University of Kentucky
UKnowledge

Theses and Dissertations--Education Sciences | College of Education

2023

# SCORE EQUATING BETWEEN AEPS-2 AND AEPS-3 FOR 0-3 YEAR OLDS

Yuyan Xia
*University of Kentucky*, xyyxiayuyan@gmail.com
Author ORCID Identifier:
https://orcid.org/0000-0003-2122-4268
Digital Object Identifier: https://doi.org/10.13023/etd.2023.334

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation
Xia, Yuyan, "SCORE EQUATING BETWEEN AEPS-2 AND AEPS-3 FOR 0-3 YEAR OLDS" (2023). *Theses and Dissertations--Education Sciences*. 134.
https://uknowledge.uky.edu/edsc_etds/134

This Doctoral Dissertation is brought to you for free and open access by the College of Education at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Education Sciences by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Yuyan Xia, Student

Dr. Kelly D. Bradley, Major Professor

Dr. Jane M. Jensen, Director of Graduate Studies

# SCORE EQUATING BETWEEN AEPS-2 AND AEPS-3 FOR 0-3 YEAR OLDS

---

DISSERTATION

---

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Education
at the University of Kentucky

By
Yuyan Xia
Lexington, Kentucky
Director: Dr. Kelly Bradley, Professor of Education Policy & Evaluation

Lexington, Kentucky
2023

Copyright © Yuyan Xia 2023
https://orcid.org/0000-0003-2122-4268

ABSTRACT OF DISSERTATION

SCORE EQUATING BETWEEN AEPS-2 AND AEPS-3 FOR 0-3 YEAR OLDS

Over the past two decades, the emphasis on educational equity in early childhood education (ECE) and early childhood special education (ECSE) has highlighted the importance of assessment through policies and regulations. Ensuring accurate assessment scores is a fundamental aspect of this trend. The release of the Assessment, Evaluation, and Programming System for Infants and Children, Third Edition (AEPS-3) in December 2021 led to a shift from the Second Edition (AEPS-2) in child development scoring. In order to harmonize the previous and updated assessment versions for children aged 0-3 across six developmental domains, a common item non-equivalent design, featuring fixed parameter calibration equating (known as 'anchoring'), is utilized within the Rasch framework.

A total of 18,411 cases from the AEPS-2 Test Level I and 317 cases from the AEPS-3 Test were utilized to assess scale quality. The psychometric properties of both assessment versions were evaluated using the rating scale Rasch model, revealing a good model-data fit. Two sets of anchor items, selected based on either identical or functional matching methods, were determined using the cosine similarity coefficient and subsequently validated through expert content analysis. These anchor item sets demonstrated acceptable quality. The research then examined the impact of different anchor sets on person parameter estimation during the anchoring process. Ultimately, the study produced person measure and observed score conversion tables between AEPS-2 and AEPS-3. The resulting conversion tables provide valuable insights into the relationship between the old and updated assessment versions.

These findings contribute to equating methodology, ECE/ECSE, and education policy. As an early implementation of functional matching anchoring equating in the ECSE field, this study provides a practical model for score equating transformation that can be applied across both early childhood education and special education sectors. In the early childhood education area, it supports the ongoing refinement of assessment tools in early childhood education, helping practitioners make more informed decisions about child development. By leveraging the psychometric model, the research contributes to improving the quality of assessment tools for early childhood education practitioners, leading to better outcomes for children in these critical developmental stages. Another important contribution of this study is that it reflects the assessment requirements in special education and connects education policy with research goals. This ensures that assessments remain consistent, fair, and accurate, enabling educators and specialists to effectively track and support children's development over time, ultimately improving educational equity.

KEYWORDS: score equating, AEPS, Rasch model, assessment, functional matching, anchoring.

Yuyan Xia

07/20/2023

Date

SCORE EQUATING BETWEEN AEPS-2 AND AEPS-3 FOR 0-3 YEAR OLDS

By
Yuyan Xia

Dr. Kelly Bradley
Director of Dissertation

Dr. Jane McEldowney Jensen
Director of Graduate Studies

07/20/2023
Date

DEDICATION

To *My beloved parents Xinhua Xia & Lianfang Zhang*

*献给我亲爱的父母 夏新华&章莲芳*

ACKNOWLEDGMENTS

The following dissertation, while an individual endeavor, has benefited from the insights and guidance of many. First and foremost, my Dissertation Chair, Dr. Kelly Bradley, embodies the high-quality scholarship I aspire to. Dr. Michael Peabody provided prompt and invaluable feedback throughout the dissertation process, enabling me to complete this project on schedule. Special thanks to Dr. Grisham, who assisted in communicating with the publisher and securing necessary data sources for the research, and Brooks Publishing for their willingness to share their data.

Next, I express my deepest gratitude to my dissertation committee and outside-reader, including Dr. Joseph Waddington, Dr. Jennifer Grisham, Dr. John Thelin, and Dr. Donald Bruce Ross III. Each committee member offered valuable insights that guided and challenged my thinking, enhancing the final product. Special recognition also goes to Dr. Shannon Sampson and Dr. Lin Yuan for their guidance, learning opportunities, and encouragement.

Beyond this academic support, I was blessed with an abundance of love, care, and encouragement from my family. My parents' unwavering faith in me provided a sturdy foundation for my work. I'm truly blessed to be their daughter. My older brother, Yuheng Xia, along with close friends Emily Nowell, Ioana Yu, Jackie Hogue, Jue Lu, Mary Boyd, Minjuan Chen, Pin Wei, Rui Jin, Siqi Liu, Tina Smith, Wei Mao, Wenbin Hu, Xiaocheng Yu, Yan Zhou, Yucong Sang, Yuxuan Zhang, and others, formed a strong emotional support system during the pandemic, propelling me to this point in my academic journey.

TABLE OF CONTENTS

ACKNOWLEDGMENTS ..... iii

LIST OF TABLES ..... vii

LIST OF FIGURES ..... viii

Chapter 1 INTRODUCTION ..... 1

*1.1 Background ..... 1*

*1.2 Theoretical Preparation and First Assessment ..... 2*

1.2.1 The Role of Assessment in ECSE ..... 3

*1.3 History of AEPS ..... 6*

1.3.1 Preparation and Meeting ..... 6

1.3.2 Early Versions of Instrument – API & EPS ..... 7

1.3.3 AEPS-2 ..... 7

1.3.4 Online System AEPSi & AEPS-3 ..... 9

1.3.5 The Needs of Conversion Table between AEPS-2 and AEPS-3 ..... 9

*1.4 Rational for Score Equating ..... 10*

*1.5 Research Questions ..... 13*

*1.6 Significance ..... 14*

*1.7 Organization of Dissertation ..... 15*

Chapter 2 LITERATURE REVIEW ..... 17

*2.1 Curriculum-based Assessment ..... 17*

*2.2 The Purposes of AEPS ..... 18*

*2.3 Psychometric Properties of AEPS ..... 20*

2.3.1 Reliability ..... 20

2.3.2 Validity ..... 21

2.3.3 Utility ..... 24

2.3.4 Psychometric Properties of AEPS under Rasch Framework ..... 25

2.3.5 Summary and Limitation of Previous Studies ..... 25

Chapter 3 METHODOLOGY ..... 27

*3.1 Equating Design Based on AEPS Structure Description ..... 27*

3.1.1 Structure Similarity ..... 27

3.1.2 Structure Changes between AEPS-2 and AEPS-3 ..... 29

*3.2 Data and Item Parameter Calibration ..... 31*

3.2.1 Data ..... 31

3.2.2 Item Calibration Procedure ..... 32

3.2.2.1 Rasch Model as the method of analyses ..... 32
3.2.2.2 The Brief History of the Application of Rasch Model in the Early Childhood Education Assessment Field ..... 33
3.2.2.3 Rasch Model with Category Data ..... 33
3.2.2.4 The Operational Definition for Each Domain ..... 36
3.2.3 Scale Validation ..... 40
3.2.3.1 Unidimensionality ..... 40
3.2.3.2 Data-model Fit ..... 41
3.2.3.3 Separation & Reliability ..... 43

