

Supporting young immigrant children and dual language learners in early
education: Examining the role of families in early childhood assessment

Saya Kawase

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Education

University of Washington

2022

Committee:

Holly S. Schindler, Chair

Nancy B. Hertzog

Program Authorized to Offer Degree:

College of Education

© Copyright 2022

Saya Kawase

University of Washington

Abstract

Supporting young immigrant children and dual language learners in early education: Examining the role of families in early childhood assessments

Saya Kawase

Chair of the Supervisory Committee:
Associate Professor Holly S Schindler
College of Education

Despite growing numbers of immigrant children and dual language learners (DLL) in the United States, their learning is often marginalized. Early childhood education (ECE) assessments, such as school readiness measures, are one example in which the strengths of immigrant and dual language learners (DLL) are missing. Further, many current ECE assessments face issues of reliability and validity with immigrant and DLL populations. In this thesis, I aim to address these issues by exploring two potential ways of creating family-based school readiness measures. I first use data from the National Household Education Surveys (NHES) Program of 2019 ($N=2,644$) to simulate a parent-reported school readiness measure with immigrant (and non-immigrant) families and find reliability and validity of evidence for this approach. Second, I propose to co-design with immigrant and DLL families a new family-based school readiness measure. Such a measure has the potential to facilitate immigrant and DLL children's learning and inform much needed changes to policies in order to leverage immigrant and DLL children's diverse assets.

TABLE OF CONTENTS

List of Figures.....	vi
List of Tables	vii
Chapter 1. Introduction	1
1.1 School Readiness for DLL Children.....	5
1.2 Family Involvement in ECE Assessment	11
1.3 Research Questions.....	13
Chapter 2. Validation of Family-based School Readiness Measure.....	13
2.1 Methods.....	15
2.2 Results.....	17
2.3 Discussion	20
Chapter 3. Validity and Reliability of School Readiness: Cross-group Comparison between Immigrant and Non-immigrant Families.....	21
3.1. Methods.....	21
3.2. Analyses	22
3.3. Results.....	23
3.4. Discussion	25
Chapter 4. Proposal for Co-designing and Implementing a Family-based School Readiness Measure	27
4.1. Introduction.....	27

4.2. Design Description	28
4.3. Methods.....	30
4.4. Anticipated Implications of the Proposed Study.....	33
Chapter 5. Discussion.....	34
References	38
Appendix A.....	45

LIST OF FIGURES

Figure 1.1. Revised conceptual framework of School Readiness (using subconstructs from Snow, 2006).....	6
Figure 2.1. Path diagram for a two-factor confirmatory factor model using KSEP (adapted from Quirk et al., 2014).....	14
Figure 2.2. Revised path diagram for a two-factor confirmatory factor model using National Household Education Surveys Program of 2019.	15
Figure 2.3. Graphical summary of two-factor confirmatory factor analysis (CFA) results using subset of data from National Household Education Surveys Program of 2019.....	19
Figure 3.1. The MIMIC model diagram from R semPlot package.....	24
Figure 3.2. Strong/scalar non-invariance (adapted from Wichet & Dolan, 2011).....	26
Figure 4.1. Conjuncture Map.....	30
Figure 4.2. Pictures of Prototype	32
Figure 5.1. Theory of Change.....	37

LIST OF TABLES

Table 1.1 Summary of Findings (regarding evidence of reliability and validity)	10
Table 2.1. Exploratory Factor Analysis with Oblimin Rotation.....	18
Table 3.1. Model Fit Statistics for Measurement Invariance Analyses from Immigrant and Non-immigrant Families.....	25

ACKNOWLEDGEMENTS

Writing this thesis is one of the most challenging, yet creative and rewarding experiences I have ever done. Without my supervisors, friends and families, I never would have finished. First of all, I would like to thank my supervisor Dr. Holly Schindler. Holly, thank you for always being available and present whenever I needed your support for my thesis and beyond. Your insightful feedback made me think of the importance of early childhood assessments more broadly. I also appreciate your uplifting comments on every draft and during our meetings, which allowed me to take deep breaths when I juggled writing, childcare and other household needs during the pandemic. I would also like to extend my gratitude to Dr. Nancy Hertzog. Nancy, thank you for your valuable and thoughtful comments on my thesis. I will keep refining my thinking of early childhood education, assessments and policy following our fruitful discussion during my thesis meeting.

I am also thankful to other UW faculty members. I have enjoyed the opportunity to complete a Graduate Certificate in Psychometrics and Applied Statistics (PAA) in addition to my degree, and I am particularly grateful to Dr. Elizabeth Sanders who helped me deepen my statistical knowledge. Liz, thank you for giving me extensive guidance over the past years.

In addition, I would like to thank my friends: Meg, Palmy, Yijie, Jessie, Sarena, Karlee, Makko, and Mandy, just to name a few. Thank you all for your support, especially for Meg and Karlee, who read and gave me feedback on my thesis; and Makko, who has been organizing parties and events for our children, which enable them to explore our diverse cultural and linguistic values.

Most of all, I want to express gratitude to my family: my son Yuhua (2 years old) and daughter Yuying (3 months old). Yuhua, your smile and “I love you” gave me the extra energy whenever I needed it. Yuying, you have already taken many classes at UW with me and hopefully enjoyed the intellectual lullabies in my belly. I look forward to creating more happiest moments with you both and seeing how you unfold your unlimited potentials! And finally, I would like to thank my husband Peng. Your love and selfless support allowed me and our children to be where we are now. Thank you for always being my biggest supporter and I am forever grateful to be next to you.

DEDICATION

This thesis is dedicated to my husband Peng and my children Yuhua and Yuying whose unconditional love, support and encouragement have inspired me to complete this project.

Chapter 1. INTRODUCTION

With a surge of immigrant populations in the United States, the children of immigrant families are poised to play critical roles in shaping the nation's future, including social and fiscal aspects. Immigrants include naturalized citizens, lawful permanent residents, asylees and refugees, people with temporary protected status, and unauthorized individuals. According to a report from the Migration Policy Institute (Esterine & Batalova, 2022), the number of immigrants in the United States has increased dramatically since 1970, in part due to increased immigration from Latin America and Asia. As of 2019, more than 44.9 million immigrants were living in the United States, which is approximately 13.7 percent of the total population, including 26 percent of all children in this country. Among the immigrant children, approximately 88 percent were born in the United States with immigrant parents, known as second-generation immigrant children, and the remaining 12 percent were born outside of the United States, known as first-generation immigrant children. Many immigrants to the United States speak a language other than English at home. Those children who are under the age of five and who have at least one parent or guardian who speaks a non-English language and are learning English while acquiring a home language, are referred to as dual language learners (DLL)¹. In this thesis, I will focus on immigrant children (both first- and second-generation), particularly DLL, who are shaping our nation through their diverse assets.

Despite their growing number, many immigrant and DLL children continue to face adversity in their lives, including, but not limited to, a lack of access to early childhood education (ECE). Decades of research on ECE indicates that high-quality ECE programs provide both short-

¹ The National Academies of Sciences, Engineering, and Medicine's (NASEM) report (2017) pointed out complex relationships between DLL and immigration and showed that not all immigrant children are DLL nor English learners (EL).

term (Camilli et al., 2010; Gormley & Gayer, 2005; Gormley et al., 2005; Gormley et al., 2008) and long-term benefits (e.g., Shonkoff & Phillips, 2000; Gormley et al., 2018; McCoy et al., 2017)². The short-term benefits include significant improvement in children's cognitive, social-emotional, and language skills, such as increased letter-word identification, spelling accuracy, and applied problem-solving (Gormley & Gayer, 2005; Gormley et al., 2008). Furthermore, McCoy et al. (2017)'s meta-analyses revealed medium- and long-term benefits of ECE, including significant reduction in special education placement and grade retention and an increase in high school graduation rates. Overall, ECE has the potential to contribute to children's learning trajectories.

In spite of its potential benefits, early education research has historically shown that immigrant children were less likely to have access to high-quality ECE programs (e.g., center-based ECE programs such as preschool) compared to their non-immigrant counterparts because immigrant families preferred family or home-based care (e.g., Halle et al., 2009; Matthews & Ewen, 2007). Following extensive examination on ECE accessibility, federal, state, and local policies have been implemented to address the gap in access through programs such as Head Start. Head Start programs support school readiness at no cost for more than a million children across the United States, especially for low-income immigrant (and other) families' children ages 3 to 5 (Head Start Program, 2020). These policies have led to higher enrollment rates of ECE, and a recent research article pointed out that children with immigrant mothers were now, in fact, more likely than children with non-immigrant mothers to be in center-based preschool programs (after controlling for variables such as socioeconomic status (SES), country of origin, and ethnicity) (Espinosa et al., 2017). This finding suggests that, currently, accessibility of ECE is not the main

² There are several studies pointing out that not all ECE programs are equally facilitative, and the magnitude of the benefits are different (e.g., Lipsey et al., 2015).

issue that immigrant children are facing; yet, they are still struggling to reach their potentials through ECE.

This begs the question: why are immigrant children still fighting to grow their assets despite the alleviation of the accessibility issue? One consideration is that, historically, immigrant children have been perceived through a deficit lens, such as the use of “achievement gaps” as a driver for ECE policy and practice. Achievement gaps are defined as “the disparities in standardized test scores between Black and White, Latina/o and White, and recent immigrant and White students” (Ladson-Billings, 2006, p.3). Following the *No Child Left Behind Act*, signed by President George Bush in 2002, the disparities in educational attainment across different groups of students (e.g., race/ethnicity, socioeconomic status, gender) have been highlighted using test scores. Through this collected data, educational achievement of immigrant children has been shown to lag behind that of monolingual English-speaking children (e.g., August & Shanahan, 2006). In ECE, several readiness gaps have also been highlighted between Hispanic/Latinx children and White children (Quirk et al., 2011; Quirk et al, 2013). This framing represents a deficit perspective towards immigrant children that emerges in public discourse even prior to school entry. Instead, we should be asking how immigrant children can thrive using their diverse assets within ECE systems.

Some ECE studies have shown preschool benefits to be particularly prominent among children of marginalized communities, such that the impact was largest among the Hispanic/Latinx immigrant children, followed by African American children (Gormley & Gayer, 2005). However, I argue that enrollment itself is not sufficient to maximize the benefit of ECE among immigrant

children. For instance, many of DLL come from immigrant families³ (Woods & Hanson, 2016), the types of offered engagement within an ECE program matters. For example, Rojas et al. (2021) showed that positive engagement with teachers, peers, and tasks among Spanish-speaking DLL children was associated positively with school readiness, including increased receptive and expressive vocabulary, phonological awareness, and print knowledge skills. That is, engagement opportunities play a crucial role in children's learning, and identifying DLL children and providing required language support is essential to maximize learning opportunities.

