



The Accommodations in Early Years Model: Theoretical Framework and Methodologies for Adapting and Evaluating Developmental Assessments

Pei-Ying Lin^{1*}, Yu-Cheng Lin², Chia Jung Yeh³

1 University of Saskatchewan, Canada. 2 University of Texas Rio Grande Valley, USA. 3 East Carolina University, USA.

Submitted on: May 18, 2018 Accepted on: August 4, 2018

ABSTRACT

Over the years, there has been an increasing interest in adapting classroom and large-scale assessments for school-aged children and adolescents. Despite the recent advances in the fields of educational measurement and special education, current accommodation research rarely involves preschool children. In this paper, we propose a theoretical model consisting of critical elements that involve a continuous and cyclical decision-making process aligned with developmentally appropriate practices addressing children and family needs. We also conducted research on a variety of methodologies that can be employed to examine the accommodated data and evaluate developmental assessments for different study purposes of early childhood special education.

Keywords: Accommodations, assessment, early childhood, special education, quantitative methodologies

INTRODUCTION

Offering appropriate accommodations (e.g., extra time, assistive technology, read-aloud) is critically important for children, adolescents, and adults with special needs when they participate in a variety of assessments (e.g., developmental, classroom-based, or large-scale assessments) across different developmental stages (K-12 and postsecondary education) because these accommodations may help children demonstrate their actual knowledge and skills. Test validity and fairness can be maintained by providing these children with the accommodations that match their special needs, given the content and skills that are being assessed. A child psychologist who precisely gauges a child's current development can provide more accurate and insightful information about the child's developmental stage and progress; the result is that more accessible and equitable learning opportunities and early interventions can then be offered to the child and his/her family. The quality of early intervention and education to the child can be enhanced. To accomplish this ultimate goal, experimental, quasi-experimental, and non-experimental research designs and data analyses have been employed in previous research to examine the effectiveness of accommodations and the validity of each item within a test.

In the United States, there is a demonstrated need to provide accommodations for young children who are experiencing developmental delays. According to the U.S. Department of Education (2016), the percentages of populations of infants and toddlers from birth to age two eligible for special education and services increased from 2.5 percent to 2.9 percent from 2005 through 2014 (N = 299,048 in 2005; N = 350,581 in 2014).

This paper first gives a concise overview of the most fundamental pedagogy for early childhood education- in particular, developmentally appropriate practices (DAP). We proposed an accommodation model, Accommodations in Early Years (AEY), encompassing fundamental theories and major principles that are applicable for developmental assessments administered to diverse young populations and families. It will then discuss the commonly used research methods for developing and validating assessments for children with special needs.

Developmentally Appropriate Practices and Developmental Assessment

The National Association for the Education of Young children (NAEYC) initially introduced Developmentally Appropriate Practice (DAP) in 1987 to describe how young children acquire knowledge (Bredekamp, 1987; Kim, & Han, 2015). The DAP framework has been inviting families and working professionals to revisit their own beliefs about child development and their learning on what is known about each individual child, not only viewing them based on their chronological age, but also considering the variation of each individual and looking at their social cultural context and how it might also have influenced their learning and development (Copple & Bredekamp, 2009). Since then, the lens of how young children learn and develop has experienced some

*Corresponding Author Email: pei-ying.lin@usask.ca

Cite as: Lin, P., Lin, Y., & Yeh, C. J. (2018). The Accommodations in Early Years Model: Theoretical Framework and Methodologies for Adapting and Evaluating Developmental Assessments. *Journal Disability Studies*. 4(1), 14-21.

©IS Publications ISSN: 2454-6623 <http://pubs.iscience.in/jds>



changes (Dickinson, 2002); the most recent position statement was updated and published in 2009 (NAEYC, 2009). The DAP serves as a comprehensive guideline encompassing knowledge of age appropriate, individually appropriate, as well as socially and culturally appropriate for all professionals, practitioners and parents who are working with young children (Bredenkamp, 2016; Copple & Bredenkamp, 2009; NAEYC, 2009). It has been implemented in many areas (e.g., curriculum, assessment, early intervention) and is broadly accepted as a framework for organizing a preschool setting and many personnel preparation programs in early childhood (Klein, & Gilkerson, 2000). The DAP also has become a critical guideline aligned with the Division for Early Childhood's (DEC) recommended practices to support the field of early childhood special education (Division for Early Childhood, 2015; Guralnick & Bruder, 2016; Klein & Gilkerson, 2000). To reach optimal development and maximize learning for young children who have been either at risk, with disabilities, and/or with developmental delays, the DEC was initiated in 1973- its most recent revision in 2015 connected research and practices to provide the latest current knowledge on how to support families and professionals who are working with these children (Division for Early Childhood, 2015).