*3.3 Equating Procedure ..... 44*
3.3.1 Selection of Anchor Items ..... 44
3.3.2 Equating Method: Fixed Parameter Calibration (Anchoring) ..... 46
3.3.2.1 The Process of Converting True Score ..... 47
3.3.2.2 The process of converting observed score ..... 47
3.3.3 Evaluation of the Equating ..... 48
3.3.3.1 General rule of evaluation in the equating ..... 48
3.3.3.2 Evaluation in this study ..... 49

*3.4 Chapter Summary ..... 50*

Chapter 4 RESULTS ..... 52

*4.1 Data Sample ..... 52*
4.1.1 Descriptive overview of AEPS-2 test level I Sample ..... 53
4.1.2 Descriptive Overview of AEPS-3 Test ..... 54
4.1.3 Descriptive Statistics of the Data Related to the Anchor Design ..... 55

*4.2 Result of Scale Calibration ..... 58*
4.2.1 Dimensionality ..... 58
4.2.2 Wright map ..... 60
4.2.3 Data fit ..... 61
4.2.4 Separation and Reliability ..... 64

*4.3 Equating ..... 66*
4.3.1 Common Item (anchor) Selection: Identical Matching ..... 66
4.3.2 Common Item (anchor) Selection: Functional Matching ..... 68
4.3.3 Item Difficult Estimation Comparison before and after Anchoring ..... 70
4.3.4 Conversion Relationship between AEPS-2 and AEPS-3 ..... 72

*4.4 Chapter Summary ..... 81*

Chapter 5 DISCUSSION AND CONCLUSION ..... 82

*5.1 Discussion ..... 82*
5.1.1 Discussion of Research Question One ..... 82
5.1.2 Discussion of Research Question Two ..... 84
5.1.2.1 Anchor item selection issue in identical matching ..... 84
5.1.2.2 Reason to Conduct Functional Matching Method ..... 85
5.1.2.3 Item parameter drift in the anchoring ..... 86
5.1.2.4 Model Data Misfit & Limited Anchor Item Pool ..... 87
5.1.3 Discussion of Research Question Three ..... 88
5.1.4 Discussion of Research Question Four ..... 90

5.1.4.1 Conversion table comparison ..... 90
5.1.4.2 Impact of the Coding Scheme and Missing Data on Equating ..... 92
5.1.4.3 Value of functional matching anchoring method in implication ..... 93

*5.2 Contribution and Implication ..... 94*

*5.3 Limitations & Future Research ..... 96*

APPENDICE ..... 98

*APPENDIX 1. WRIGHT MAP IN THE SIX DEVELOPMENTAL AREAS IN AEPS-2 ..... 98*

*APPENDIX 2. WRIGHT MAP IN THE SIX DEVELOPMENTAL AREAS IN AEPS-3 ..... 104*

*APPENDIX 3. LIST OF ITEMS IN AEPS-2 ORDER BY ITEM DIFFICULTY PARAMETER ..... 110*

*APPENDIX 4. LIST OF ITEMS IN AEPS-3 ORDER BY THE ITEM DIFFICULTY PARAMETER ..... 116*

*APPENDIX 5. ITEM CALIBRATION AND FIT RESULTS BY DEVELOPMENTAL AREA FOR THE AEPS-2 ..... 124*

*APPENDIX 6. ITEM CALIBRATION AND FIT RESULTS BY DEVELOPMENTAL AREA FOR THE AEPS-3 ..... 133*

*APPENDIX 7. PERSON ABILITY MEASURE CONVERSION TABLE IN EACH AREA ..... 143*

**REFERENCES** ..... 156

**VITA** ..... 162

LIST OF TABLES

Table 1 degree of similarity between eligibility decisions and norm-referenced assessments ..... 23
Table 2 Psychometric Properties of AEPS under Rasch Framework ..... 26
Table 3 Number of AEPS Test Items per Developmental Area ..... 30
Table 4 The content of six developmental areas ..... 36
Table 5 Descriptive Statistics of Raw Scores for Both Versions: Total Items and Common Items in Six Developmental Areas ..... 57
Table 6 Dimensionality of AEPS-2 and AEPS-3 ..... 59
Table 7 Summary of item infit and outfit indices for AEPS-2 and AEPS-3 ..... 63
Table 8 Summary of separation and reliability of AEPS-2 ..... 65
Table 9 Summary of separation and reliability of AEPS-3 ..... 65
Table 10 Identical matching Anchor items were used in Fixed calibration equating (anchoring) ..... 66
Table 11 Functional matching Anchor items were used in Fixed calibration equating (anchoring) ..... 69
Table 12 Recommended Convert Table in Six Developmental Areas for Implementation ..... 76

LIST OF FIGURES

Figure 1: Hierarchical arrangement and organizational structure of AEPS test ..... 28
Figure 2 The Histogram of the Age Distribution for Level I Cases ..... 53
Figure 3 The histogram of the age distribution for AEPS-3 cases ..... 54
Figure 4. the comparison of item difficult measures (identical matching anchoring vs. functional matching anchoring) ..... 72
Figure 5 The comparison of person ability measures (identical matching anchoring vs. functional matching anchoring) ..... 74

# CHAPTER 1 INTRODUCTION

## 1.1 Background

Children with physical and mental disabilities long for high-quality education in their lives. Assessment is a critical aspect of high-quality education that requires attention from policy makers, educators, researchers, and parents. Understanding major policies and legislation will offer insight into the history, current state, and future progression of child assessment. When discussing assessment in Early Childhood Education (ECE) and Early Childhood Special Education (ECSE), it is crucial to explore the history of ECSE regarding educational equity and quality, particularly the policies and legislation that impact services and assessments for young children with disabilities. In the ECSE field, it is necessary to assess child development from the very beginning, consistently, and with high-quality assessment tools.

It is important to understand the brief development of assessment in the ECE and ECSE fields, particularly the Assessment, Evaluation, and Programming System (AEPS). Established in 1974, the AEPS is an integral part of the assessment development history in ECSE in the United States and has been influenced by educational changes. A comprehensive understanding of ECSE policies and legislation, especially the development and history of assessment, will set the context for studying score equating between two versions of AEPS. Therefore, this section focuses on understanding the brief development of assessment in the history of ECSE through policy and legislation, starting in the 1960s.

## 1.2 Theoretical Preparation and First Assessment

In the 1960s, with the emergence of far-reaching implications for future child development (e.g., Piaget's theory of cognitive development and Vygotsky's concept of the zone of proximal development), the "nature vs. nurture" debate in child development gradually shifted from "genetic determinism" (emphasizing the role of genes or elements of human physiology) to "the theory of gene-environment interaction" (emphasizing the role of gene-environment interaction). These changes provided the theoretical basis for ECE policies for children with disabilities and transformed the view of child development in laws and legislation.