Unfortunately, DLL children are not gaining appropriate support across the country. Although Head Start provides recommendations to support DLLs' learning, Illinois is the only state which explicitly requires bilingual instruction if there are 20+ DLL with the same home language in a preschool program (Bipartisan Policy Center, 2020). Thus, most immigrant children strive to thrive in English-focused learning environments with little to no first-language support. There are several factors hindering the promotion of DLL children's language support in ECE. First, there is no systematic way to identify DLL children in ECE settings (Park & Pompa, 2021). Second, neither sufficient research nor practical applications (e.g., teacher training) have been conducted to promote DLL learning in ECE settings. In addition, it is difficult to assess DLL children's learning. That is, reliable and valid school readiness assessments for DLL children are yet to be developed. This lack of appropriate measurement has also hindered potential revisions to curriculum and instruction as well as policy and system-level change. Thus, this thesis focuses on new directions in school readiness assessments for DLL children to address the above issues.

³ While Woods and Hanson (2016) showed a high rate of bilingualism among immigrant families, the details of DLL in ECE have not been identified. See the issue in Park and Pompa (2021).

1.1 SCHOOL READINESS FOR DLL CHILDREN

Over the past decade, substantial attention has been paid towards school readiness⁴. The notion of school readiness emerged from the beginning of compulsory education within a maturation's perspective, that is, to examine whether children are matured enough to enter school. However, the current definition focuses on assessing an early phase of continuous development such as “the state of child competencies at the time of school entry that are important for later success” (Snow, 2006, p. 9). In other words, the construct of school readiness is used not only to assess a variety of skills at the time of kindergarten entry, but also to predict children's learning trajectory (see Figure 1.1. for the conceptual framework).

A series of school readiness measures have been developed and introduced by states following early childhood education policies, namely, Race to the Top - Early Learning Challenge (RTT-ELC) awards (2013-2016). These measures include Washington Kindergarten Inventory of Developing Skills (WaKIDS), Oregon Kindergarten Assessment, Maryland/Ohio Kindergarten Readiness Assessment and North Carolina Kindergarten Entry Assessment. Even by regional levels, school readiness measures have been developed, such as Kindergarten Student Entrance Profile (KSEP), which is commonly used in central California, and is a central focus of this thesis. While many of the tests share similar subconstructs, it is important to point out that the factors each assessment addresses are quite varied (e.g., from having two subsets of skills, i.e., cognitive and socioemotional skills in KSEP, to up to nine in Illinois Kindergarten Individual Development Survey).

⁴ School readiness is also known as kindergarten entry assessments (KEA) but I use school readiness throughout this thesis.

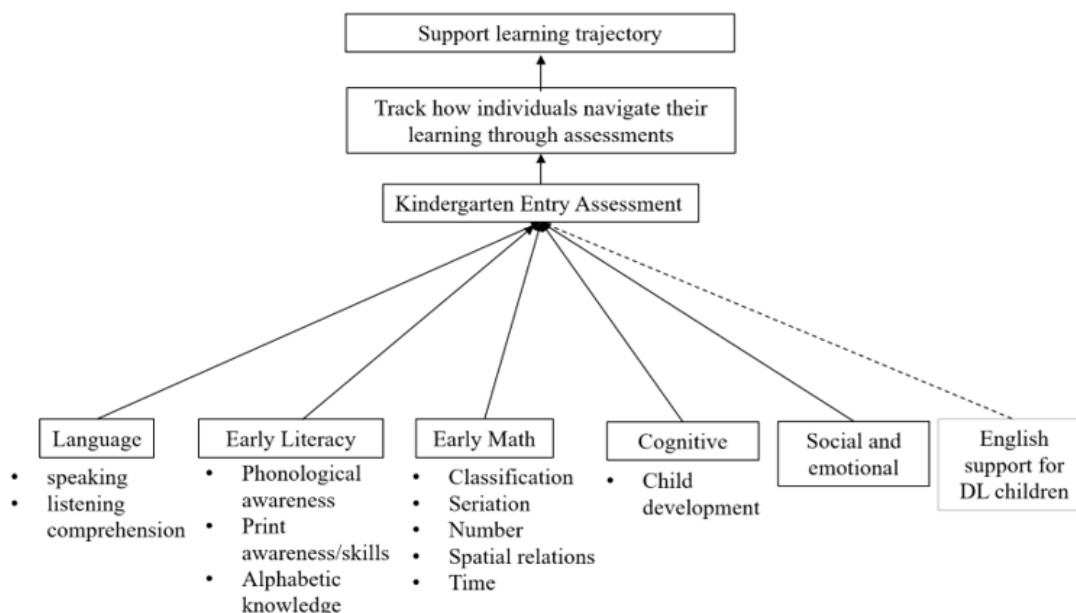


Figure 1.1. Revised conceptual framework of School Readiness (using subconstructs from Snow, 2006).

Prior research suggests that some school readiness measures, such as KSEP, could reliably be used for immigrant children. This measure was developed by a group of educational psychology researchers at UC Santa Barbara in collaboration with Santa Maria-Bonita School District and addresses the current notion of school readiness (i.e., observation-based rating to assess kindergarten students' knowledge and skills to support their learning trajectory). In particular, researchers have examined the use of KSEP with students who may have limited access to early childhood education, such as low-income Latinx children. For example, Quirk et al. (2014) observed a high internal consistency in two factors in KSEP: Socioemotional and Cognitive skills, for Latinx children in central California. Edyburn et al. (2017) showed that KSEP was able to demonstrate measurement invariance for kindergarten and prekindergarten students, suggesting that KSEP could potentially inform necessary interventions prior to kindergarten entry. Furthermore, a series of longitudinal studies suggest KSEP could predict learning trajectories when

compared to Grade 2 achievement (Quirk et al., 2013) as well as Grade 5 literacy achievement (Quirk et al., 2016).

While the above studies indicate the possible effective use of school readiness measures such as KSEP for DLL children, there are several validity and reliability issues that many of them overlooked. In this thesis, reliability refers to the extent to which items/variables provide consistent and repeatable measurement, such as across items (i.e., internal consistency) and across different raters (i.e., inter-rater reliability). Validity refers to the extent to which the scores from a measure represent the variable/construct that they are intended to (i.e., construct validity). One overarching issue comes from how to conduct assessments, particularly in ECE settings. Miller-Bains et al. (2017) examined the validity of a multidimensional performance-based assessment (also known as observation-based, naturalistic, or work sampling assessment), which refers to an assessment in which teachers rate a child's skills on a scale of items in their classroom context, including student work, observations, and classroom tasks. This is a primary framework for assessment, as opposed to a direct assessment format (e.g., interview, pencil-and-paper tests) which requires more resources (i.e., training, cost, time), and may not be suitable for young children. The authors examined internal consistency of one of the widely used performance-based assessments, TS GOLD, and found that teachers were likely to rate individual students similarly across all learning constructs (e.g., Language, Math, Socioemotional). That is, teachers were less able to differentiate students' skills within the same classroom, potentially because teachers need to assess individual children while multiple children are engaging in the same activity. Consequently, it could be difficult to disentangle individual performance from group performance, suggesting a potential challenge of ECE assessment reliability and validity due to the clustering effect with peers during the assessment.

There are other risks of assessment that are unique to DLL children. For example, teachers may not be able to tell whether an error made by DLL children (e.g., math) reflects their learning progress (e.g., not fully understanding a concept yet) or need for language support (e.g., not understanding a question itself). Given that teachers conduct their assessment during an ongoing activity with multiple children, it is extremely difficult to identify the cause of error, even for experienced teachers. Another issue when teachers rate DLL children is scoring reliability. Joseph et al. (2020) reported that DLL children were more likely to be misclassified as not meeting the language domain compared to other English-speaking children. This could be due to several linguistic mismatches between teachers as raters and children, such as different phonological inventories. For example, DLL children are likely to differentiate phonetic details of their two languages to maximize their efficiency of language processing (e.g., Oh et al., 2011), resulting in their production being different from monolingual English-speaking children. However, ECE teachers, particularly for those who are monolingual English speakers and have limited exposure to diverse talkers, might detect the difference as an “error” although this is not the case⁵.

Overall, potential validity and reliability issues in teachers’ ratings for DLL children exist in current assessments. Thus, it is essential to develop a method to assess children who have diverse assets, such as immigrant and DLL children. One way to do this may be through family-based school readiness assessments. By using parents as children’s raters, we can address classroom clustering effect as parents can carefully observe their child without being impacted by distractions or resource issues that group-conducted assessments face (Miller-Bains et al., 2017). In addition, parents and children often share the same linguistic characteristics, which could also help eliminate the bias observed in Joseph et al. (2020). Furthermore, prior research

⁵ See the details of findings reviewed in Table 1.1.

suggests a potential benefit of parent-led assessments and have found a high correlation between teacher-rated and parent-rated assessment with homogeneous samples, i.e., monolingual Chinese mothers and teachers, (Xie and Li, 2019). This study suggests that parents as raters could be as reliable as teachers with professional training and experiences, but further examination is necessary with data including diverse demographic backgrounds to simulate the US context. In chapters 2 and 3, I will take up this issue. I will examine whether having parents as raters is feasible by simulating family-based assessment using a national survey and providing reliability and validity evidence. I will then examine whether parents' ratings of school readiness differs between immigrant and non-immigrant families.