To enhance the effectiveness of teaching, learning and curricula, professionals and researchers in the field of early intervention, early childhood education, and early childhood special education have identified many challenges that come from using standardized tests and procedures in relation to the unique characteristics of young children (Bagnato, Goins, Pretti-Frontczack, & Neisworth, 2014; Bracken, & Nagle, 2007; Squires, 2015). Nagle (2014) recognized several exclusive features of young children during assessment, including swift developmental changes, variations in their development, behavioral fluctuations, and situational variables. For example, young children might perform differently in different settings (Squires, 2015). It might be more challenging to identify children who are atypical in their development because of their exclusive characteristics in early childhood. However, it is crucial to identify early developmental delays so that children's health and well-being can be properly supported (Singh, Squires, Yeh, Heo, & Bian, 2016). To address the challenges discussed above, various developmental assessments have been used for assessment and screening purposes to obtain quick information about development. A variety of screening assessments can be utilized in numerous authentic settings (e.g., preschool classrooms, at home, clinic)(Squires, 2015; Singh et al., 2016). Several research studies have revealed that parents' reports about their children's development are often accurate and provide reliable information with low costs and are more time-efficient compared to other standardized testing (Johnson, & Marlow, 2006; Singh et al., 2016).

Developmental assessment is an ongoing process of gathering information from parents, caregivers and/or applicable professionals in order to provide early identification, eligibility determination, monitoring of child progress, and

optimizing child development (McLean, Wolery, & Bailey, 2004; Snow, & Van Hemel, 2008). Developmental assessment consists of a linear developmental trajectory with a series of different competencies and various skills, which might demonstrate different developmental status across a range of different developmental domains (Leung, Mak, Lau, Cheung, & Lam, 2010). Researchers have indicated that developmental assessments should be age-appropriate not only for the identification but also for monitoring of young children who are at risk or with developmental difficulties (Johnson & Marlow, 2006).

Compared to school age children learning in a K-12 school setting, appropriate educational opportunities for all infants, toddlers, and preschoolers are lacking (Macy, Marks, & Towle, 2014). Some infants and toddlers might stay at home and receive care from their parents and/or relatives. Parents or primary caregivers may complete parental questionnaires, checklists, or assessments with respect to their child's development. One of the parent-completed screening assessments, which has been extensively used globally, is the Ages and Stages Questionnaires, Third Edition (ASQ-3) (Squires & Bricker, 2009; Singh, Yeh, & Blanchard, 2017). This questionnaire can be used by parents and implemented by a wide range of professionals like home visitors, early educators, and health care providers who possess a 6th grade reading level (Ringwalt, 2008). Researchers reported that the ASQ-3 has been translated into several languages—including Chinese, French, Hindi, and Spanish—for early identification and early screening of developmental delays (Singh, Yeh, & Blanchard, 2017). It comprises five developmental domains: communication, gross motor, fine motor, problem solving, and personal-social skills. Completing this questionnaire with 30 items for each age range take less than 20 minutes (Ringwalt, 2008).

Accommodation Principles and Theory

In recent years, there has been an increasing amount of literature on accommodations for students with diverse needs, including those with exceptionalities and English language learners. Based on our synthesis of major hypotheses and existing evidence, this paper proposes the *Accommodations in Early Years Model* as a theoretical framework for adapting developmental assessments and facilitating developmentally appropriate practices in early childhood special education. Given that current accommodation research rarely involves preschool children, we believe that the accommodation theory and principles discussed in this paper can be applied to developmental assessments in early childhood special education. Our theoretical model presents critical components in Figure 1. In this model, the developmentally appropriate practices (the inner circle) integrates with the five surrounding elements: children's and caregivers' special needs, assessment characteristics and purposes, test validity, accommodation types, and developmental milestones and learning outcomes. In addition, each element contains its own sub-elements. For example, the "types" element consists of four sub-elements: setting, timing, presentation, and response modalities. This

framework is represented in concentric circles as the decision-making process is cyclical and continuous. The developmentally appropriate practices at the center of the model lay out a critical foundation of all elements and sub-elements. Specifically, a decision-making process begins at Element 1 (children's and care givers' special needs); the process then continues around the circle to Element 5 (developmental milestones). However, the process does not end here; rather, it continues around the circle to re-evaluate the effectiveness and impact of the accommodation policy and practices. (see Figure 1)