Beginning in the 1960s, the federal government and many universities started to expand their role in early childhood education. In 1965 (revised in 1966), the Elementary and Secondary Education Act (ESEA), referred to as Public Law 89-10, was enacted as the first broad-scale education act, which significantly impacted early childhood special education. The ESEA stipulated that states and localities could use federal funds to provide funding for all children (including infants and toddlers). The amendment explicitly stated that the bill should protect children with disabilities and set the exact amount of funding, which guaranteed the bill's administrative implementation and provided full financial support (e.g., 1.3 billion). However, despite this, the first program evaluation regarding services for children with disabilities did not receive attention. Some members from the Cooke Head Start Planning Committee voiced their disagreement with special education evaluation: "The medical people felt that the purpose of Head Start was to feed children, get their teeth fixed, and offer them a pleasant experience. What was there to evaluate? It was clear Head Start would do no harm." (Vinovskis, 2008) In 1965, after applying the program assessment, Zigler said the measure was "a disaster" (2008). In summary, during this period, special education received considerable financial support at the

policy level, but the evaluation aspect was not prioritized in terms of both awareness and execution.

### 1.2.1 The Role of Assessment in ECSE

During the late 1960s and 1970s, assessments began to play a crucial role in ECE for the first time. In 1968, the Handicapped Children's Early Education Assistance Act (Public Law 90-538) was issued to find suitable and meaningful education approaches for children aged 0-6 and provide relevant information and guidance for education programs. This act established the Handicapped Children's Early Education Program (HCEEP), the first federal ECE program focused on serving the entire population of young children with disabilities (Hebbeler, Smith, & Black, 1991). Assessment became an essential component of ECSE programs. The federal government funded 120 experimental centers as part of the First Chance Program and their demonstration intervention models. One main goal of this program was to develop, test, and publicize assessments for young children with developmental disabilities or those at high risk. For example, a comprehensive early childhood project in Cedar Rapids and a model preschool center for disabled children with professional training, research, and service component & mental retardation center in Seattle were funded to develop practical assessments (Black, 1974). Later, in 1972, the Economic Opportunity Act Amendment was introduced, guaranteeing that ten percent of Head Start enrollment opportunities were reserved for children with disabilities. However, despite these policies aimed at ensuring inclusivity, assessment requirements still varied among states during that time. This lack of consistency caused significant delays in the development of assessments for early intervention (Sandall, McLean, & Smith, 2000, p. 32).

From the 1980s onwards, various amendments and laws were passed to expand provisions and support for children with disabilities. The Amendment to the Education for All

Handicapped Children Act of 1983, the Education of the Handicapped Act Amendments (EHAA, Public Law 99-457) in 1986, the Americans with Disabilities Act (ADA) in 1990, and the Individuals with Disabilities Education Act (IDEA) all aimed to provide better education and related services to disabled children. These Acts broadened the target population for services (e.g., federal funds under Preschool Award Grants served all children with disabilities), the age range (e.g., EAHCA included 0-3-year-olds), and the enforcement (e.g., EHAA mandated implementation in the preschool component) (Hebbeler et al., 1991). Assessments became an essential part of early education, ensuring eligibility and effectiveness of intervention programs for specific sub-groups during 1980s (Bailey & Bricker, 1986). Despite legal protections, people with disabilities continue to face discrimination in various aspects of life. These laws and amendments have been essential in promoting early identification measures and emphasizing the importance of assessment in early education for children with disabilities.

Into the early 2000s, more than a decade later, legislation surrounding assessment in ECE/ECSE remained largely unchanged. However, in the first decade of the 21st century, transformative acts such as the No Child Left Behind Act (NCLB) and the Race to the Top Plan (RtT) emerged, significantly underscoring the importance of assessment in ECSE. NCLB emphasized the importance of evidence-based practice (EBP) in education. Assessment is an important part in EBP, focusing on children's performance monitoring and data-based decision-making (Reichow, Boyd, Barton, & Odom, 2016, p. 16). During this period, early intervention assessment shifted towards developmental appropriateness and family involvement (Sandall, McLean, & Smith, 2000, pp. 33-34). The recommended practice of early childhood assessment from DEC also indicates that the assessment should be helpful, acceptable, authentic, collaborative, convergent, equitable, sensitive, and harmonious (Bagnato & Neisworth, 1999). In

2009, the RtT Plan aimed to improve education quality across the country by offering funding through competitive grants. As a result, attention was given to early childhood special education assessment, leading to significant federal investments (e.g., The Enhanced Assessment Grants refer to a $15 million fund targeting kindergarten entry assessments) in the development of psychometrically sound instruments. Examples of these initiatives include the Quality Rating and Improvement System (Schachter, Piasta, & Justice, 2020) and the Race to the Top-Early Learning Challenge grants, which helped states develop and enhance comprehensive early childhood assessment systems. The National Research Council (2008) defined a complete early childhood assessment system (CECAS) as an integrated assessment system that includes developmental screening measures, formative assessments, environmental quality measures, adult-child interactions, and a kindergarten entry assessment. These efforts facilitated the monitoring of young children's learning and development and the evaluation of the effectiveness of early childhood learning programs.

Over the past decade, the U.S. Congress promulgated Every Student Succeeds Act (ESSA), which paid more attention to the quality of education and the educational equity of disadvantaged students and emphasized the importance of early intervention. A preschool development fund has been established to encourage states to promote the coordination and quality of early intervention services and access early intervention opportunities for infants and young children with special needs. The bill also emphasizes continuous and comprehensive evaluation to monitor the quality of early intervention (Dennis, 2017). Simultaneously, at the statewide level, the assessment has started to pursue the needs of systemization and standardization; for example, the Virginia Department of Education issued the quality standards of inclusive school self-assessment (Education, 2019). This need is not limited to the USA;

Australia and Canada are developing national early childhood assessments. Instrumentation is needed to monitor young children's growth globally; however, young children's assessment is fraught with challenges. Psychometricians and educational researchers must work together with the early childhood community to develop these instruments.

## 1.3 History of AEPS

Educational policies and regulations have influenced the development of the Assessment, Evaluation, and Programming System (AEPS), which was designed in the late 1970s by Diane Bricker, Ph.D., and her colleagues as a comprehensive tool to assess and support the development of young children. Originally, its purpose was to evaluate infants, toddlers, and preschoolers who were at risk of developmental delays or had disabilities. The first edition of AEPS was published in 1984 and it has been revised and updated four times until 2021. This section presents the four stages of AEPS development, from the initial preparation to the publication of the first edition-API & EPS, the release of the second edition (AEPS-2), and the latest and current edition, AEPS-3.

### 1.3.1 Preparation and Meeting

The delay in the development of assessments for early intervention (Sandall et al., 2000, p. 32) motivated professionals to work on long-term child assessment development. The original idea of AEPS came from a meeting of the American Association for the Education of the Severely and Profoundly Handicapped (now known as The Association for Persons with Severe Handicaps) in October 1974. A group of professionals discussed providing a functional measurement tool for children with special needs. This discussion was fascinating because everyone working with young children was eager to find an alternative to standard normative

reference tests or homemade assessments, the validity and reliability of which were questionable at the time.

