, and thus each domain measures related aspects of program quality. Most importantly, the total number of points in each of the four rating domains predicted observed classroom quality. With respect to total rating

points, the difference in points between a 2 Star program (average 8.6 points) and 4 Star program (average 28.8 points) predicts a 1.2 point difference in ERS scores, which translates into quite a substantial effect, given the amount of observed variation across programs (over a standard deviation). This suggests that as a measure of child care quality, the YoungStar rating system has achieved validity.

Figure 1: Overall ECERS-R/FCCERS-R Scores by YoungStar Rating Level



Note: Significant difference in the ERS scores comparing low-quality programs (2 Star) to high-quality programs (3 Star or above).

Results related to whether YoungStar ratings predicted children's school readiness showed that, on average, children in YoungStar programs were meeting developmental expectations and learning an important range of skills during the time period of the study. This suggests that on average these children were likely to enter formal schooling ready to learn. However, analyses of the data did not support the conclusion that children in more highly rated YoungStar programs, whether measured by star level or total rating points, predicted children's school readiness in the spring of the study year. In general, the differences in school readiness between children in 2 Star and 3 Star or higher rated programs were not statistically significant (see Figures 2 and 3). This pattern of a lack of association between the YoungStar rating and children's outcomes was found across both academic skills and teacher's rating of children's learning related behaviors.

The finding that rating scales do not differentiate children's levels of early skills and behavior is consistent with most validation studies of other state or local QRIS rating systems that have examined child outcomes. Additionally, due to the comparatively small sample of 4 and 5 Star programs, the study was not designed to test for differences in observed quality, or child outcomes, between programs at the higher end of the rating scale. In addition, the results are most conclusive in showing that there are no meaningful differences between 2 and 3 Star programs.

Figure 2: Children’s School Readiness Assessment Scores, by YoungStar Rating Level

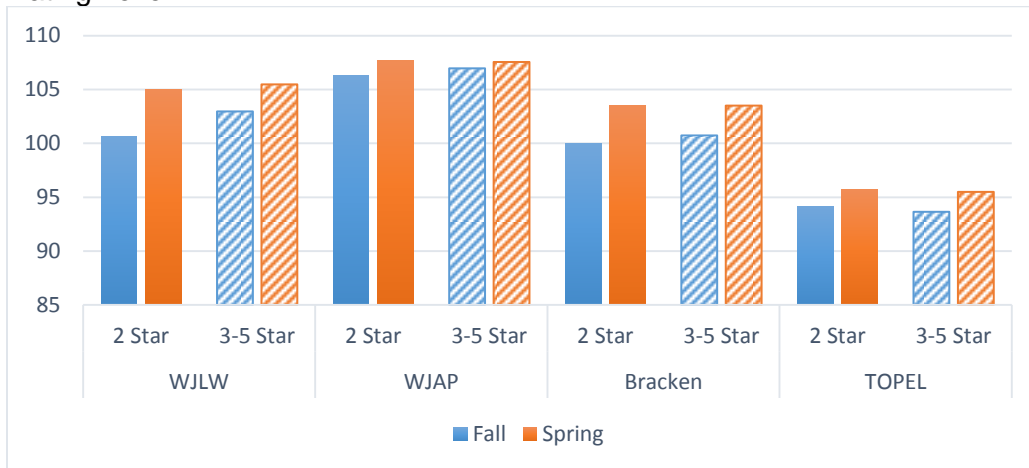
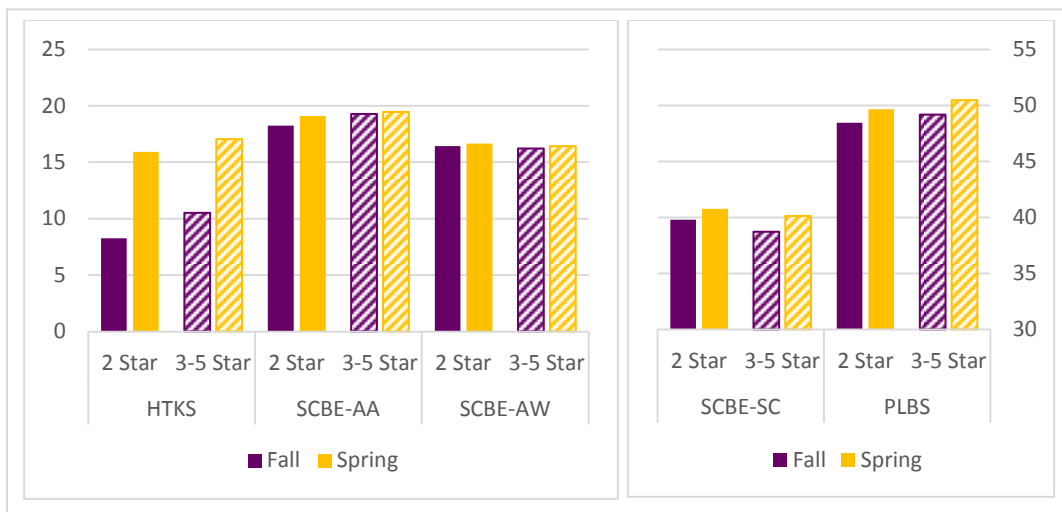


Figure 3: Teacher’s Rating of Children’s Learning Related Behaviors, by YoungStar Rating Level



Prior studies of other preschool settings indicate that the aspects of classrooms that are specifically aligned with particular skill-building activities and interactions, such as implementing specific curriculum or supportive instructional practices, produce increased gains in early academic skills and learning related behaviors. Thus, although efforts to improve broad dimensions of quality of child care environments are important for many reasons, they may not directly translate into higher levels of measurable school readiness skills and behaviors. Given the fact that YoungStar ratings are aligned with differences in observed child care quality but do not predict differences in specific domains of children’s early academic and behavioral skills, it may suggest that the broad dimensions of child care quality assessed by YoungStar are not the key inputs

related to the specific goal of furthering children's pre-academic skills and learning behaviors.

Taken together, the findings from the WECCS study are best understood as providing mixed support for the validity of the YoungStar rating scale. The scale was developed to identify where programs fell along a continuum of child care quality from a set of objective indicators that could be measured sufficiently well in a large at-scale system. Analyses of the WECCS data finds that YoungStar has succeeded in this task. However, analyses also suggest that higher quality child care, within the range of moderate to good care, is necessary but not sufficient for intentionally and specifically developing children's early school readiness. Additional policy efforts and new programmatic practices may be needed to ensure that ECE programs are able to facilitate and support children's early learning of academic and behavioral skills.

For more details about the study design and findings related to the analysis of observational quality ratings, see the WECCS full reports found at: [website1](#) and [website2](#) .