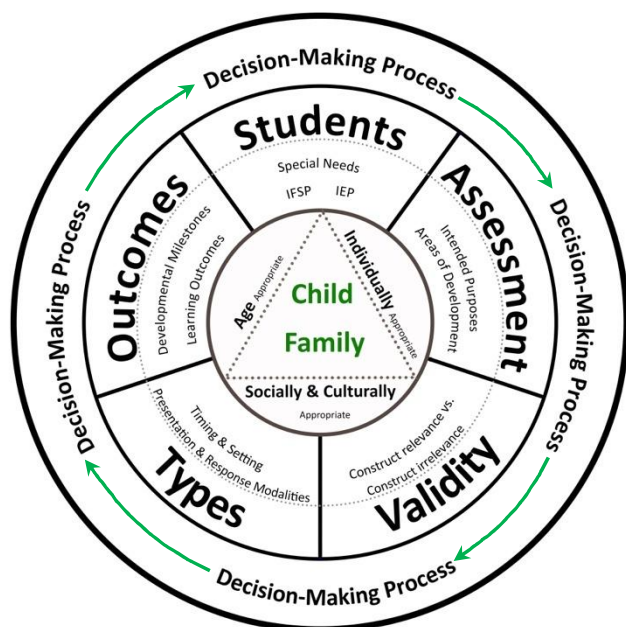


Figure 1. The *Accommodations in Early Years Model*.

Element 1: Children's and Caregivers' Special Needs

To illustrate our points, we use a case example to elaborate on the *Accommodations in Early Years Model* we propose in this paper.

Elise is a Chinese mother with a learning disability who speaks English as a second language; she has a child with Down's syndrome in the United States. Her special needs make it difficult for her to read quickly. Providing extra time, a translator, and assistive technology (e.g., text-to-speech, speech-to-text software) may permit Elise to read and respond to the questions on a parent questionnaire of early childhood motor skills for her child. In addition, the child can also have additional time to complete the activities asked of her by her mother (e.g., "draw a picture of your Mom," "pull the zipper up for me", "button the buttons of your coat", or "use your chopsticks to eat noodles"), making the assessment results more valid. However, if Elise were simply provided with a visual magnification device, large-print, or a sign language interpreter, these accommodations wouldn't be helpful for her to read and respond to her child.

As seen in the example above, it is important to match children's and caregivers' special needs with specific types of accommodations to maintain test validity. To align with the

guidelines of developmentally appropriate practices, the assessment practices should take a child's age, social, and cultural background into consideration to ensure that the practices are age-appropriate, individually appropriate, as well as socially and culturally appropriate (Bredenkamp, 2016; Copple, & Bredenkamp, 2009; NAEYC, 2009). For instance, learning how to use chopsticks is essential for children raised in Chinese culture. It is recommended that accommodations that are documented in students' individual education programs (IEP) and have been used in a child's classroom should be provided in large-scale or standardized assessments (IDEA, 2004; Johnson & Monroe, 2004). In the case of young learners, we suggest that appropriate accommodations should be documented in each child's individualized family service plan (IFSP) or IEP. Moreover, they should be provided consistently across a child's natural environments (e.g., home, community, school), or least restrictive environments for a variety of activities (e.g., play activities, classroom instruction and assessments). The IFSP for infants and toddlers before age three and IEP for preschoolers after age three developed collaboratively by a multidisciplinary team must be family-centered, and deliver varied services and programs to meet the child's and family needs- such as parent counseling and training and social work (Bruder, 2010; Corr & Danner, 2014; Tomasello, Manning, & Dulmus, 2010). The Individuals with Disabilities Education Improvement Act (IDEIA, 2004) in the United States mandates that the IFSP or IEP be in place for each child with a disability who is eligible for special education and services. Even though current research on test accommodations doesn't discuss the needs of care givers, in this paper we emphasize that appropriate accommodations should not only be available for young children, but also for their primary caregivers, especially for those with special needs and English language learners. In other words, equitable and appropriate accommodations need to be documented in each child's IEPs and IFSP. Without appropriate accommodations, primary caregivers who fill out a questionnaire or an assessment concerning their child's progress and development may not be able to provide accurate information to support optimal development.

Element 2: Purposes of Assessment

The purposes of different types of assessments may differ from one individual to another (e.g., norm-referenced, criteria-referenced, authentic, or portfolio assessment). To meet individual and family needs, appropriate diagnostic instruments and procedures should be employed to identify if an infant or toddler is experiencing developmental delays in one or more of five major areas of development including physical, cognitive, communication, social-emotional, and adaptive development (IDEA, 2004). As such, the purposes of each assessment for developmental delays should be clearly identified and defined prior to making the accommodation decisions, because each may substantially influence the selection, design and implementation of accommodations. Moreover, the interpretations of the results gathered from a questionnaire or an

assessment instrument should be reliable and valid to support the claims about what knowledge and skills the assessment is intended to assess (American Educational Research Association et al., 2014). In other words, with certain changes to the regular test administration, the assessment can still measure what it is supposed to measure (test construct, for example, demonstrating fine motor skills independently without an adult's assistance in Elise's case described above); this idea directly connects to our third element—test validity.