### 1.3.2 Early Versions of Instrument – API & EPS

In 1980, the University of Idaho supported the project through a supplemental award to the Handicapped Children's Early Education Project grant. (By this time, Gentry had moved to Idaho, and Bricker to Oregon.) Dale Gentry, along with Katie McCarton, created the Adaptive Performance Instrument (API). As an extensive collection of more than 600 items, API provided detailed and valuable descriptions of children's behavior from birth to two years old. However, a long administration time (8 to 10 hours) frustrated implementer. The instrument's operational issues, the termination of additional federal subsidies, the lack of psychological test data, and persistent time management problems led to API becoming part of the final project report.

From 1983 to 1984, significant modifications (i.e., rewriting the items) led to the measure being renamed as the Comprehensive Early Evaluation and Programming System. Two changes were made during this work period. First, the number of items was reduced from over 600 to fewer than 300; second, the assessment extended the target developmental range to 36 months.

### 1.3.3 AEPS-2

The EPS, serving as the blueprint for AEPS, evolved over a decade into the commercialized product known as 'AEPS for Birth to Three Years.' Subsequently, the second version of AEPS underwent another ten-year development phase, ensuring comprehensive refinement. In 1993, Paul H. Brookes Publishing Company commercialized the AEPS Test for Birth to Three Years and its associated curriculum. Like the first edition, the AEPS for Birth to Three Years was composed of a test and an associated curriculum. Later, due to the urgent need

for an adequate assessment tool for children aged three to six years, the development of AEPS for this age group became a crucial plan to cover. Since the test aimed to focus on the entire range of early childhood, the development of the test for three-to-six-year-olds and the related curriculum began in 1985. By 1987, Slentz conducted the first field study for the first version (1987). The results of this study laid the foundation for extensive revisions of the test.

During the same period of time, between 1992 and 1995, the researchers developed a curriculum linked to Level II of the AEPS. The researchers began calling the test Assessment, Evaluation, and Programming System Test for Three to Six Years. In 1996, Brookes Publishing Company added Volumes 3 and 4: AEPS Assessment for Three to Six Years (Bricker & Pretti-Frontczak, 1996), and AEPS Curriculum for Three to Six Years (Bricker & Waddell, 1996), respectively.In 1999, the researchers' group studied and discussed the data gathered from the AEPS® test and the outreach training projects. Based on these discussions, the participants (Kristie Pretti-Frontczak, KJ Slentz, Elizabeth Straka, Betty Capt, Jane Squires, Natalya McComas, and Diane Bricker) modified the test. After one year of discussion and revision, in the fall of 2000, the team completed the second edition, which was released in 2002.

The second edition of AEPS® contains four parts: Administration Guide, Test for Birth to Three and Three to Six Years, Curriculum for Birth to Three, and Curriculum for Three to Six Years. Additionally, some components of the AEPS have been disseminated to other countries in Spanish, French, Chinese, and Korean. The release of the second edition expanded training efforts. The authors of AEPS and other AEPS experts provided a series of training opportunities for domestic and international users.

### 1.3.4 Online System AEPSi & AEPS-3

In the fall of 2006, Brookes Publishing initiated an online electronic management system for AEPS® called AEPSinteractive (AEPSi). AEPSi is a secure, online system for individualized child data files that enhances the efficiency and effectiveness of AEPS®. The features of AEPSi include automatic scoring, monitoring developmental progress, and creating personalized reports to meet various levels of standards and requirements (i.e., local, state, and federal levels). The AEPSi was continuously improved and used throughout the following decade. Systematic data collection provides enhanced data resources for the study.

Additionally, the AEPS® authors have established a non-profit company: Early Intervention Management and Research Group (EMRG). EMRG is committed to supervising the AEPS®'s continuous commitment to quality to ensure its ongoing improvement. After issuing the second edition, researchers began to put more effort into modifications and improving the validity of the AEPS test. Following nearly two decades of diligent efforts and numerous evidence-based studies (Grisham, Waddell, Crawford, & Toland, 2020; Johnson & Macey, 2019; M. Macy, Chen, & Macy, 2019; M. Macy, Pool, Chen, Rusiana, & Sawyer, 2021; Toland, Grisham, Waddell, Crawford, & Dueber, 2021) that carefully examined the content, psychometric properties, and cut-off scores, the AEPS-3 was released in December 2021.

### 1.3.5 The Needs of Conversion Table between AEPS-2 and AEPS-3

In early childhood education, the continual development of child developmental score documentation is an urgent need and a vital component of assessment. The development of a conversion table for child development is crucial for tracking progress and comparing scores across different versions of an assessment tool. For example, the history of AEPS development

demonstrates the importance of a conversion table, as the second version includes a conversion table that converts raw scores from the first and second editions to percentages.

Despite the importance of conversion tables, they have received less attention than they deserve. Many assessment tools lack awareness of the need for score equating during the transition between different versions, making it challenging to find score exchange tables for commonly used child development assessments like the Battelle Developmental Inventory and the Ages and Stages Questionnaires.

Further research is necessary to identify best practices for score equating across different versions of child development assessments. While conversion tables exist to varying degrees, such as the simple raw score percentage method employed in AEPS-2 and the z-score percentile method used in the Ages and Stages Questionnaires, scholars have yet to explore score equating under the Rasch framework in the early childhood assessment scenario. Ensuring that conversion tables are available and accessible can help guarantee that child development assessments are effective and reliable tools for tracking progress and promoting healthy development in early childhood education.

### 1.4 Rational for Score Equating

In understanding the significance of accountability in early childhood education, let's delve deeper into how various accountability measures can enhance educational quality and effectiveness, contribute to policy development, and better prepare children for their academic journey ahead.

Accountability in early childhood education is essential for ensuring quality and effectiveness, helping bridge readiness gaps for kindergarten. This hinges on accountability measures that maintain high teaching standards and learning goal achievement (Wright, 2007). Key to this are assessments, both formal and informal, tracking children's progress. Although challenging due to children's uneven development, assessments need to be culturally, linguistically, and developmentally responsive. Authentic assessments consider factors such as language comfort and familiar settings, while standardized ones meet reliability and validity standards. Efforts to enhance program quality involve methods like direct regulations, setting minimum standards such as class sizes, teacher qualifications, and safety requirements. These often create a quality floor, not encouraging improvement. Outcomes-based approaches like QRIS focus on outcomes, improving quality, though they sometimes narrow educational goals. Assessment tools must be valid, reliable, developmentally appropriate, and culturally sensitive, accommodating varied skill acquisition rates, multiple languages, and dual-language learners, ensuring relevance for a diverse child population.

In conclusion, accountability entails effective, appropriate assessment tools, aligned with K-12 systems. Culturally sensitive, valid, and reliable assessments alongside quality improvement form an effective accountability system, crucial for future academic success and lifelong learning.

Federal usage of assessment scores significantly influences funding allocation, policy formulation, and intervention programs. Consistent scores across assessments foster educational protocols development, reflecting real student performance changes.

Recently, U.S. Departments of Education and Health and Human Services used data to improve young children’s social-emotional development and mental health, emphasizing data-

driven policy. International comparisons, like the OECD's 'Education at a Glance 2022' report (Indicators, 2023), inform policy discussions, showing lower early education program participation in the U.S. compared to the OECD average. The 'Early Learning in the United States: 2021' report (Cascio, 2021) highlights challenges like early learning affordability, accessibility, and early educator compensation. Assessment scores gauge proposed solution success, such as public investments in childcare, aimed at reducing educational disparities and enhancing economic stability.