Name of Measures	Abbreviation	Purpose of Study	Methods (A: Subjects, B: Analyses)	Findings	Reference
Kindergarten Student Entrance Profile	KESP	To examine the factor structures and reliability of KESP (revised)	A. 579 Latino students from low income households, central CA B. Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA)	1. Two-Factor CFA showed that the 12-item KESP is best conceptualized as measuring two distinct aspects of children's school readiness—Socio Emotional Readiness (SER) and Cognitive Readiness (COG). 2. Internal consistency yielded evidence supporting each factor's reliability; SE ($\alpha = .88$) and COG ($\alpha = .81$).	Quirk, Reibelz and Furlong (2014)
Kindergarten Student Entrance Profile	KESP	To examine measurement invariance of KESP for kindergarten and prekindergarten students	A. 572 kindergarten and 548 prekindergarten students in central California. B. Multiple-groups confirmatory factor analysis models with parameter constraints across groups	1. Measurement and structural parameter of the KESP were invariant across K and Pre-K students 2. Suggests KESP could be used for Pre-K to inform interventions prior to kindergarten entry	Erdyoun, Quirk and Furlong (2017)
Teaching Strategies GOLD	TS GOLD	To examine (i) TS GOLD for its internal consistency and its convergent and discriminant validity in a sample of kindergarten classrooms; (ii) correlations between widely used performance-based assessment (i.e., the Woodcock-Johnson III Tests of Achievement) and TS GOLD	A. 58 teachers and 1,066 kindergarten students (73% White, 16% Black, 10% Hispanic, 4% Asian; 40% free lunch), SE states, US B. Confirmatory factor analysis and estimated the interdomain correlations for TS GOLD; and examined convergent and divergent validity by comparing children's school readiness scores (TS GOLD by teachers) to scores as measured by independent data collectors using direct assessments	1. Showed TS GOLD's scale reliability in terms of factor structure and the convergent validity 2. Literacy, language and cognitive (except Math) had comparable or larger associations with direct assessments of different constructs 3. Teachers tended to rate individual students more similarly across all learning constructs 4. TS GOLD were less able to differentiate students' skills within a classroom (i.e., students' scores are more correlated within a same classroom by performance-based compared to direct assessment)	Miller-Baums et al. (2017)
Washington Kindergarten Inventory of Developing Skills	WaKIDS	To examine inter-rater reliability of WaKIDS, particularly with (i) agreement with the master code (developed by experts); (ii) frequency of inaccurate assessment and (iii) teacher's influences on the ratings	A. 54 teachers (43 full- and 11 part-day) from 26 school districts in WA B. Watched and rated four kindergarten student portfolios (Black, White, Latino, Mixed race; three of them are English-speaking including one DLL and one student with developmental delay); Conducted inter-rater reliability	1. Teacher ratings were more frequently in agreement with the master code for the typically developing, native English-speaking students (Black 86%, White 85%), but not for DLL (66%) nor developmental delay (74%) 2. Cognitive, Physical, and Mathematics domains tended to be more different from the master code than were the Social-emotional, Language, and Literacy; teacher's ratings for students with developmental delay tends to be inflated (i.e., overestimated); DL children was more misclassified as not meeting the language domain 3. Teachers' training and professional experiences positively affect their ability to accurately assess children	Joseph et al. (2020)
Chinese Preschool Readiness Scale	CPRS	To develop and assess CPRS	Study 1: A. 29 teachers and 288 children B. EFA Study 2: A. 46 teachers and 765 children; 110 parents B. ESEM and CFA	1. Five factors were determined: self-care abilities and emotional maturity, cognitive and communications skills, social competence, learning dispositions, and classroom rules 2. Concurrent validity was established: CPRS scores were correlated with parent-rated self-regulation (but not for all)	Xie and Li (2019)

Table 1.1 Summary of Findings (regarding evidence of reliability and validity)

1.2 FAMILY INVOLVEMENT IN ECE ASSESSMENT

While using parents as raters for existing school readiness frameworks may help alleviate some challenges of current assessments for DLL and immigrant children, a more transformative approach is to reconsider the current assessment frameworks altogether in order to capture the voices of families. The latter part of this thesis thus aims to address issues surrounding immigrant children's assessment by proposing a study to co-design a new school readiness measure with families. Families play significant roles in their children's learning. Traditional tests and measurements have historically examined different constructs (e.g., literacy, math skills) in classroom settings with minimal family involvement; yet, recent research in the field of learning sciences promotes the integration of family practices into children's learning. The following section briefly summarizes recent learning theories, namely, asset-based pedagogies, and discusses how we can integrate some of the concepts to design a new family-based school readiness.

One of the solutions to address issues of traditional readiness measures is to integrate an asset-based approach. For example, Gloria Ladson-Billings' "culturally relevant pedagogy (CRP)" and Django Paris's and H. Samy Alim's "culturally sustaining pedagogies (CSP)" are two influential curriculum and instruction theories which take into consideration power, equity, and justice. Culturally relevant pedagogy (CRP; Ladson-Billings, 1995) aims to 1) promote academic success, 2) promote willingness to nurture and support cultural competence, and 3) develop a socio-political or critical consciousness in the early version of CRP (Ladson-Billings, 1995). Later, Ladson-Billings revised the first component of CRP to read as "focus on learning" (Ladson-Billings, 2017, p. 142). Culturally sustaining pedagogy (CSP; Paris, 2012; Paris & Alim, 2017) shares a similar view to culturally relevant pedagogy but builds on CRP by emphasizing the importance of retaining the heritage languages and cultures; and culturally enriching strengths,

stated as “CSP seeks to perpetuate and foster—to sustain—linguistic, literate, and cultural pluralism as part of schooling for positive social transformation” (Paris & Alim, 2017, p. 1). These asset-based approaches provide a framework to value racial, linguistic, and cultural diversities that individual students carry, and there is positive empirical evidence showing the effect of cultural integration on learning outcomes, including with Latino children (e.g., Matthews and Lobe, 2019).

While we have seen curriculum and instruction incorporate the asset-based approach, assessments tend to withstand change. That is, little consideration has been made to reflect diverse immigrant and DLL children’s assets in both research and application of assessments. Indeed, some assessments have ‘bilingual’ versions (e.g., Woodcock–Muñoz (WM) Language Survey Revised, i.e., Spanish and English Forms), which expanded the language use but not the conceptual framework. Evans (2021) also discussed the difficulties of applying asset-based assessment to standardized testing:

“...it seems there is a significant challenge ahead for large-scale standardized tests as the level of individualization needed is in direct opposition to the way these tests are designed and the intended uses of these tests within state or district accountability systems.”

There are foreseen challenges in applying the asset-based approach to standardized testing or any large-scale measurement due to the lack of generalizability. Yet, this does not mean we need to endure the issues of testing which could further marginalize children and families. Indeed, Souto-Manning and Rabadi-Raol (2018) highlight the issue by using the theoretical framework of critical racial theory (CRT) such that “how traditional notions of quality in early childhood education are exclusionary, rooted in White monolingual and monocultural values and experiences, and apply deficit paradigms to frame the developmental trajectories of multiply

minoritized children” (p. 204). To disrupt deficit-based assessment used to compare a white middle-class “standard” to racialized immigrant children, it is crucial to extend the concepts of assessment by integrating racial, cultural, and linguistic diversities and strengths that each child brings into and beyond classroom.

1.3 RESEARCH QUESTIONS

The current thesis examines two possible approaches for incorporating families into ECE assessment. First, I examine whether parents are valid and reliable raters (see Chapter 2), and whether those ratings differ by immigrant and non-immigrant families (see Chapter 3). Second, I explore a more transformative approach by proposing the co-design with immigrant families of a new family-based school readiness measure (see Chapter 4).

Chapter 2. VALIDATION OF FAMILY-BASED SCHOOL READINESS MEASURE

In this chapter, I examine whether parents are valid and reliable raters by simulating parent-reported assessment using a national survey. I chose KSEP as a model of school readiness given that prior research showed a high reliability using it with DLL children, specifically Latinx children (Quirk et al., 2014), the number of items is smaller, and its rubrics (see Appendix) are relatively easy to use with minimum training compared to other measures³. Furthermore, as prior research suggests the importance of examining both cognitive and non-cognitive school readiness (Davis et al., 2016), this measure covers two factors, cognitive and socioemotional skills, which were identified by Exploratory and Confirmatory Factor Analyses (CFA/EFA) in Quirk et al. (2014) (See Figure 2.1). For this study, I tested a parent-reported school readiness measure following the same factor structure (socioemotional and cognitive skills) using items and data from

National Household Education Surveys (NHES) Program of 2019. Figure 2.2. shows the conceptual model with the items extracted from NHES. Hence, the current chapter addresses the following questions:

1. What is the reliability evidence for measuring school readiness using parents as raters?
2. In addition, what is the construct validity evidence for measuring school readiness using parents as raters?

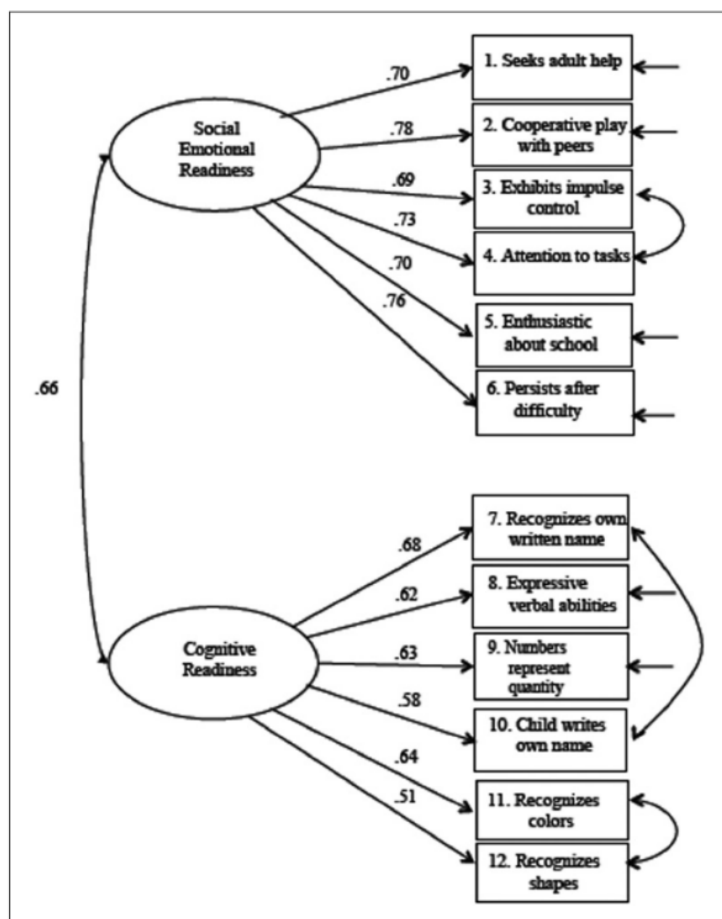


Figure 1. Standardized parameter estimates and factor correlations for final two-factor CFA model. Note. All estimates were significant at the $p < .001$ level. CFA = confirmatory factor analysis.

Figure 2.1. Path diagram for a two-factor confirmatory factor model using KSEP (adapted from Quirk et al., 2014).