Element 3: Test Validity

Appropriate and valid accommodations that are used to address these preschool children's diverse needs should be used to enhance student learning in the inclusive classroom and strengthen the validity of developmental assessments. Invalid instructional and assessment accommodations, however, should be avoided because they may violate the validity of an assessment and cannot produce meaningful test results for children with special needs (Fuchs et al., 2000a, 2000b). Therefore, it is critical to outline allowed and not-allowed accommodations in test manuals and policy documents as well as to provide solid justifications for such guidelines and policies (National Research Council, 2004). In Elise's case, the parent questionnaire was supposed to measure her child's motor skills without an adult's assistance; holding the child's hand to assist her in completing the activities (e.g., drawing a picture, buttoning the buttons, pulling up a zipper, or using chopsticks to eat noodles) would be deemed as inappropriate because the test validity of this parent questionnaire may be compromised thereby. In other words, the assistance from the parent (a modification) threatens the validity of this assessment. To avoid such validity threats, the use of test modifications is not permitted for a vast majority of developmental assessments (Johnson & Marlow, 2006). While accommodations have received little attention in either research or practices of developmental assessments, the BDI-2 (Battelle Developmental Inventory-2) (Newborg, 2010) has paid explicit attention to accommodations for students with disabilities. The BDI-2, an assessment for children from birth to age eight, assesses a child's adaptive, motor, communication, cognition, and personal-social skills. Unlike many other tests, it provides guidelines for adapting the standardized test administration for young children with visual, hearing, or speech impairments, emotional or behavioral disorders, physical disabilities and multiple exceptionalities to ensure that the adaptations do not compromise the validity of this norm-referenced assessment.

Element 4: Types of Accommodations

Theoretically, a test accommodation is a change in an aspect of the test administration that is not related to the knowledge or skill the test is intended to measure (e.g., American Educational Research Association et al., 2014; Bolt & Thurlow, 2007; Fuchs, Fuchs, Eaton, Hamlett, & Karns, 2000a; Lindstrom & Gregg, 2007). These changes are usually in one or more of four aspects of the test administration: timing, setting, presentation modality, and response modality (Fuchs, Fuchs, & Capizzi, 2005; National Research Council, 2004). Changes that do affect what

the test measures are sometimes referred to as modifications. Examples for each of four major categories are (1) timing: extended testing time, frequent supervised breaks; (2) setting: alternate testing room, individual or small group administration, preferential seating within the regular classroom; (3) presentation modality: large print, Braille, read-aloud, sign language or oral interpreter, assistive technology such as text-to-speech software; and (4) response modality: scribe, tape recording of responses, Braille, sign language or oral interpreter, assistive technology such as speech-to-text software, computer or word processor.

Element 5: Developmental Milestones and Learning Outcomes

An infant or a toddler may experience significant developmental delays if h/she is not achieving milestones in or more of the areas of development (physical, cognitive, communication, social-emotional, and adaptive development) (IDEA, 2004). As the child grows, h/she is expected to progressively achieve integrated objectives and outcomes across the five areas of development indicated in his/her IFSP or IEP. Throughout each child's pathway used to achieve each developmental milestone, we suggest stakeholders make good use of accommodations that contain all elements discussed in this paper. In addition, it is critical to evaluate and review the effectiveness and appropriateness of an accommodation and multiple combinations of accommodations for each child (Kettler, 2012; Lin & Lin, 2015, 2016). Furthermore, to reduce the potential adverse effects of inconsistent use of accommodations, we recommend that educators pay close attention to including accommodations into the transition planning process for the young children.

Research Methodologies for Evaluating Accommodated Data and Assessments

In previous sections, we proposed and discussed fundamental theory and principles for accommodations that provide a comprehensive framework for teaching practices and research. In this section, we will shift our focus from a theory to research methods for validating the assessment data using accommodations. The accommodation research to date has focused on two main aspects: (1) examining the effectiveness of specific accommodations (e.g., extended time, read-aloud) for children with varied special needs; and (2) test validity and fairness for all children, including children with and without special needs. This section begins by discussing each of the two aspects as well as examining the pros and cons related to each aspect.

To investigate the effects of single or multiple accommodations, the interaction hypothesis and differential boost are two major hypotheses that have been utilized to compare the group differences in test results (e.g., Bolt & Ysseldyke, 2006; Sireci, Scarpati, & Li, 2005). According to the interaction hypothesis, only the children with special needs who need the accommodations would benefit from them, whereas the test performances of children without special needs should be the same or almost the same in both accommodated and non-