In conclusion, federal score usage influences policy, funding allocation, and intervention program design. Domestic and international score data guide decisions affecting children, families, and educators, emphasizing consistency and equating in scoring.

Scoring consistency and equating standardize procedures across evaluators and adjust scores across versions, ensuring score comparability and accurate child development understanding. Equating allows comparison of student performance over time or across groups. Training assessors, data reviews, and standardized protocols improve consistency, while equating aids in interpreting score changes, comparing cohorts, and evaluating interventions. These methods identify developmental trends and improvement areas, helping educators tailor teaching approaches. Assessing children requires consistent, equatable methods sensitive to child development, culture, language proficiency, and personal traits, and assessments should guide teaching practices, curriculum planning, and program evaluations, allowing educators to refine strategies based on children's learning progress (Grisham-Brown, Hallam, & Pretti-Frontczak, 2008; Grisham, Waddell, Crawford, & Toland, 2020).

## 1.5 Research Questions

The rise in educational equity and the introduction of AEPS-3 have sparked new concerns regarding the fairness of transitioning child development measures from AEPS-2 to AEPS-3. The aim of this study is to provide psychometric evidence for score equating between AEPS-2 and AEPS-3. First, a psychometric property examination was conducted in terms of the developmental areas of AEPS-2 and AEPS-3 using a Rasch model. Rasch analyses were chosen because they provide a psychometric method for assessing items within a measure and ensuring items differentiate children at various points along the continuum (Bond, Fox, & Lacey, 2007; Boone, 2016). Second, this study aims to use two anchor item designs—identical anchor items and functional matching anchor items (refer to section 3.3.1 for definitions)—to analyze the impact of the anchor item on equating results. Subsequently, the score equating procedure will be executed to determine the relationship between the six developmental areas in AEPS-2 and AEPS-3. The research questions for this study are as follows:

1. To what extent do AEPS-2 and AEPS-3 instruments fit the Rasch Rating Scale Model?

2. What is the most efficient set of the common items in the six developmental areas, respectively, for the purpose of equating across two measures?

3. How adequate was the fixed parameter calibration, in terms of the accuracy of equating?

4. What score conversion table is provided on the six developmental areas from AEPS-2 to AEPS -3?

## 1.6 Significance

By addressing several questions, this study aims to contribute valuable insights to existing literature. The first insight offered by this study concerns score equating for comparability in developmental areas. Assessing children using different scales may pose challenges due to potential variations in the difficulty levels of the scales. Score equating is a statistical process employed to adjust scores on assessment forms, enabling the interchangeability of scores from two different versions. By conducting score equating, this study ensures that scores are comparable within each developmental area, allowing for more accurate comparisons and interpretations of children's progress.

The second insight offered by this study is the importance of generalizing the system between two versions. Children's development is characterized by continuity and instability, necessitating the collection of extensive long-term data for reliable evaluations of their abilities. Score equating enables the generalization of the assessment system, preserving children's longitudinal developmental information across different versions of the assessment tool. This allows researchers and educators to track children's development consistently, despite potential changes in assessment versions.

The third insight presented by this study is ensuring fairness in decision-making. AEPS is an evidence-based assessment tool, with scores playing a crucial role in determining eligibility for interventions. As AEPS-3 is introduced, there is a transition period between the new and old versions. Score equating ensures fairness in decision-making when children's developmental scores are derived from different versions of the assessment. By eliminating practice effects, score equating helps maintain equitable decision-making processes, such as determining

eligibility for special education services, even when children of the same age have scores from different AEPS versions. Fourth. Preequating for Raw-to-Scale Score Conversion in Future Versions: This study employs the Rasch model equating, a method that differs from classical test theory. As all items in AEPS-2 and AEPS-3 are administered in this study, they remain invariant when applied to new groups. This process allows for the preparation of raw-to-scale score conversion tables before the future "new" form is issued, facilitating rapid score reporting. By establishing these conversion tables, the study lays the groundwork for future versions of the assessment tool, ensuring that scores can be quickly and accurately converted and reported.

In summary, this study contributes to the existing literature by addressing the challenges associated with the introduction of AEPS-3 and the need for score equating. It ensures comparability in developmental areas, generalizes the system between different versions, guarantees fairness in decision-making processes, and paves the way for seamless raw-to-scale score conversions in future versions. By tackling these questions, the study aims to enhance the accuracy, consistency, and fairness of assessments for children with disabilities, ultimately leading to better support and outcomes for this population.

### 1.7 Organization of Dissertation

In this dissertation, the information is organized into five chapters, which include the introduction, literature review, methodology, results, and discussion. This chapter focused on the importance of assessment in the areas of ECE and ECSE. It provided a theoretical foundation and examined the historical context of assessment in these areas, supported by evidence from policies and regulations. Additionally, the chapter reviewed the development of the AEPS as a representative child developmental assessment tool in the field. This example highlighted the

need for score exchange between different versions of the assessment and emphasized the significance of score equating in enhancing assessment utility and application. Besides the introduction chapter, my dissertation consists of another four chapters.

Chapter two focuses on reviewing the purposes of the AEPS as a curriculum-based assessment tool. It provides a brief summary of the psychometric properties of AEPS under classical testing theory, specifically examining reliability, validity, and utility. Additionally, chapter two reviews the literature on the psychometric properties of AEPS under the Rasch framework, and the closing of this chapter highlights the limitations found within the previous literature on the topic.

Chapter three presents the research plan to solve each question, in which I covered several key aspects. Firstly, I discussed the equating design, which was based on examining the structural similarities and differences between the two assessments. Secondly, I outlined the process of parameter calibration under the Rasch framework. Additionally, I described the common item selection, scale transformation method, calibration linking, and evaluation strategies employed during the equating process.

Chapter four first provides an overview of the data sample for both the AEPS-2 and AEPS-3 tests. It then focuses on the dimensionality of the AEPS-2 and AEPS-3, addressing research question 1. Sections three and four analyze the item structure of both AEPS tests, specifically addressing research question 2.a. The fourth section presents the results of the scale analysis, pertaining to research question 2. Finally, the findings of the scaling equating investigation, addressing research question 3, are discussed. Chapter five provides the discussion, conclusions, contributions and implications, limitations, and suggestions for future studies.

# CHAPTER 2 LITERATURE REVIEW

A substantial body of literature has extensively explored the psychometric characteristics of AEPS, encompassing studies conducted within the traditional testing theory as well as those employing the Rasch model framework. However, upon reviewing the available literature, it becomes evident that no direct investigations have been conducted specifically examining the score exchange between AEPS-2 and AEPS-3. As a result, the primary focus of this literature review will be on the current research pertaining to the psychometric characteristics of the assessment.

## 2.1 Curriculum-based Assessment

Bagnato and Neisworth (2000) suggested that standard early childhood assessment practices should encompass eight essential characteristics: functional, acceptable, authentic, collaborative, harmoniously convergent, sensitive, and equitable. Numerous assessments for ECSE have adopted the author's practice recommendations. Curriculum-based assessment (CBA) is one of the most practical and effective approaches, embodying a developmental orientation (Bagnato, Neisworth, & Capone, 1986). Renowned examples of CBAs include the High/Scope Child Observation Record (COR), Hawaii Early Learning Profile (HELP), SCERTS Model, Carolina Curricula, Creative Curriculum, and the AEPS Test (Gao & Grisham-Brown, 2011).