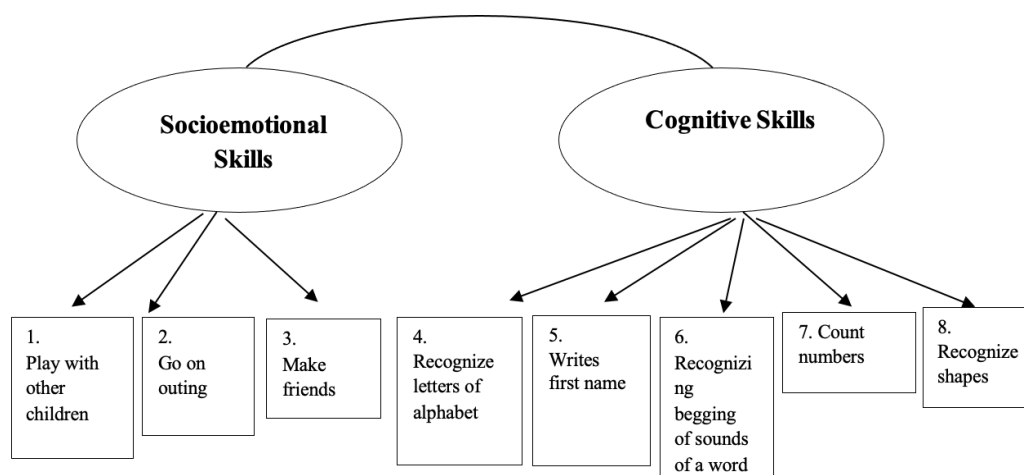


Figure 2.2. Revised path diagram for a two-factor confirmatory factor model using National Household Education Surveys Program of 2019.

2.1 METHODS

Dataset

Data comes from the public use National Household Education Survey (NHES) Program of 2019. NHES provides descriptive data on the educational activities of the US population and has been conducted every two to three years from 1991. In this study, the most recent data, from 2019, was utilized. The data was collected by completing the paper form (via mail), the web survey or by phone across 50 states and the District of Columbia. This parent-reported survey includes information about early care and education for a total of 7,092 children from birth through age 6. The ratio of child's sex is almost even (51.9% Male, 48.1% Female), and their race and ethnicity includes White, non-Hispanic (49.2%), Hispanic (25.6%), Black (12.7%), Asian or Pacific Islander (5.6%) and other race (7%). While most of school readiness data is available in aggregated versions (e.g., Washington Kindergarten Inventory of Developing Skills), this data provides more detailed information about subscales and questions. In this study, a subset of data ($N = 2,644$) was

used that included all parents' reports with children ages 4 and 5 in order to simulate a school readiness measure and be compatible with Quirk et al. (2014).⁶

Analyses

Using the above database, construct validity of parent-reported school readiness was assessed using exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA determined a factor structure based on the data and the explained amount of variance. CFA then confirmed the number of factors and relationships with associated items. Thus, EFA and CFA enabled me to determine whether the samples could fit the two-factor model as originally developed in Quirk et al. (2014). The total sample ($N = 2,644$) was randomly split in two. The first half was used for EFA and the other half was used for CFA. All the data analyses were performed in R (R Core Team, 2013).

Phase I: Reliability

Internal consistency was performed to examine reliability. This evaluates the correlations between items and factors, and Cronbach's alpha was measured. An alpha value of 0.70 and greater indicates acceptable internal consistency.

Phase II: EFA and CFA

As for EFA, parallel analysis with oblimin rotation was performed on the first subset of sample ($N = 1,322$). In addition, examination of the scree plot and eigenvalues were conducted to identify the factor structure. Items with loadings below 0.30 were excluded. As for CFA, the other half of the dataset was used for the analysis. The key indices to assess goodness-of-fit were Comparative Fit Index (CFI) where values of 0.90 or greater indicate acceptable fit (0.95+ are great); Root Mean Square Error of Approximation (RMSEA), where values of 0.06 or lower

⁶ Given that limited ECE data is available, many education researchers have been using a census data to address various issues in ECE.

indicate acceptable fit; and Standardized Root Mean Square Residual (SRMR), where values of 0.08 or lower indicate acceptable fit.

2.2. RESULTS

Phase I: Reliability

High internal consistency was found for the total scale (Cronbach's $\alpha = 0.89$), and this indicated a good internal consistency.

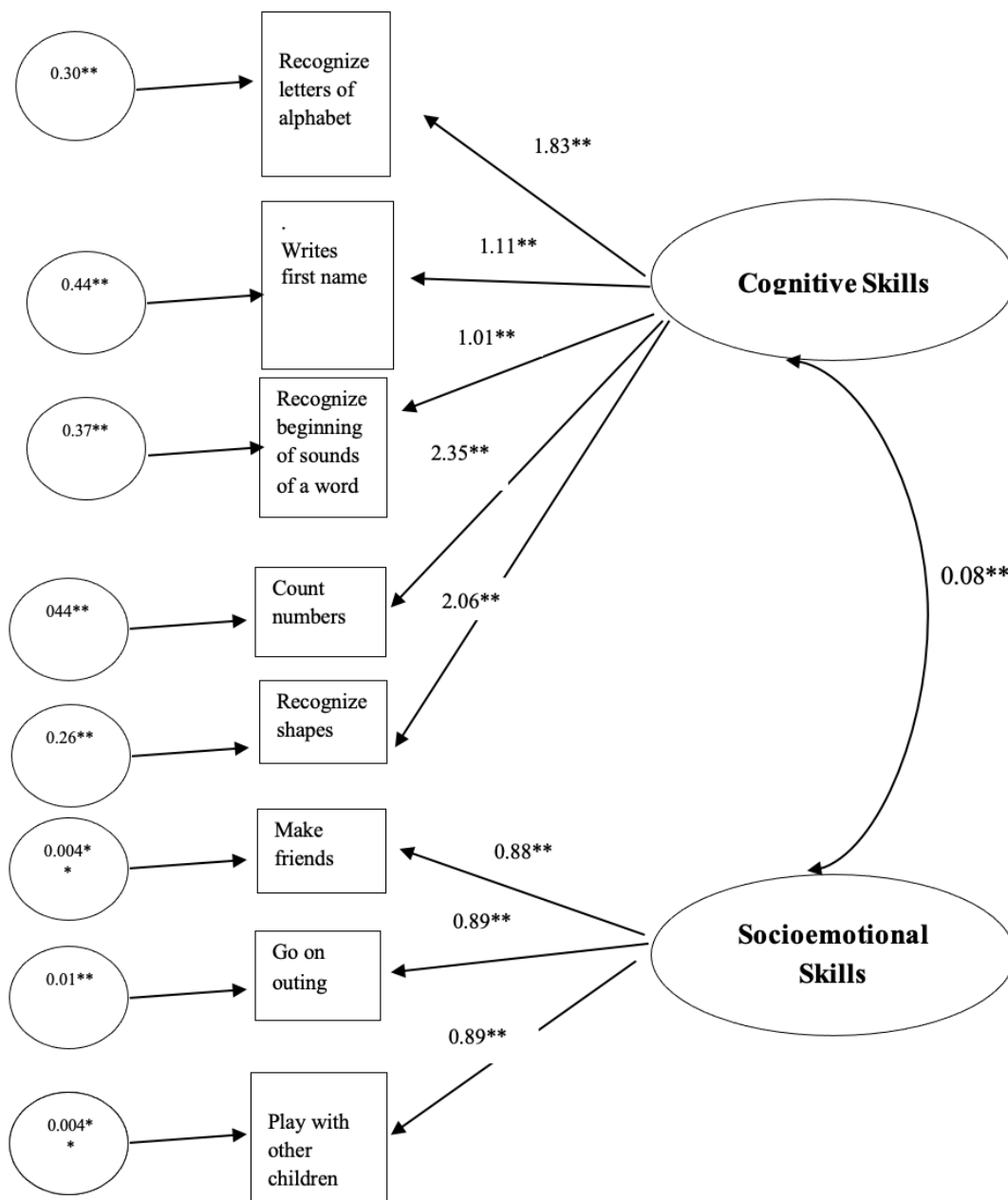
Phase II: EFA and CFA

As for EFA, parallel analysis, examination of the scree plot, and eigenvalues suggested that two common factors are the best fit. See Table 2.1. for EFA with oblimin rotation. Furthermore, CFA was performed using the Weighted Least Square Means (WLSM). First, unidimensional confirmatory factor analysis was performed, and the model fit was not acceptable ($\chi^2 = 35206.35$, $RMSEA < 0.05$, $SRMR = 0.287$, $CFI = 0.408$). Following this, as suggested by EFA, two-factor confirmatory factor analysis was performed, and the model fit was acceptable ($\chi^2 = 3478.27$, $RMSEA < 0.05$, $SRMR = 0.038$, $CFI = 0.942$). See Figure 2.3. for the graphical summary of two-factor confirmatory factor analysis (CFA) results.

Item	Description	Communalities	Factor1 Loading	Factor2 Loading
1	LETTER	0.92	0.96	
2	NAME	0.73	0.85	
3	SOUND	0.74	0.86	
4	COUNT	0.93	0.97	
5	SHAPE	0.94	0.97	
6	PLAY	0.99		0.99
7	OUT	0.99		0.99
8	FRIEND	0.99		0.99
			Factor1	Factor2
<u>Oblimin Rotated Variance Accounted For</u>			53%	37%

Note. $N=3,546$. Estimates from maximum likelihood with Oblimin rotation; statistically significant loadings shown in boldface. names. LETTER = recognize letters of alphabet; NAME = write own name; SOUND = Recognizing begging of sounds of a word; COUNT = count numbers; SHAPE = recognize shape; PLAY = enables to play with other children; OUT = enjoy outing; FRIEND = enables to make friends

Table 2.2. Exploratory Factor Analysis with Oblimin Rotation



Note. $N = 3546$. * $p < 0.05$, ** $p < 0.01$

Figure 2.3. Graphical summary of two-factor confirmatory factor analysis (CFA) results using subset of data from National Household Education Surveys Program of 2019.

2.3. DISCUSSION

This study tested to see whether a parent-reported assessment (i.e., parents as raters) could be valid and reliable. In order to address this investigation, a series of reliability and validity analyses were conducted using a national survey, NHES, 2019. Prior research has shown potential validity and reliability issues for teacher-reported school readiness measures, especially for immigrant children. This study addresses the issues by adapting a widely used two-factor school readiness construct to simulate a parent-reported school readiness measure.

The findings showed that parent-reported school readiness could be valid and reliable given that the subset of items from the national survey could maintain the two-factor model provided by Quirk et al (2014). Although this study was able to offer a great deal of insight into the measurement of school readiness for DLL children, it also has limitations. First, the original national survey was not designed to address school readiness but rather to ask a series of questions regarding early childhood program participation and parent and family involvement in education. For future studies, a careful design of surveys is essential to fully understand the construct of school readiness rated by immigrant (and non-immigrant) families. In the next chapter, I extend this work by examining whether school readiness can be meaningfully assessed by parents of both immigrant and non-immigrant children equivalently through looking at measurement invariance.