accommodated conditions. Compared with the interaction hypothesis, a less stringent standard has been employed by the differential boost. It suggests that both groups of children can benefit from the accommodations; however, those effective and appropriate accommodations would help accommodated children with special needs make larger gains than accommodated children without special needs in both testing conditions. Based on these key concepts, two main hypotheses have been widely used to evaluate the effectiveness of particular accommodations in previous studies. Studies have often compared the overall test scores of accommodated and non-accommodated groups in experimental or quasi-experimental studies (Cahalan-Laitusis, 2010; Fletcher et al., 2006). Common statistical analyses have been used to examine the between- and within- group variances and interaction effects between groups and an accommodation, including the analysis of variance (ANOVA)/repeated measures ANOVA (Cahalan-Laitusis, 2010; Lewandowski, Lovett, Parolin, Gordon, & Coddling, 2007; Lewandowski, Lovett, & Rogers, 2008; Simth & Riccomini, 2013), and the analysis of covariance (ANCOVA)/repeated measures ANCOVA (Bolt & Thurlow, 2007; Cahalan-Laitusis, 2010; Fletcher et al., 2006).

However, previous research on the effects of different types of accommodations often yielded mixed results, such as extended time and read-aloud (e.g., Fuchs et al., 2000b; Lewandowski et al., 2007, 2008; Thompson, Blount, & Thurlow, 2002). The inconsistent findings reflect a wide range of variations among the studies, such as the heterogeneous student characteristics, the sample sizes, the age of the studied groups, the subject domains, and the research methods employed (Cahalan-Laitusis, 2010; Lewandowski et al., 2008). Accommodated and non-accommodated data in previous studies were often examined by the main hypotheses discussed above. There have been concerns mainly with how the mixed results may weaken the assumptions about which accommodation benefits a group with special needs more than the non-disabled group, or which accommodation only improves the test performances of accommodated children with special needs (Sireci et al., 2005). More importantly, comparing the sum of accommodated test scores with non-accommodated scores may not address the concerns about test validity. Such a total score comparison approach may not indicate whether the test constructs remained the same, or whether they were simply inflated by altering the testing procedures, conditions or environment. However, it is imperative to maintain the test validity and fairness for all children; accommodations should not give an unfair advantage or disadvantage to a particular group of children (e.g., accommodated or non-accommodated group) and should not inflate the test scores (Fuchs & Fuchs, 2001). To scrutinize the construct validity of accommodated assessments, alternative statistical methods have been developed. The following section reviews the most widely adopted analysis, differential item functioning (DIF), for modeling item-level responses of accommodated and non-accommodated examinees at the same trait level (e.g., math

performance). As Furr and Bacharach indicate, “[p]erhaps the best way to evaluate construct bias is a procedure called differential item functioning analysis” (p. 308).

A range of differential item functioning (DIF) methods has been used to detect if the probability of examinees in two groups (e.g., females versus males; accommodated versus non-accommodated group) answering an item correctly differs-given that they have the same latent traits or abilities. One of the most popular methods is the Mantel-Haenszel (MH) procedure for detecting an item that functions differently for a focal (accommodated) or reference group (standard testing condition) (Mazor, Clauser, & Hambleton, 1993). It is important to note that an item does not show DIF if the group differences were found in groups with heterogeneous abilities. The estimates of MH are the odds ratio calibrated from a multi-way contingency table ($2 \times 2 \times S$), where it represents the item score of a dichotomous item (0 or 1), respondents’ memberships (focal or reference group), and the possible scores of a test (Holland & Thayer, 1988; Linacre & Wright, 1989; Zwick, 2002).

In contrast to MH, the methods for detecting a DIF item based on item response theory (IRT) are independent of examinees’ abilities and do not confound with the characteristics of samples. By applying the IRT-based methods, the estimation of each item can be calibrated precisely- thus, they are useful for identifying DIF items through an examination of group differences in item difficulty and detecting possible items that might be biased for a given group of examinees and warrant further investigation (Hambleton, Swaminathan, & Rogers, 1991). Essentially, IRT is a psychometric approach estimating an individual’s response to an item based on qualities of the person and the item; thus, its procedures provide rich information about each test taker, item, and test (Furr & Bacharach, 2014). The fundamental IRT model is often called the Rasch model, known as a one-parameter logistic model (1PL), and is often presented as:

$$P(X_{ij} = 1 | \theta_j, b_i) = \frac{\exp(\theta_j - b_i)}{1 + \exp(\theta_j - b_i)}$$

where P represents the probability of a test taker j with ability responding to a given item i correctly. Moreover, θ_j refers to the examinee j ’s level of latent ability and b_i reflects the difficulty of item i . The details of Rasch and extended IRT models (2PL, 3PL, and 4PL) have been discussed at great length in a large number of publications in the measurement literature (e.g., Embretson & Reise, 2000; Hambleton, Swaminathan, & Rogers, 1991; Loken & Rulison, 2010).