The AEPS is a CBA designed to align closely with the curriculum, accurately reflecting the content being taught (Grisham-Brown & Pretti-Frontczak, 2011). The AEPS evaluates the mastery of content within a logical hierarchy (Vanderheyden, 2005) and combines the curriculum goals and assessment questions using the same set of items. For instance, the AEPS

test's identification system features a hierarchical structure consisting of strands, goals, and objectives (refer to Figure 1 in Section 3.1.1, Page 25). This organizational structure signifies the sequential arrangement of the strands, goals, and objectives. Furthermore, the same set of items serves as both the curriculum goal and the assessment question.

## 2.2 The Purposes of AEPS

The primary usage of the assessment results is planning for instruction, reporting children's developmental progress, and continuously evaluating the program's quality. According to the principle of designing a qualified assessment by NAEYC, "assessment of young children's progress and achievements is ongoing, strategic, and purposeful." The assessment results are used to inform the planning and implementation of experiences, communicate with the child's family, and evaluate and improve teachers' and the program's effectiveness."

The AEPS assessment serves four primary purposes, which are utilized on various occasions: 1) screening for design intervention content, 2) eligibility determination, 3). accountability monitoring 4) and program assessment.

Specifically, for the purpose of screening design intervention content, the AEPS test is designed to measure six critical developmental areas: gross motor skills, fine motor skills, cognitive abilities, adaptive skills, social skills, and social communication. In line with the hierarchical structure of child development, the test examined diverse learning tasks, the number of items, and the scope of content areas. The AEPS test follows the standard of a good assessment as the DEC recommended, and the children's scores on the AEPS test show similar patterns as their scores on other high-quality assessment instruments. When the educators plan and monitor the intervention content, AEPS can be a potential supplement or replacement.

The second purpose of the test is to determine eligibility. Some researchers consider the AEPS one of the most effective and efficient assessments to establish eligibility (Grisham-Brown et al., 2008). AEPS offers authentic assessment, such as collecting data from observing children's behavior in the natural environment. Professionals use the assessment results to verify eligibility for services and determine high-quality goals and intervention content (Bricker et al., 2008; Lee, Bagnato, & Frontczak, 2015).

The third purpose is accountability monitoring. In 2005, OSEP started the Results Driven Accountability process to assess the quality of ECE programs. OSEP's accountability system shifted from compliance-oriented to results-oriented. All federal early childhood agencies are responsible for reporting child outcome data, and the evaluation of the child outcomes determines the distribution of funds from Part B and Part C of Section 619. Through engaging the stakeholders, three targeted child outcomes were identified in the State Child Outcomes Measurement System Framework (S_COMS) (Early Childhood Technical Assistance Center & Center for IDEA Early Childhood Data Systems, 2017). The program should report data on children's positive social-emotional skills (including social relationships). The program also needs to measure the quality of children acquiring and using knowledge and skills (including early language and communication), and the part C preschool requires an early literacy area. The last quality requirement is that children use appropriate behaviors to meet their needs.

The AEPS's social and social-emotional areas encompass all the requirements of the positive social-emotional skills as per S_COMS. Also, the six developmental areas in the AEPS-2 and eight developmental areas in the AEPS-3 cross coordinate the second and third targets of child's outcomes: acquire and use knowledge and skills; use appropriate behavior to meet their

needs. Also, the professionals proved that the result of the AEPS test meets the federal accountability requirements (2015).

The fourth purpose is program assessment. Many school psychologists and experts, when deciding on early childhood assessment tools, are more familiar with the normative reference test than with the curriculum-based authentic assessment, such as AEPS, for program assessment. The agreement between AEPS and norm-referenced tests should assure the validity of using an observational evaluation to plan and monitor effectiveness.

## 2.3 Psychometric Properties of AEPS

Assessments are tailored to a specific purpose and used only for the purpose demonstrated to produce reliable, valid information. This review of psychometric properties of AEPS is conducted under the Classical Test Theory (CTT) framework and Rasch framework. Below, I will review the AEPS assessment reliability, validity, and utility based on the existing studies (see table 2) under CTT framework first.

### 2.3.1 Reliability

Scale reliability is a measure of the consistency and stability of a test or assessment tool. It refers to the extent to which the scores obtained from the scale are consistent across multiple administrations of the test or when different items from the same construct are used. A highly reliable scale produces similar results under consistent conditions, indicating that the assessment tool effectively measures the intended construct with minimal error. Under the classical testing theory, several common methods, such as test-retesting, internal consistency, and split-half reliability, are conducted to measure the scale's reliability. Correlation and Cronbach's alpha are

measures produced by the methods, and a higher value of these measures indicates more consistent results. The reliability of a scale is often measured using the person separation reliability index (PSR) or the person separation index (PSI) within the Rasch model framework. These two parameters assess internal consistency reliability, similar to classical test theory, and evaluate the ability to distinguish between different person ability levels. Both PSR and PSI range from 0 to 1, with higher values indicating better reliability and greater precision in differentiating among individuals.

Studies of the various versions of AEPS have consistently demonstrated good psychometric properties. Utilizing classical test theory, two studies (Gao & Grisham-Brown, 2011; Grisham-Brown et al., 2008) indicated the AEPS-2 is a reliable measure of child development. At the same time, other efforts have resulted in the construction of AEPS-2, which aims to provide a precise measurement on the individual level and predict various outcomes in child performance ((Bricker et al., 2008; Bricker, Yovanoff, Capt, & Allen, 2003; Castaneda-Villa & James, 2007; Gao & Grisham-Brown, 2011; Wang, Sandall, Davis, & Thomas, 2011).

### 2.3.2 Validity

Validity refers to "an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores or other modes of assessment" (Kane, 2010).

Six studies have investigated the degree of similarity between eligibility decisions and norm-referenced assessments. Bricker et al. (2003) discovered that the chance of AEPS over-identifying children's eligibility ranges from 5% to 25% by age interval, while the risk of under-identification is between 0% and 8%. Macy, Bricker, and Squires (2005) found that, when

compared to norm- referenced assessments like Gesell or BDI, the average agreement rate in eligibility is as high as 94%. Bricker et al. (2008) findings align with Bricker et al. (2003) findings. In a study involving children aged 4 to 66 months, Bricker et al. (2008) found that the probability of AEPS over-identifying children's eligibility is between 9% and 30% by age interval, while the likelihood of under-identifying ranges from 0% to 12%. Toland et al.'s (2021) study found that the eligibility classification accuracies were consistent with Bricker et al. (2008). Additionally, the precisions of eligibility classification were compatible with Bricker et al. (2008). It is important to emphasize that one of the providers' preferences is to maximize the sensitivity of eligibility determination to ensure that no child in need of services is overlooked.

AEPS cut-off score in Hallam, Lyons, Pretti-Frontczak, and Grisham-Brown (2014) demonstrated similar eligibility determination decisions as observed in previous studies. The researchers compared AEPS cut-off scores with BDI (Newborg & Company, 2005) standard deviation scores to examine the consistency of decisions from each test. The results indicated a 78% agreement in the developmental status of the children assessed.