Chapter 3. VALIDITY AND RELIABILITY OF SCHOOL READINESS: CROSS-GROUP COMPARISON BETWEEN IMMIGRANT AND NON- IMMIGRANT FAMILIES

To examine potential usage of parent-reported school readiness measures for both immigrant and non-immigrant families, this study examined measurement invariance across immigrant and non-immigrant families. Measurement invariance refers to “whether an instrument is interpreted in the same way across different groups of individuals” (Holden et al., 2020). Assessing measurement invariance is an essential process in this thesis to ensure that interpretations of the latent construct (i.e., school readiness) with this national survey are valid across subgroups (i.e., immigrant and non-immigrant families). For example, parents from immigrant families might respond to the national survey differently due to not understanding items/questions or due to some other issues. Thus, examining measurement invariance of school readiness across the families allows us to determine the appropriateness of using the parent-reported items from the national survey as an alternative school readiness measure. Hence, the current chapter addresses the following questions:

1. To what degree is school readiness measured invariantly across children from different family backgrounds (i.e., immigrant and non-immigrant families)?
2. If variance tests suggest that latent mean comparisons are appropriate across groups, how does school readiness differ based on family immigration status?

3.1. METHODS

Dataset

The same dataset was used from Chapter 2, but the data was divided by immigration status. In total, 2,319 non-immigrant and 325 immigrant family data were analyzed⁷. The data analyses were performed in R (R Core Team 2013). To address school readiness, five items from the survey were used. These items are used to measure cognitive aspects of school readiness as defined in Quirk et al. (2014). Items related to the socioemotional factor were dropped due to a planned missingness. Thus, this investigation treats school readiness as a one-factor model by using items related to cognitive domain.

3.2. ANALYSES

Phase I: MIMIC Model

Prior to the measurement invariance analysis, I looked at whether there were differences at the latent structure using the Multiple Indicators, Multiple Causes (MIMIC) Model. This is an extension of CFA with covariates and used to contextualize the latent variable of interests (e.g., school readiness) using the demographic variables (e.g., family background). Using the dummy-coded predictor variable, the effect of group membership on the latent structure was examined.

Phase II: Measurement invariance analysis using Lavaan

In the second step, I tested for configural invariance by allowing all model parameters to vary freely across groups. In the third step, weak/metric invariance was tested by constraining the factor loading equally across groups. As for the fourth step, strong/scalar invariance, which builds upon metric invariance, was tested by constraining item intercepts equal across groups as well as the factor loading. At last, strict invariance, built upon the strong/scalar invariance, was tested by further constraining item residual variances to be equal across groups. All the steps of the

⁷ The current study did not use any measures to control sampling issues such as simulation nor sampling weight. See the caveats in Discussion.

measurement invariance analysis were conducted using `measurementInvariance()` function from a package, 'semTools.'

Model fit was evaluated using a variety of fit indices, as recommended by Brown (2006). Measurement invariance is tested by evaluating how well the specified model (e.g., the model set up by the researcher) fits the observed data. Current practice emphasizes the importance of using multiple fit statistics to assess model fit (Kline, 2015). As research suggests the issues of using the chi-square statistics in measurement invariance analyses (e.g., Maede et al., 2008) such that SEM requires a large sample size, the current study relied on other goodness-of-fit (i.e., approximate) indices: Root Mean Square Error of Approximation (RMSEA) values less than .05 (Browne & Cudeck, 1992), Standardized Root Mean-square Residual (SRMR) values less than .01, and Comparative Fit Index (CFI) values greater than .90 (Marsh et al., 2004). Chen (2007) also suggested a criterion of a -.01 change in CFI, paired with changes in RMSEA of .015 and SRMR of .030 (for metric invariance) or .015 (for scalar or residual invariance).

3.3. RESULTS

Phase I: MIMIC Model

The MIMIC Model showed that there was no significant effect of group membership on the latent structure ($p > .05$) while the model fit indices showed that only the CFI meet the guideline ($\chi^2(9) = 83.93$, RMSEA = 0.06, SRMR = 0.028, CFI = 0.980). See Figure 3.1. for the MIMIC model diagram from R `semPlot` package. This suggests that while these items of the school readiness measure (observed variables) predict school readiness (latent variable), the group difference between immigrant and non-immigrant families is not observed.

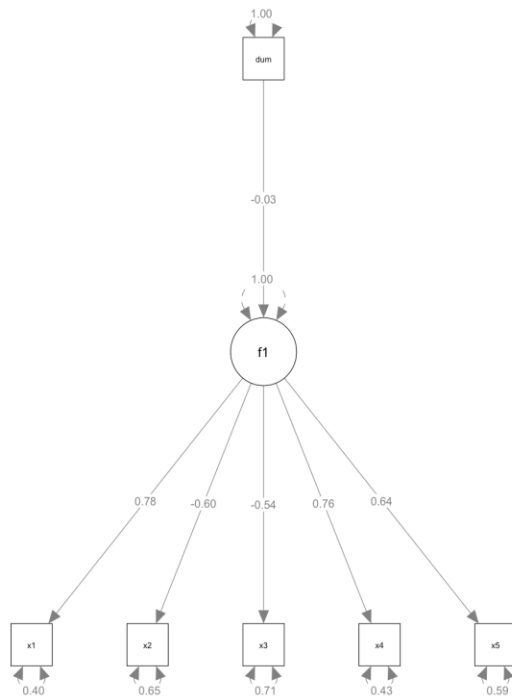


Figure 3.1. *The MIMIC model diagram from R semPlot package*

Phase II: Measurement invariance analysis using Lavaan

Further, a series of measurement invariance analyses were conducted. The configural invariance model showed acceptable levels of model fit (Robust: $\chi^2(10) = 22.36$, $p = 0.013$, CFI = 0.996, RMSEA = 0.031, SRMR = 0.013) through the two indicators of approximate fit indices (CFI and RMSEA). While the invariance was failed by the robust chi-square test, there is enough invariance that we could practically proceed with further steps to increase the constraints. The analysis of measurement invariance showed that the measurement invariance failed at strong/scalar invariance ($p < .001$), which are built upon metric invariance and was tested by constraining item intercepts equal across groups as well as the factor loading. See Table 3.1. for the summary of model fit statistics. This finding suggests that two groups of respondents (immigrant and non-

immigrant families) interpret a given measure (school readiness) in a conceptually similar manner, yet immigrant families are systematically rating the items lower.

Model	RMSEA	Δ RMSEA	CFI	Δ CFI	χ^2 (df)	P value for χ^2 difference
1. Configural Invariance	0.039		0.998		16.6 (10)	
2. Weak/Metric Invariance	0.03	0.008	0.998	0	31.4 (15)	0.18
3. Strong/Scalar Invariance	0.067	0.036	0.985	0.013	108.3 (22)	***
4. Strict Invariance	0.054	0.012	0.989	0.005	121.8 (23)	0.25

Note: ***p <.001

Table 3.3. Model Fit Statistics for Measurement Invariance Analyses from Immigrant and Non-immigrant Families

3.4. DISCUSSION

This chapter examined measurement invariance across immigrant and non-immigrant families using a simulated parent-reported school readiness measure. Overall, several important findings were shown. In the measurement invariance analyses, partial invariance was confirmed, such that a non-invariant intercept was observed (See Figure 3.2. for the graphic demonstration). This suggests that one group tends to give systematically higher (or lower) item responses as a default. In this data, items collected from non-immigrant family data showed a higher intercept compared to immigrant family data, potentially because immigrant children are more likely to encounter issues in school readiness, such as DLL children's language skills (item 1 and item 2) (Magnuson et al., 2006). Given that many standardized assessments, including school readiness measures, were originally developed for English monolingual (non-immigrant) populations, we need to further consider how we can accommodate early childhood learning and assessment considering the fast-growing population of DLL and immigrant children.

Another important implication of this study is that parent-reported assessment can be useful through approximate (but not perfect) fit indices. As parents' involvement appears to benefit

children of immigrants such as an increase in the level of English proficiency and other cognitive scores (Lahela, 2008), there may be advantages of using parent-reported assessment not only for scoring (e.g., Joseph et al., 2020) issues, but also to promote parent involvement in immigrant children's learning and development.

Lastly, it is also important to address several issues in this set of analyses. First, there is a large sample difference between two groups. For the chi-square test, for instance, it is unclear whether there was enough power to detect model misfit due to the small sample size, especially the data from immigrant families. In the future, Monte Carlo Simulation and/or sample weight needs to be considered in order to address the two crucial issues, i.e., the small sample size as well as the imbalance of sample size across groups. Regardless, the findings from these analyses contribute important insights regarding the reliability and validity of a parent-reported school readiness measure across immigrant and non-immigrant families.

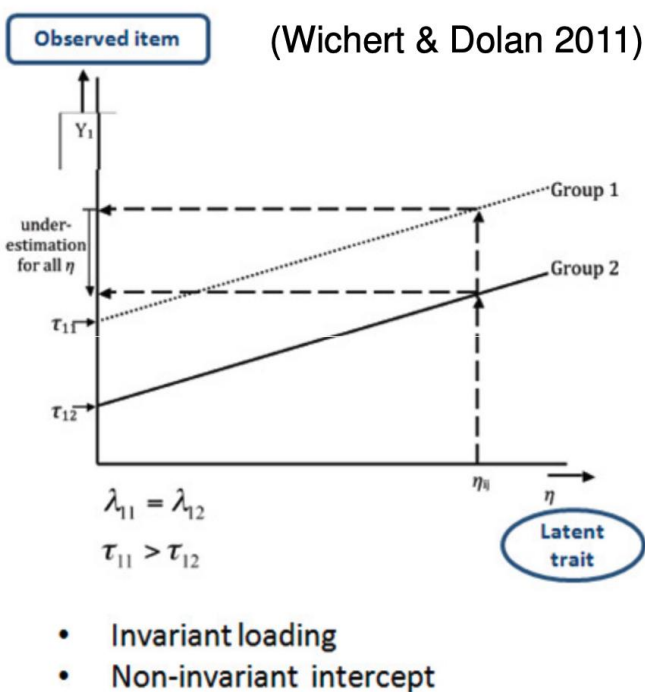


Figure 3.2. *Strong/scalar non-invariance (adapted from Wichert & Dolan, 2011)*

Chapter 4. PROPOSAL FOR CO-DESIGNING AND IMPLEMENTING A FAMILY-BASED SCHOOL READINESS MEASURE

4.1. INTRODUCTION

To dismantle the pervasive issues in educational assessment, I propose designing a new family-based school readiness measure with immigrant families. Given that prior studies have pointed out potential risks for immigrant and DLL children being assessed in less valid and reliable manners, it is crucial to explore potential solutions. This thesis argues for greater family involvement in assessment as one potential direction. Chapters 3 and 4 simulated the possibility of parents as reporters, and I have confirmed that a parent-reported school readiness measure can have acceptable validity and reliability.