Variations of IRT-based methods have been employed to examine the accommodation-related DIF, including the mixture Rasch model (MRM)(Cho, Lee, & Kingston, 2012; Cohen, Gregg, & Deng, 2005; Scarpati, Wells, Lewis, & Jirka, 2009), Item Response Theory Likelihood-Ratio Test for Differential Item Functioning (IRTLRDIF) (Bolt & Ysseldyke, 2006, 2008), the hierarchical generalized linear model (HGLM) (Randall, Cheong, & Engelhard, 2011), the multilevel multidimensional

mixture IRT (Finch & Finch, 2013), and the multilevel measurement modeling (MMM) (Beretvas, Cawthon, Lockhart, & Kaye, 2012; Lin, Childs, & Lin, 2016; Lin & Lin, 2014).

CONCLUSION

This paper set out to discuss the fundamental principles of accommodations that are widely applicable for adapting development assessments for young preschool children with special needs. We propose a theoretical model, *Accommodations in Early Years*, consisting of key elements, including a child's and caregivers' special needs, the intended purposes of a developmental assessment, the test validity of the assessment, types of accommodations, the child's developmental and learning outcomes, and an ongoing decision-making process (Figure 1). Moreover, each element in the model is embedded within the framework of developmentally appropriate practices in early childhood education. We recommend that these key elements in this model should be taken into consideration when adapting a developmental assessment for a young child with special needs. Furthermore, we conducted research on a variety of methodologies that have been employed to examine the accommodated data for different study purposes. The quantitative methodologies involve statistical analyses based on the classical test theory and item response theory. Our discussion suggests that it would be useful to utilize the same, modified, or extended methodologies in a number of possible future studies.

This paper makes several contributions to the current literature review. First, our theoretical accommodation model is a comprehensive assessment framework that is applicable and practical for a wide range of developmental assessments. This model is generalizable across children with varied characteristics and special needs. In addition, we have also taken the family needs of primary caregivers of young children into account- in this we differ from existing literature on accommodations. Second, our accommodation model considers different aspects of accommodations such as the intended purposes of a developmental assessment and its test validity, in addition to a traditional view of accommodations. Finally, we synthesized the current literature on varied methodologies that can be directly applied to evaluate, develop, and improve the quality of a given developmental assessment for young children who are in need of accommodations throughout their learning and developmental trajectories. The conclusions in this paper have valuable implications for both practice and research.

REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: Author.
- Bagnato, S. J., Goins, D. D., Pretti-Frontczak, K., & Neisworth, J. T. (2014). Authentic assessment as "best practice" for early childhood intervention: National consumer social validity research. *Topics in Early Childhood Special Education, 34*(2), 116-127.
- Beretvas, S. N., Cawthon, S. W., Lockhart, L. L., & Kaye, A. D. (2012). Assessing impact, DIF, and DFF in accommodated item scores: a comparison of multilevel measurement model parameterizations. *Educational and Psychological Measurement, 72*(5), 754-773.
- Bolt, S. E., & Thurlow, M. L. (2007). Item-level effects of the read-aloud accommodation for students with reading disabilities. *Assessment for Effective Intervention, 33*(1), 15-28.
- Bolt, S. E., & Ysseldyke, J. E. (2006). Comparing DIF across math and Reading/Language arts tests for students receiving a read-aloud accommodation. *Applied Measurement in Education, 19*(4), 329-355.
- Bolt, S. E., & Ysseldyke, J. E. (2008). Accommodating students with disabilities in large-scale testing: A comparison of differential item functioning (DIF) identified across disability types. *Journal of Psychoeducational Assessment, 26*(2), 121-138.
- Bracken, B. A., & Nagle, R. J. (2007). *Psychoeducational assessment of preschool children* (4th ed). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cahalan-Laitusis, C. (2010). Examining the impact of audio presentation on tests of reading comprehension. *Applied Measurement in Education, 23*(2), 153-167. doi: 10.1080/08957341003673815
- Cho, H., Lee, J., & Kingston, N. (2012). Examining the effectiveness of test accommodation using DIF and a mixture IRT model. *Applied Measurement in Education, 25*(4), 281-304.
- Clauser, B. E., Mazor, K. M., & Hambleton, R. K. (1993). The effects of purification of the matching criterion on the identification of DIF using the Mantel-Haenszel procedure. *Applied Measurement in Education, 6*, 269-279.
- Cohen, A. S., Gregg, N., & Deng, M. (2005). The role of extended time and item content on a high-stakes mathematics test. *Learning Disabilities Research & Practice, 20*(4), 225-233.
- Copple, C., & Bredekamp, S. (Eds.). (2009). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8* (3rd ed). Washington, DC: NAEYC.
- Corr, C., & Danner, N. (2014). Court-appointed special advocate strong beginnings: raising awareness across early childhood and child welfare system. *Early Child Development and Care, 184* (9-10), 1436-1446. doi: 10.1080/03004430.2013.845564
- Dickinson, D. K. (2002). Shifting images of developmentally appropriate practice as seen through different lenses. *Educational Researcher, 31*(1), 26-32. doi:10.3102/0013189X031001026
- Division for Early Childhood (2015). *DEC recommended practices: Enhancing services for young children with*