Table 1 degree of similarity between eligibility decisions and norm-referenced assessments

| Author (year) | Version /Participants | Content | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Concurrent validity | Construct validity | Internal consistency | Inter-rater agreement | Inter-rater reliability | Rest-retest reliability | social validity | utility |
| **Bailey & Bricker (1986)** | EPS-I; 32 (10 w/ disabilities) | ✓ | | | | ✓ | ✓ | | ✓ |
| **Slentz (1986)** | EPS-II; 56 (15 w/disabilities, 22 at-risk) | | ✓ | ✓ | | ✓ | | | |
| **Bricker, Bailey, & Slentz (1990)** | EPS-I 335 (152 w/disabilities, 93 at-risk) | ✓ | | ✓ | | ✓ | ✓ | | ✓ |
| **Hsia (1993)** | EPS-II; 82 (disability status not Specified) | | ✓ | ✓ | | ✓ | | | |
| **Sher (2000)** | AEPS, 1st Ed., Level I; 20 (10 w/disabilities) | | ✓ | | | ✓ | | | |
| **Noh (2005)** | AEPS, 2nd Ed., Level II; 65 (31 w/disabilities) | ✓ | ✓ | ✓ | | ✓ | | ✓ | |
| **Gao& Grisham-Brown (2011)** | AEPS, 2nd Ed., Level II; (children w/o disabilities); 5 preschool teachers | ✓ | | | | ✓ | | | ✓ |
| **Wang, Sandall, Davis, & Thomas, 2011** | | | | | | | | | |

### 2.3.3 Utility

In a series of utility studies on various AEPS versions, participants found AEPS testing valuable for program planning, monitoring, and setting beneficial developmental goals for children. Several assessment utility studies reveal that teachers and daycare providers believe the AEPS test (formerly called EPS) aids in developing educational plans (Bailey & Bricker, 1986) and contributes to high-quality learning goals (Bricker & Pretti-Frontczak, 1998; Pretti-Frontczak & Bricker, 2000). A comparative study between AEPS and the Oregon Project (Hamilton, 1995) found that goals written with AEPS were of higher quality than those from the Oregon Project. In another similar study, Notari and Drinkwater (1991) discovered that AEPS was more effective and convenient for integrating into a child's routines compared to developmental goals generated from individual education plans.

Straka (1996) also observed similar results when comparing developmental goals from the Communication and Symbolic Behavior Scales, with AEPS producing higher quality goals and objectives for young children. In Gao and Grisham-Brown (2011), Head Start teachers preferred AEPS for classroom planning compared to norm-referenced tests. D. Lee, Bagnato, and Pretti-Frontczak (2015) found that AEPS, as an authentic assessment, was superior in monitoring children's progress, eligibility determination, individual program planning, and meeting federal accountability requirements compared to conventional assessments based on professional preference. However, the study also noted that the high level of utility attributed to AEPS by professionals might be due to it being a required assessment rather than the quality of the measure itself. itself.

### 2.3.4 Psychometric Properties of AEPS under Rasch Framework

Bricker's (2003) study first used the dichotomous logistic model under the Rasch framework to explore the technical adequacy of AEPS, which lost the partial credit value information of the data. This operation causes an additional scoring error. For the child development area, a child may be able to demonstrate the difficult item rather than master the primary skill. For example, in the gross motor area, the child can skip the crawling part of the development process and directly master walking skills. As shown in Bricker (2008) and Winchell's (2011) studies, the three-point rating score was used in the Rasch analyses. All three studies I mentioned considered all developmental skills as one whole trait. However, McLean (2005) suggested exploring other methods based on the model-data fit statistics from Bricker's (2003) study. In Toland, Grisham, Waddell, Crawford, and Dueber (2021) study, instead of applying the Rasch model for the total score, the rating scale Rasch model was used in six developmental areas, respectively (for more details, see table 3).

### 2.3.5 Summary and Limitation of Previous Studies

While scholars have conducted extensive studies on the AEPS's psychometric properties under the classical test theory and the Rasch framework, their focus has primarily been on validating the scale and providing ongoing evidence for its reliability, validity, and utility. However, there appears to be a research gap regarding the exploration of score exchange between the two most recent versions of the AEPS. It is necessary to investigate the comparability and equivalence of scores between these versions, which would contribute valuable insights to the field of early childhood assessment.

Table 2 Psychometric Properties of AEPS under Rasch Framework

| Author(year) | Version | Participant | Age | Purpose | Rasch Model | dimensionality | Model fit | Reliability |
|---|---|---|---|---|---|---|---|---|
| **Bricker, D., Yovanoff, P., Capt, B., & Allen, D. (2003).** | AEPS-2 | Level I: 436 children<br>Level II: 425 children | 1-72 mo | corroborate eligibility decisions. | Rasch dichotomous<br>one-parameter logistic ( 1 PL) model | Unidimensionality | No fit evidence provided | No reliability evidence provided |
| **Bricker, D., Clifford, J., Yovanoff, P., Pretti-Frontczak, K., Waddell, M., Allen, D., & Hoselton, R. (2008).** | AEPS-2 | Level I: 732 children<br>Level II: 649 children | 0-66 mo | Eligibility determination | Partial credit model | Unidimensionality | All the item fit the model quite well | All the areas reliability >0.79(except Adaptive) (level I);<br>All the areas reliability >0.67(level II) |
| **Winchell, B. (2011).** | AEPS-2 | Level I:7,162 children<br>Level II: 17,194 children | 0-96 mo | examine the technical adequacy | Rasch model | Unidimensionality | 96.8% item fit the model (level I) ;<br>99.5% item fit the model (level II)<br>Note: 0.5-1.7 | Good reliability: person separation reliability = 0.96; item separation reliability = 0.98(level I)<br>person separation reliability = 0.99;<br>item separation reliability = 1.00(Level II) |
| **Toland, M. D., Grisham, J., Waddell, M., Crawford, R., & Dueber, D. M. (2021).** | Field-test version of AEPS-3 | 874 children | 2–83 mo | (1). Evaluate the AEPS-3<br>(2). Eligibility Determination | Rating scale model | Mulitdimensionality | 80.8% item fit the model<br>Note:0.5-1.5 | Person reliability >7.4 |

# CHAPTER 3 METHODOLOGY

In the methodology section, I will discuss 1) the equating design based on the structure similarities and differences of the two assessments, 2) parameter calibration under the Rasch framework, and 3) the selection of common items, scale transformation methods, calibration linking, and evaluation strategies in the equating process.