The next question to look at is how we might change the assessment content and process altogether. As chapter 4 disclosed, the traditional school readiness assessments may continue to lead to a deficit perspective towards DLL children, even with parents as reporters, particularly if the same items are used. While this approach could help introduce additional language support, we need to further think of innovative procedures to help immigrant children thrive with their diverse assets, which traditional education systems (e.g., curriculum, instruction, and assessments) have yet to uncover. Co-designing and implementing a family-based school readiness measurement not only could lead to more accurate assessment for immigrant children and families, but could possibly address and change pervasive systematic issues surrounding DLL children.

Accordingly, the proposed research investigates the following questions:

- 1) How can I help create a healthy and productive environment for co-design where immigrant families' voices are heard in school readiness assessment and beyond?

- 2) What are some effective practices for co-designing with immigrant families in ways that leverage immigrant children's assets?

4.2. DESIGN DESCRIPTION

This research is categorized as a social design experiment. Gutiérrez and Jurow (2016) define the goals of social design experiments as ones that “include the traditional aim of design experiments to create theoretically grounded and practical educational interventions, the social agenda of ameliorating and redressing historical injustices, and the development of theories focused on the organization of equitable learning opportunities” (p. 565). In the area of early childhood education research, immigrant and DLL children’s learning and development has been framed through a deficit lens by constant comparisons to a White middle-class ‘standard’. Despite recent progress towards greater attention to equity, particularly after Garcia-Coll and her colleagues advocated a resilience perspective which emphasized the “diversity and strengths” of marginalized groups (1996, p. 1891), little practical considerations have been made regarding how ECE assessment is developed and used to value immigrant children’s assets. Thus, the present work emerged following a series of quantitative analyses (see the previous chapters) investigating the issues of school readiness measures surrounding immigrant children.

As a part of the design process, I have created a conjuncture map that “conjectures through a learning environment design, distinguishing conjectures about how the design should function from theoretical conjectures that explain how that function produces intended outcome” (Sandoval, 2014, p. 18). In other words, this conjuncture map specifically addresses the examination of this co-design process. See Figure 4.1. for the details. First, there are high level conjectures about how to support the intended learning, i.e., 1) designing a culturally

relevant/sustaining assessment for immigrant children and families that would value their cultural assets, and 2) positioning assets in early childhood assessment to disrupt deficit mindsets rooted in traditional notions of learning. The conjectures become concrete through a specific design ('embodiments') by preparing tools and materials needed 1) to organize informal gatherings (i.e., an outdoor picnic) with immigrant families to prepare a stage (scaffolding) for co-design, and 2) to design an assessment with immigrant families (e.g., recording device). Furthermore, the embodiments generate mediating processes such as developing trusting relationships between immigrant families and researchers; sharing children's assets with immigrant families; analyzing those children's assets; and co-designing a strengths-based, culturally responsive assessment by immigrant families and a researcher. The intended outcomes include 1) building immigrant families' trust that their children will be cared for, 2) engaging in immigrant families learning and leveraging their untapped resources, 3) developing an early childhood assessment which values diverse cultural assets, and 4) developing critical thinking/perspective towards "traditional" assessment by immigrant parents.

DESIGN AND THEORETICAL CONJECTURES

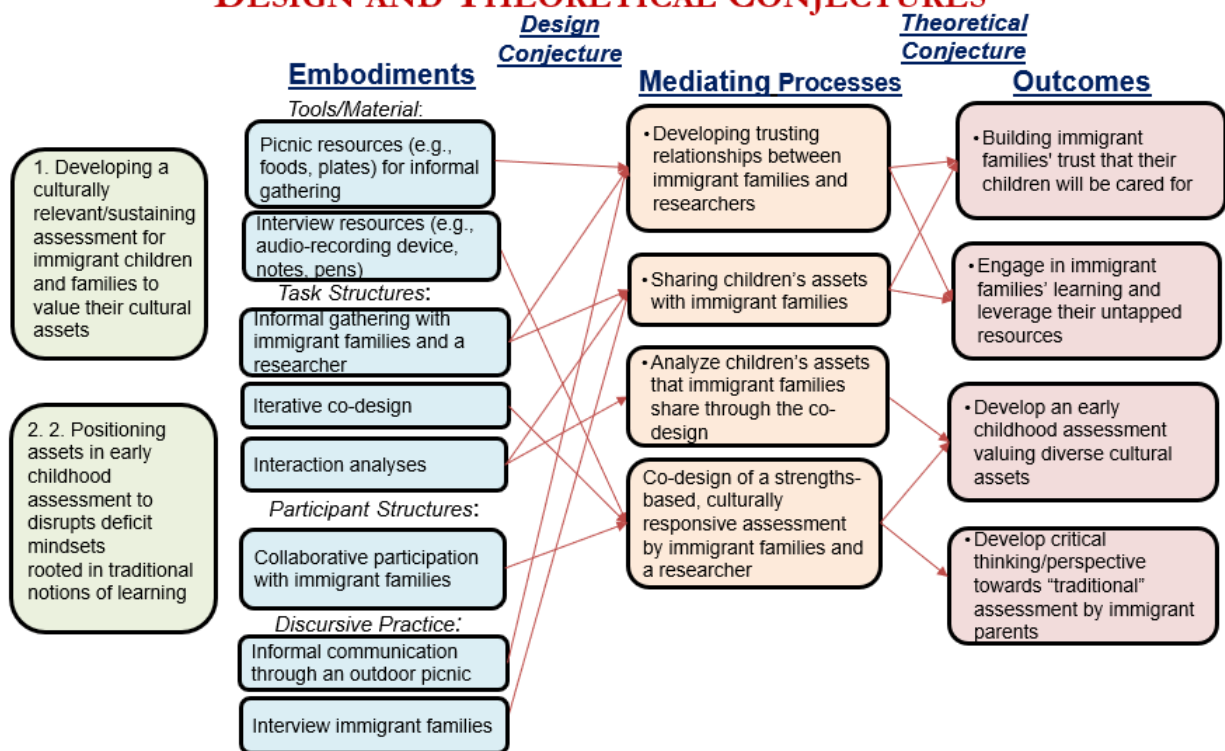


Figure 4.1. Conjuncture Map

4.3. METHODS

RQ1: How can a researcher help create a healthy and productive environment for co-design where immigrant families' voices are heard?

In this proposed co-design, I will organize a series of informal gatherings with immigrant families to scaffold the following co-design work. This stage is essential, as without building a trusting relationship, a co-design process cannot be successfully completed. That is, there could be a potential power barrier between immigrant families and researchers, which could hinder meaningful and authentic co-design.

As a prototype, two immigrant families who live in greater Seattle areas were invited to a casual picnic to discuss their children's strengths. While we enjoyed pizza and other foods at Anderson Park in Redmond, WA, I conducted a semi-structured interview, which attempted to capture strengths through items listed in the publicly available Washington State Early Learning and Development Guidelines (2012). See Figure 4.2. for the pictures of the prototype session. Based on the prototype session, there are a few things I observed. First, parents may be skeptical to participate in any research, especially answering questions regarding their children's learning and development, potentially because they feel judged. Thus, it is very important to develop trusting relationships between immigrant families and researchers prior to the co-design process. Even after developing a relationship, the co-design process needs to be carefully planned so that each question is navigated in a comfortable and respectful way to ensure the potential power balance between the families and researcher. Another powerful factor I observed was language use. The parents were all bilingual (English and Japanese), and I told them to answer in any comfortable language, but since I presented the inquiries in English, this led them to think more on the "English" side of their children's development and learning, such as behaviors conducted when their children were speaking in English. While this study was aimed to advocate for an asset-based, family-based assessment, I realized that even one component, the language use in this case, could affect the co-design process significantly regardless of the carefully designed theoretical and conceptual frameworks. Together, future co-design sessions need to take this into account to fully engage in learning about immigrant families and their children's assets including language use.

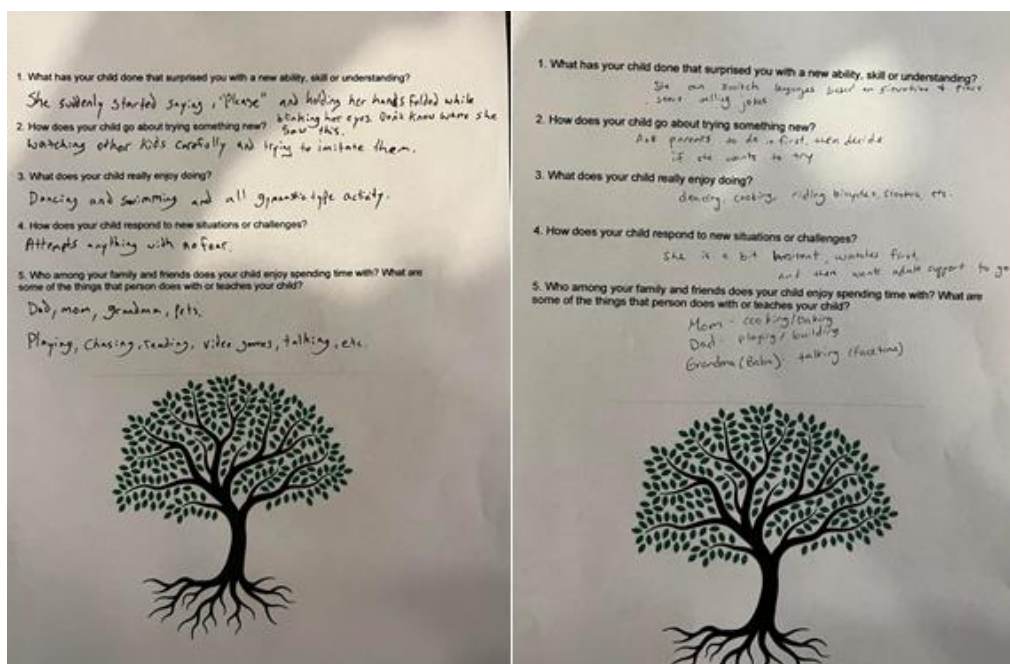


Figure 4.2. Pictures of Prototype

RQ2: What are some effective practices for co-designing with immigrant families in ways that leverage immigrant children's assets?