- disabilities and their families* (DEC Recommended Practices Monograph Series No. 1). Los Angeles, CA: DEC.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Hillsdale, NJ: Erlbaum.
- Finch, W. H., & Hernández Finch, M. E. (2013). Investigation of specific learning disability and testing accommodations based differential item functioning using a multilevel multidimensional mixture item response theory model. *Educational and Psychological Measurement, 73*(6), 973-993.
- Fletcher, J., Francis, D. J., Boudousquie, A., Copeland, K., Young, V., Kalinowski, S., & Vaughn, S. (2006). Effects of accommodations on high-stakes testing for students with reading disabilities. *Council for Exceptional Children, 72*(2), 136-150.
- Fuchs, L. S., & Fuchs, D. (2001). Helping teachers formulate sound test accommodation decisions for students with learning disabilities. *Learning Disabilities Research & Practice, 16*(3), 174-181.
- Fuchs, L. S., Fuchs, D., Eaton, S. B., Hamlett, C. L., & Karns, K. M. (2000). Supplemental teacher judgments of mathematics test accommodations with objective data sources. *School Psychology Review, 29*(1), 65-85.
- Fuchs, L. S., Fuchs, D., & Capizzi, A. M. (2005). Identifying appropriate test accommodations for students with learning disabilities. *Focus on Exceptional Children, 37*(6), 1-8.
- Furr, R. M., & Bacharach, V. R. (2014). *Psychometrics: An Introduction* (2nd ed.). Thousand Oaks, CA: Sage.
- Guralnick, M. J., & Bruder, M. B. (2016). Early childhood inclusion in the United States: Goals, current status, and future directions. *Infants and Young Children, 29*(3), 166-177.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the mantel-haenszel procedure. In H. Wainer, & H. I. Braun (Eds.), *Test validity* (pp. 129-145). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Individuals with Disabilities Education Act of 2004, 20 U.S.C. § 1400. (2004).
- Johnson, S., & Marlow, N. (2006). Developmental screen or developmental testing? *Early Human Development, 82*(3), 173-183. doi: 10.1016/j.earlhumdev.2006.01.008
- Johnson, E., & Monroe, B. (2004). Simplified language as an accommodation on math tests. *Assessment for Effective Intervention, 29*(3), 35-45.
- Kim, H. K., & Han, H. S. (2015). Understanding early childhood teachers' beliefs and self-stated practices about social competence instructional strategies in the context of developmentally appropriate practice: A comparison of preservice and in-service teachers in the United States. *European Early Childhood Education Research Journal, 23*(4), 476-496. doi:10.1080/1350293X.2015.1087152
- Klein, N. K., & Gilkerson, L. (2000). Personnel preparation for early childhood intervention programs. In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of Early Childhood Intervention*. (2nd ed., pp. 454-483). New York, NY: Cambridge University Press.
- Leung, C., Mak, R., Lau, V., Cheung, J., & Lam, C. (2010). Development of a preschool developmental assessment scale for assessment of developmental disabilities. *Research in Developmental Disabilities, 31*(6), 1358-1365. doi:10.1016/j.ridd.2010.07.004
- Lewandowski, L. J., Lovett, B. J., Parolin, R., Gordon, M., & Coddling, R. S. (2007). Extended time accommodations and the mathematics performance of students with and without ADHD. *Journal of Psychoeducational Assessment, 25*(1), 17-28.
- Lewandowski, L. J., Lovett, B. J., & Rogers, C. L. (2008). Extended time as a testing accommodation for students with reading disabilities: Does a rising tide lift all ships? *Journal of Psychoeducational Assessment, 26*(4), 315-324.
- Lin, P., & Lin, Y. (2014). Examining student factors in sources of setting accommodation DIF. *Educational and Psychological Measurement, 74*(5), 759-794. doi:10.1177/0013164413514053
- Lin, P., Childs, R. A., & Lin, Y. (2016). *Untangling complex effects of disabilities and accommodations within a multilevel IRT framework*. *Quality & Quantity, 50* (6), 2767-2788. doi: 10.1007/s11135-015-0288-8
- Linacre, J. M., & Wright, B. D. (1989). Mantel-Haenszel DIF and PROX are equivalent!. *Rasch Measurement Transactions, 3*, 52-53.
- Lindstrom, J. H., & Gregg, N. (2007). The role of extended time on the SAT® for students with learning disabilities and/or attention-deficit/hyperactivity disorder. *Learning Disabilities Research & Practice, 22*, 85-95.
- Loken, E., & Rulison, K. L. (2010). Estimation of a four-parameter item response theory model. *British Journal of Mathematical and Statistical Psychology, 63*(3), 509-525.
- Macy, M., Marks, K., & Towle, A. (2014). Missed, misused, or mismanaged: Improving early detection systems to optimize child outcomes. *Topics in Early Childhood Special Education, 34*(2), 94-105. doi:10.1177/0271121414525997
- McLean, M. E., Wolery, M., & Bailey, D. B. (2004). *Assessing infants and preschoolers with special needs* (3rd ed.). Upper Saddle River, NJ: Merrill.
- Moore-Brown, B. J., Montgomery, J. K., Bielinski, J., & Shubin, J. (2005). Responsiveness to intervention: Teaching before testing helps avoid labeling. *Topics in Language Disorders, 25*(2), 148-167.
- Nagle, R. J. (2014). Issues in preschool assessment. In B. A. Bracken, & R. J. Nagle (Eds.), *Psychoeducational assessment of preschool children* (4th ed., pp. 29-48). Mahwah, NJ: Lawrence Erlbaum Associates.
- National Association for the Education of Young Children (NAEYC). (2009). *Developmentally appropriate practice in early childhood programs serving children from birth*