I propose a series of interviews with immigrant families to be conducted as a part of designing an asset-based school readiness measure. The goal of these interviews will be to engage in immigrant children's learning and leverage their untapped resources. For recruiting immigrant families, I will meet different organizational leaders and post recruitment flyers to different SNS platforms. The interview sessions will be recorded (audio-only), and the data will be analyzed to determine the items used for the assessment.

Following a series of interviews, final items will be determined for the assessment by the researcher. The newly developed assessment will be shared with immigrant families ($n = 30$) and further discussions will be made for the revisions. In addition, different language versions will be prepared so that immigrant parents can express their ideas with their preferred language, rather than use 'English-mode'.

Data Collection Plan

As for the research design timeline, I will first meet with different organization leaders (e.g., Latino Community Association) as well as immigrant family friends to recruit potential participants. After agreeing to participate in the research, informed consent as well as other background questionnaires will be obtained. Following the participant recruitment, two stages of data collection will be conducted. First, a series of informal gatherings will be planned to scaffold the following co-design process by building trusting relationships with immigrant families. These will be informal events, and immigrant families will be invited to voluntarily participate during the first week. The following week, co-designing assessment items will be determined with the previously invited immigrant families. The first day will be to gather ideas to brainstorm what kinds of assets children hold, and why it is important to reflect these assets in ECE assessment. Then, the researcher will select items that fit the above-mentioned asset-based approach. Following data analyses including transcribing, coding, and reflecting immigrant families' voices, I will select assessment items with immigrant families and receive feedback from them.

4.4. ANTICIPATED IMPLICATIONS OF THE PROPOSED STUDY

This research proposal was developed to co-design a new school readiness measure with immigrant families to dismantle pervasive issues surrounding immigrant and DLL children's learning. As DLL and immigrant children have been invisible in ECE contexts (Park & Pompa, 2021), reflecting their voices is essential to tackle systematic changes. Findings of this study are thus intended to contribute by 1) spotlighting DLL children and their families whose voices need to be heard, 2) informing teachers and schools about how they can help DLL children with diverse assets thrive through revised curriculum and instruction, and 3) enabling systems (governments

and policies) to identify DLL children and their learning situations to implement more inclusive policies and targeted funding.

Chapter 5. DISCUSSION

Despite the rapid increase of immigrant and DLL children in the US, their learning and growth opportunities have been marginalized in ECE settings. This is because there is no system or comprehensive research to identify immigrant and DLL children in the various ECE settings. Hence, their needs and voices, including in revisions of curriculum, instructions, and assessment, have been silenced (Park & Pompa, 2021; Souto-Manning & Rabadi-Raol, 2018). In this thesis, I focused on assessment, which has significant implications for ECE including curriculum, instruction and policies. Indeed, findings of potential rating biases towards DLL children (e.g., Joseph et al., 2020) call into question widely used, teacher-based school readiness measures. Such measures currently lack validity and reliability, especially for DLL children, potentially due to linguistic and cultural mismatches between teachers and DLL children. Thus, I examined whether parent-reported school readiness could be reliable and valid.

Findings of the two simulation studies showed that parents-as-raters could be reliable and valid. Using a national survey of early childhood education reported by parents, the first study showed reliability and validity evidence that parents' responses could maintain a two-factor model, originally proposed by KESP with teacher response.

Another significant finding is that immigrant parents responded to the school readiness items differently compared to non-immigrant parents. The second study with the group-comparison showed that while both immigrant and non-immigrant parents addressed school readiness in a similar manner, immigrant parents gave systematically lower item responses. This

could indicate that DLL children should receive additional support (e.g., language support) to acquire some school readiness skills. It is less likely that this lower score was due to misunderstanding of questions due to parent's limited English proficiency given that the national survey was conducted using multiple languages; however, there is a possibility that this school readiness measure was not suitable to reflect their children's skills as it was obtained through the items that traditional concepts of school readiness hold. That is, there is a possibility that immigrant parents assessed their children only from an 'English' side of their behaviors. In other words, some children might be able to count numbers in Spanish but not in English, which could lead the parents to evaluate it as not meeting the learning goal. The standardized assessments have limitations to accommodate diverse learning styles, and further investigation is necessary to reflect diverse DLL children's assets in addition to the family involvement in assessments.

Thus, the third study extended the studies by exploring how I, the researcher, and immigrant families could co-design an asset-based school readiness measure to reflect their authentic learning performance and progress as well as to reveal needs of learning support. Interestingly, my brief pilot examination demonstrated the importance of how the co-design needs to be navigated as my pilot immigrant families also used a lens of English as a standard. Regardless, this proposal provides anticipated implications such that this approach could reflect immigrants and DLL children's voices and leverage their assets. Furthermore, findings could contribute to changes in schools and systems such that instruction, curriculum and policies could be revised accordingly.

There are several limitations to the investigations in this thesis. First, we need to address the construct of school readiness. As summarized in Chapter 1, factor structures of school readiness are varied from having two subsets of skills, i.e., cognitive and socioemotional skills

(e.g., KSEP), up to nine (e.g., Illinois Kindergarten Individual Development Survey). Future research should examine the face/content validity with ECE researchers, educators, and families to determine what children need to learn to prepare for their longer-term learning. Second, while the first SEM study suggests that parent-reported school readiness measures can be valid and reliable, it does not mean teachers and parents rate in an equal manner. To address the direct comparisons, further examination is necessary using more controlled experimental settings, which would allow us to determine similarities and differences in their scoring. Third, this thesis targets DLL children broadly and does not specify a group such as Latinx children. Future studies should address whether and how family-based assessment work “for whom” by focusing on a specific ethnic group or choosing other demographic factors such as intersectionality of gender and race to capture the needs of DLL children.

There are also potential issues regarding consequential validity. Indeed, research has reported different degrees of perception towards school readiness based on parental socioeconomic and education status. For instance, parents without college degrees are likely to perceive school readiness measures as being more important compared to those with college degrees (West et al., 1995). In addition, low-income parents (e.g., whose children are in Head Start) tend to expect more from their children (Piotrkowski, 2004). It is thus important to consider how developing and implementing family-based school readiness measures affects immigrant families with diverse backgrounds.

In spite of the above limitations, I believe that implementing family-based school readiness measures would provide opportunities 1) for immigrant children to answer items in authentic and comfortable manners (e.g., language, environment), 2) for parents to assess their children’s learning, and 3) for teachers/schools to engage more with immigrant families. By meeting these targets, 1) immigrant parents and teachers could share needs of additional support, 2) teachers and

schools could navigate diverse assets of DLL children through accommodating the needs and revising curriculum and instructions, and 3) local/state/federal governments can identify detailed characteristics of DLL children in ECE and implement additional policies and/or fundings (See Figure 5.1. for the details described through a theory of change). In conclusion, this thesis has helped identify potential benefits of developing and implementing family-based school readiness measures with important implications for DLL children and ECE more broadly.

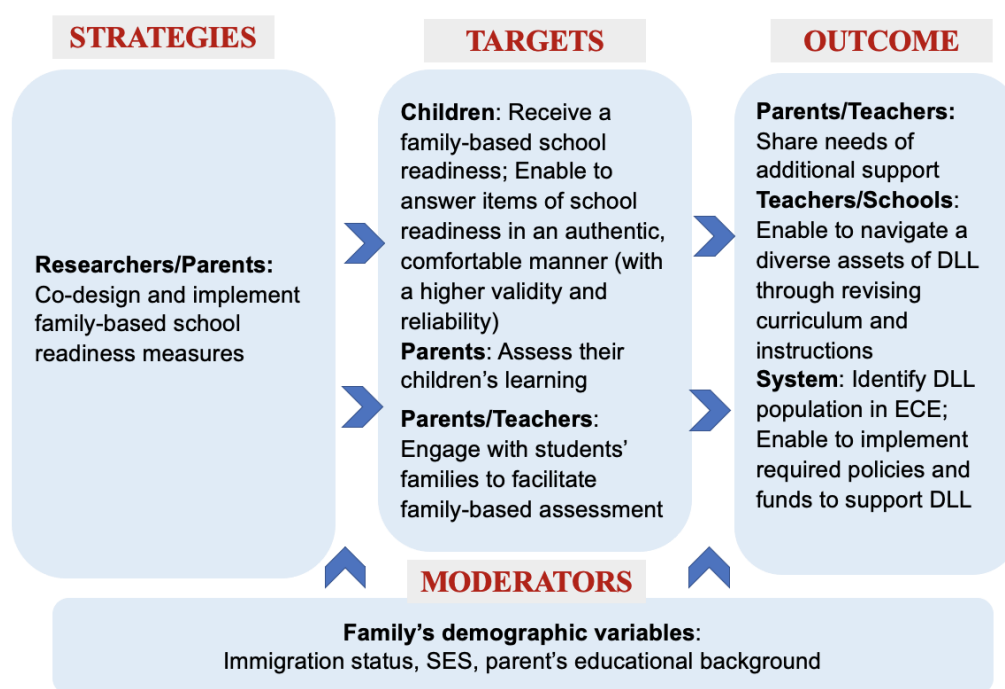


Figure 5.1. Theory of Change.

REFERENCES

- August, D., Shanahan, T., & Escamilla, K. (2009). English language learners: Developing literacy in second-language learners—Report of the National Literacy Panel on Language-Minority Children and Youth. *Journal of literacy research, 41*(4), 432-452.
- Bipartisan Policy Center. (2020). Expanding access to bilingual learning for dual language and English learners: What we know, what we don't know, and what we should do about it.
- Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological methods & research, 21*(2), 230-258.
- Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers college record, 112*(3), 579-620.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural equation modeling: a multidisciplinary journal, 14*(3), 464-504.
- Coll, C. G., Crnic, K., Lamberty, G., Wasik, B. H., Jenkins, R., Garcia, H. V., & McAdoo, H. P. (1996). An integrative model for the study of developmental competencies in minority children. *Child development, 67*(5), 1891-1914.
- Davies, S., Janus, M., Duku, E., & Gaskin, A. (2016). Using the Early Development Instrument to examine cognitive and non-cognitive school readiness and elementary student achievement. *Early Childhood Research Quarterly, 35*, 63-75.
- Espinosa, L. M., LaForett, D. R., Burchinal, M., Winsler, A., Tien, H.-C., Peisner-Feinberg, E. S., & Castro, D. C. (2017). Child Care Experiences Among Dual Language Learners in the United

States: Analyses of the Early Childhood Longitudinal Study–Birth Cohort. *AERA*

Open. <https://doi.org/10.1177/2332858417699380>

Edyburn, K. L., Quirk, M., & Furlong, M. (2017). Measurement invariance of a school readiness screener for use in preschool and kindergarten. *Early Education and Development, 28*(7), 810–821. <https://doi.org/10.1080/10409289.2017.1282802>

Esterline, C. E., & Batalova, J. B. (2022, March 23). *Frequently requested statistics on immigrants and immigration in the United States*. migrationpolicy.org. Retrieved May 9, 2022, from https://www.migrationpolicy.org/article/frequently-requested-statistics-immigrants-and-immigration-united-states?gclid=CjwKCAjwjtOTBhAvEiwASG4bCOQUzIKW4rAW5PFmlkCat4QVohEzdTByYumqIqCuFCJdl7gukx7ikhoC9qoQAvD_BwE

Evans (2021, November 3) Culturally Sensitive, Relevant, Responsive, and Sustaining Assessment: Are There Limits to Making Large-Scale Standardized Testing Culturally Responsive? <https://www.nciea.org/blog/culturally-responsive/culturally-sensitive-relevant-responsive-and-sustaining-assessment>

Gormley, W. T., & Gayer, T. (2005). Promoting school readiness in Oklahoma an evaluation of Tulsa's pre-k program. *Journal of Human resources, 40*(3), 533-558.

Gormley, W. T., Phillips, D., & Gayer, T. (2008). Preschool programs can boost school readiness. *Science, 320*(5884), 1723-1724.

Gormley, W. T., Phillips, D., & Anderson, S. (2018). The effects of Tulsa's Pre-K program on middle school student performance. *Journal of Policy Analysis and Management, 37*(1), 63-87.

Gutiérrez, K. D., & Jurow, A. S. (2016). Social design experiments: Toward equity by design. *Journal of the Learning Sciences, 25*(4), 565-598.

- Halle, T., Forry, N., Hair, E., Perper, K., Wandner, L., Wessel, J., & Vick, J. (2009). Disparities in early learning and development: lessons from the Early Childhood Longitudinal Study–Birth Cohort (ECLS-B). *Washington, DC: Child Trends*, 1-7.
- Head Start Programs. The Administration for Children and Families (2020). Retrieved May 9, 2022, from <https://www.acf.hhs.gov/ohs/about/head-start>
- Holden, G. W., Gower, T., & Chmielewski, M. (2020). Chapter 9 - Methodological considerations in ACEs research. In G. J. G. Asmundson & T. O. Afifi (Eds.), *Adverse Childhood Experiences* (pp. 161–182). Academic Press. <https://doi.org/10.1016/B978-0-12-816065-7.00009-4>
- Joseph, G., Soderberg, J. S., Stull, S., Cummings, K., McCutchen, D., & Han, R. J. (2020). Inter-rater reliability of washington state’s kindergarten entry assessment. *Early Education and Development*, 31(5), 764–777. <https://doi.org/10.1080/10409289.2019.1674589>
- Kline, R. B. (2015). Principles and practice of structural equation modeling. Guilford publications.
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American educational research journal*, 32(3), 465-491.
- Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in US schools. *Educational researcher*, 35(7), 3-12.
- Lipsey, M. W., Farran, D. C., & Hofer, K. G. (2015). A Randomized Control Trial of a Statewide Voluntary Prekindergarten Program on Children's Skills and Behaviors through Third Grade. Research Report. *Peabody Research Institute*.
- Magnuson, K., Lahaie, C., & Waldfogel, J. (2006). Preschool and school readiness of children of immigrants. *Social science quarterly*, 87(5), 1241-1262.

- Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First-and higher order factor models and their invariance across groups. *Psychological bulletin*, 97(3), 562.
- Matthews, J. S., & López, F. (2019). Speaking their language: The role of cultural content integration and heritage language for academic achievement among Latino children. *Contemporary Educational Psychology*, 57, 72-86.
- Matthews, H., & Ewen, D. (2006). Reaching all children? Understanding early care and education participation among immigrant families. *Center for Law and Social Policy, Inc.(CLASP)*.
- McCoy, D. C., Yoshikawa, H., Ziol-Guest, K. M., Duncan, G. J., Schindler, H. S., Magnuson, K., ... & Shonkoff, J. P. (2017). Impacts of early childhood education on medium-and long-term educational outcomes. *Educational Researcher*, 46(8), 474-487.
- Meade, A. W., & Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, 7(4), 361-388.
- Miller-Bains, K. L., Russo, J. M., Williford, A. P., DeCoster, J., & Cottone, E. A. (2017). Examining the validity of a multidimensional performance-based assessment at kindergarten entry. *AERA Open*, 3(2), 2332858417706969. <https://doi.org/10.1177/2332858417706969>
- National Academies of Sciences, Engineering, and Medicine. (2017). *Promoting the educational success of children and youth learning English: Promising futures*. National Academies Press.
- NHES. (n.d.). National Center for Education Statistics. Retrieved June 3, 2021, from <https://nces.ed.gov/nhes/dataproducts.asp>
- Oh, G. E., Guion-Anderson, S., Aoyama, K., Flege, J. E., Akahane-Yamada, R., & Yamada, T. (2011). A one-year longitudinal study of English and Japanese vowel production by Japanese

adults and children in an English-speaking setting. *Journal of Phonetics*, 39(2), 156–167.

<https://doi.org/10.1016/j.wocn.2011.01.002>

Park, M., & Pompa, D. (2021). Ending the Invisibility of Dual Language Learners in Early Childhood Systems: A Framework for DLL Identification. Issue Brief. *Migration Policy Institute*.

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational researcher*, 41(3), 93-97.

Paris, D., & Alim, H. S. (Eds.). (2017). *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world*. Teachers College Press.

Piotrkowski, C. S. (2008). MSW Students' Stereotypes of Mothers Receiving Means-Tested Public Assistance. *Arete*, 32(1).

Quirk, M., Furlong, M., Lilles, E., Felix, E., & Chin, J. (2011). Preliminary development of a Kindergarten School Readiness Assessment for Latino Students. *Journal of Applied School Psychology*, 27(1), 77–102. <https://doi.org/10.1080/15377903.2010.540518>

Quirk, M., Mayworm, A., Edyburn, K., & Furlong, M. J. (2016). Dimensionality and measurement invariance of a school readiness screener by ethnicity and home language. *Psychology in the Schools*, 53(7), 772–784. <https://doi.org/10.1002/pits.21935>

Quirk, M., Nylund-Gibson, K., & Furlong, M. (2013). Exploring patterns of Latino/a children's school readiness at kindergarten entry and their relations with Grade 2 achievement. *Early Childhood Research Quarterly*, 28(2), 437 –449. <https://doi.org/10.1016/j.ecresq.2012.11.002>

Quirk, M., Rebelez, J., & Furlong, M. (2014). Exploring the dimensionality of a brief school readiness screener for use with Latino/a children. *Journal of Psychoeducational Assessment*, 32(3), 259–264. <https://doi.org/10.1177/0734282913505994>

- Rojas, N. M., Yoshikawa, H., & Morris, P. (2021). Preschool Children's Engagement and School Readiness Skills: Exploring Differences between Spanish-speaking Dual Language Learners and Monolingual English-speaking Preschoolers. *Early Education and Development*, 1-25.
- Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the learning sciences*, 23(1), 18-36.
- Shonkoff, J., & Phillips, D. (2000). Committee on Integrating the Science of Early Childhood Development Board on Children, Youth, and Families. From Neurons to Neighborhoods: The Science of Early Childhood Development.
- Snow, K. L. (2006). Measuring school readiness: Conceptual and practical considerations. *Early education and development*, 17(1), 7-41.
- Souto-Manning, M., & Rabadi-Raol, A. (2018). (Re) Centering quality in early childhood education: Toward intersectional justice for minoritized children. *Review of Research in Education*, 42(1), 203-225.
- Washington State Department of Learning (2012). Washington State Early Learning and Development Guidelines Birth through 3rd Grade. Retrieved from https://www.dcyf.wa.gov/sites/default/files/pubs/EL_0015.pdf
- West, J., Germino-Hausken, E., & Collins, M. (1995). Readiness for kindergarten: Parent and teacher beliefs (Statistics in brief). Washington, DC: National Center for Education Statistics.
- Wichert, J. M., & Dolan, C. V. (2010). Measurement invariance in confirmatory factor analysis: An illustration using IQ test performance of minorities. *Educational Measurement: Issues and Practice*, 29(3), 39-47.
- Woods, T., & Hanson, D. (2016). *Demographic trends of children of immigrants*. Washington, DC: Urban Institute.

Xie, S., & Li, H. (2019). Development and validation of the chinese preschool readiness scale. *Early Education and Development, 30*(4), 522–539. <https://doi.org/10.1080/10409289.2019.1596462>

APPENDIX A

KSEP Item	Not Yet (1)	Emerging (2)	Almost Mastered (3)	Mastered (4)	Rating
1. Seeks adult help when appropriate	When unable to complete tasks, does not seek and will not accept adult assistance	When unable to complete tasks, does not seek but will accept adult assistance	When unable to complete tasks, asks for assistance without being able to express a specific need	When unable to complete tasks, seeks adult assistance and can express specific needs related to the tasks	
2. Engages in cooperative play activities with peers	Rarely engages with other children and/or may use aggression to disrupt others' cooperative play	Plays alongside other children, but does not engage in interactions with them	Plays next to other children and may initiate conversations or share toys but actions are not part of an ongoing shared activity	Engages with other children in ongoing shared activities (e.g., imaginative play, tag, blocks, etc.)	
3. Exhibits impulse control and self-regulation	Unable to delay having wants and needs met	Distracted by getting wants and needs met, yet able to be redirected by others	Distracted by getting wants and needs met but redirects self	Able to delay wants and needs until appropriate time	
4. Maintains attention to tasks (attention focus, distractibility)	Does not sustain attention to tasks; is frequently distracted by external stimuli	Stays with tasks for a short time but becomes easily distracted by external stimuli	Distracted from tasks by external stimuli but can redirect self back to tasks after becoming distracted	Consistently attends to tasks without becoming distracted by external stimuli	
5. Is enthusiastic and curious about school	Is lethargic, unenthusiastic, and/or indifferent towards school activities and rarely responds to prompting	Is not self-initiating to participate in school activities but will participate when prompted	Shows some excitement and interest in school activities, but only for selected activities	Expresses excitement and interest in participating in a wide range of school activities	
6. Persists with tasks after experiencing difficulty (task persistence, coping with challenges)	Easily becomes visibly frustrated and quits with tasks when difficulty is encountered	Persists with tasks for a short amount of time but quits when difficulty is encountered AND child does not respond to teacher's attempts at redirection	Persists with tasks for a short amount of time but quits when difficulty is encountered BUT child does respond to teacher's attempts at redirection	Usually persists with tasks until completed even after difficulty is encountered; no teacher redirection is needed	