- through age 8. *A position statement of the National Association for the Education of Young Children*. Retrieved from <http://www.naeyc.org/files/naeyc/file/positions/PSDAP.pdf>
- National Research Council. (2004). *Keeping score for all: The effects of inclusion and accommodation policies on large-scale educational assessments*. Committee on Participation of English Language Learners and Students with Disabilities in NAEP and Other Large-Scale Assessments. Judith A. Koenig and Lyle F. Bachman, Editors. Board on Testing and Assessment, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Phillips, S. E. (1994). High-stakes testing accommodations: Validity versus disabled rights. *Applied Measurement in Education*, 7(2), 93-120.
- Randall, J., Cheong, Y. F., & Engelhard, G. (2011). Using explanatory item response theory modeling to investigate context effects of differential item functioning for students with disabilities. *Educational and Psychological Measurement*, 71(1), 129-147.
- Ringwalt, S. (2008). *Developmental screening and assessment instruments with an emphasis on social and emotional development for young children ages birth through five*. Chapel Hill: The University of North Carolina, FPG Child Development Institute, National Early Childhood Technical Assistance Center.
- Scarpati, S. E., Wells, C. S., Lewis, C., & Jirka, S. (2011). Accommodations and item-level analyses using mixture differential item functioning models. *The Journal of Special Education*, 45(1), 54-62.
- Singh, A., Squires, J., Yeh, C. J., Heo, K. H., & Bian, H. (2016). Validity and reliability of the developmental assessment screening scale. *Journal of Family Medicine and Primary Care*, 5(1), 124-128. doi: 10.4103/2249-4863.184636
- Singh, A., Yeh, C., & Blanchard, S. B. (2017). The Ages and Stages Questionnaire: A global screening scale. *Boletín Médico del Hospital Infantil de México*, 74(1), 5-12. doi: 10.1016/j.bmhmx.2016.07.008
- Sireci, S. G., Scarpati, S. E., & Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research*, 75(4), 457-490.
- Smith, G. W., & Riccomini, P. J. (2013). The effect of a noise reducing test accommodation on elementary students with learning disabilities. *Learning Disabilities Research & Practice*, 28(2), 89-95.
- Snow, C. E., & Van Hemel, S. B. (Eds.). (2008). *Early childhood assessment: Why, what, and how*. Washington, DC: The National Academies Press.
- Squires, J. (2015). Assessment: Guiding principles for accurate and efficient decision making. In R. M. Santos (Ed.), *DEC recommended practices: Enhancing services for young children with disabilities and their families* (pp. 37-52). Los Angeles, CA: DEC.
- Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaires, Third Edition (ASQ-3)*. Baltimore, MD: Brookes Publishing.
- Thompson, S., Blount, A., & Thurlow, M. (2002). A summary of research on the effects of test accommodations: 1999 through 2001 (Technical Report34). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Tomasello, N. M., Manning, A. R., & Dulmus, C. N. (2010). Family-centered early intervention for infants and toddlers with disabilities. *Journal of Family Social Work*, 13(2), 163-172.
- U.S. Department of Education. (2016). 38th annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2016. Washington, DC: Office of Special Education and Rehabilitative Services. Retrieved from <https://www2.ed.gov/about/reports/annual/osep/index.html>
- Zuriff, G. E. (2000) Extra examination time for students with learning disabilities: An examination of the maximum potential thesis. *Applied Measurement in Education*, 13, 99-117.
- Zwick, R. (2002). The assessment of differential item functioning in computer adaptive tests. In W. Linden, & G. W. Glas (Eds.), *Computerized adaptive testing: Theory and practice* (pp. 221-244) Springer Netherlands.

Source of Funding: None

Conflict of interest: None