## 3.1 Equating Design Based on AEPS Structure Description

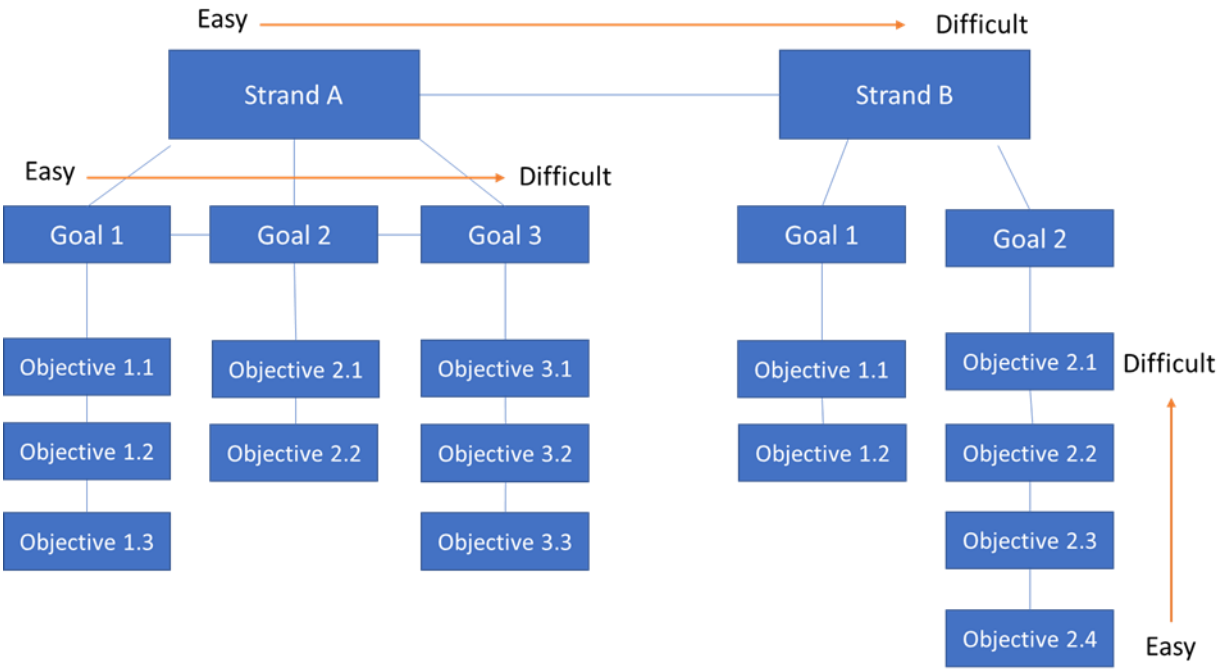
Several approaches can be employed to gather data for equating purposes. The group of test-takers included in an equating study should adequately represent the population that will take the test under normal testing conditions. The Random Groups design, Single Group with Counterbalancing design, and Common-Item Non-equivalent Group design are three popular designs used in score equating. The choice of design takes into account both practical and statistical considerations. In this study, the Common-Item Non-equivalent Group equating design is applied, based on the similarities and differences in the assessment structure, which relate to the data structure. This will be discussed in the subsequent section.

### 3.1.1 Structure Similarity

Both versions of AEPS utilize logical hierarchy structures and a three-point rating scale system. As a curriculum-based assessment (CBA), AEPS is closely linked to the curriculum, effectively reflecting the content being taught (Grisham-Brown & Pretti-Frontczak, 2011). It evaluates the mastery of content within a logical hierarchy (Vanderheyden, 2005) and seamlessly integrates the curriculum goals and assessment

questions using the same set of items. For instance, the AEPS test identification system features a hierarchical structure consisting of strands, goals, and objectives for each developmental area (Figure 1).

Figure 1: Hierarchical arrangement and organizational structure of AEPS test




Note: Adapted from Assessment, evaluation, and programming system (AEPS) for infants and children. Volume 1: AEPS measurement for birth to three years. (Bricker, D. 1994, page.65).

Assessment plays a crucial role in guiding curriculum design, supporting developmental gains across various domains such as language, cognitive, social and emotional, and physical development (NAEYC, 2003). The organizational structure in AEPS-2 and AEPS-3 demonstrates the sequential arrangement of strands, goals, and objectives across six and eight domains, respectively. Each domain comprises a set of strands organized from simple to complex. A group of related items forms the strands,

which represent the goals and objectives arranged hierarchically by the AEPS developers. The data collection method for the AEPS test involves observing a child in their natural environment by a teacher or other trained professionals.

In the AEPS, all items are scored using a three-point rating scale (2, 1, or 0). When a child's behavior consistently meets the expert-defined criterion, the item is scored as 2, indicating that the child can perform the functional skill independently across time, materials, settings, and people. A score of 1 is assigned when the child demonstrates the behavior with assistance, specific people, a particular environment, or inconsistently. The rating scale model will be employed to calibrate the two scales initially, based on the data structure. Additionally, a child receives a score of 0 when they have not yet developed functional skills or cannot be observed exhibiting the target behaviors consistently across time and settings.

### 3.1.2 Structure Changes between AEPS-2 and AEPS-3

Alterations between AEPS-2 and AEPS-3 involve the merging of different age groups and the addition of novel developmental domains. The AEPS-2 includes six developmental content areas for two age levels, with this study focusing on the three to six age level scale (refer to table 3.1). At the birth to three years level, the test consists of 33 items in the fine motor area, 55 items in the gross motor area, 32 items in the adaptive area, 58 items in the cognitive area, 46 items in the social-communication area, and 25 items in the social-emotional area. The AEPS-3 features eight developmental content areas for a single age level (see table 3.1). The fine motor area contains 31 items, the gross motor area has 65 items, the adaptive area comprises 54 items, the social-emotional area includes 61

items, the social-communication area consists of 49 items, the cognitive area has 50 items, the literacy area features 57 items, and the math area contains 41 items.

The AEPS-3 introduces several significant changes. First, it combines the two previous age divisions (birth to 3, and 3 to 6) into a single age range (0-6). Second, AEPS-3 incorporates two new test areas focusing on math and literacy. Third, it includes additional scoring instructions for a 1-point rating. Fourth, the AEPS curriculum has been updated, guided by multi-tiered support systems, blended practices, and activity-based intervention frameworks. These blended practices consist of three levels: growing (birth to 18 months), ready (18 months to 3 years), and skills (3 to 6 years). The curriculum content is associated with 18 routines and activities found in the Ready set volume, ensuring comprehensive coverage of key developmental stages.

Table 3 Number of AEPS Test Items per Developmental Area

| **Version** | **AEPS-2** | **AEPS-3** |
|---|---|---|
| **Scale** | Level I (Birth to Three Years) | Birth to Six |
| **Fine Motor** | 33 | 31 |
| **Gross Motor** | 55 | 65 |
| **Adaptive** | 32 | 54 |
| **Cognitive** | 58 | 61 |
| **Social-Communication** | 46 | 49 |
| **Social** | 25 | 50 |
| **Literacy** | | 57 |
| **Math** | | 41 |
| **Total Number of Items** | 249 | 408 |

In summary, while there were modifications in the age and developmental domains in both versions (such as merging age groups and adding domains), the general structure of the assessment instruments remained similar, and a significant number of items in each developmental domain were retained. However, it is important to note that the data were

not collected from the same group of target population. Therefore, the Common-Item Non-equivalent Group equating design was applied in this study.

## 3.2 Data and Item Parameter Calibration

### 3.2.1 Data

Data collection for the AEPS test involves observing a child in their natural environment by a teacher or another trained professional. In the AEPS, all items are scored on a three-point Likert scale (2, 1, or 0). If the child's behavior consistently meets the criterion described by experts, the item is scored as 2, indicating that the child can perform the functional skill independently across different contexts. A score of 1 is given when the child performs the behavior inconsistently or requires assistance, while a score of 0 indicates that the child has not yet developed the functional skill or fails to demonstrate the target behaviors consistently.

The current data include samples from both AEPS-2 and AEPS-3, provided by AEPS Publishing (i.e., Brookes Publishing). Generally, the sample consists of children with or without disabilities who range in age from birth to 6 years and 11 months. The eligible group refers to children considered at-risk for delays or disabilities, while the ineligible group refers to developmentally typical children who do not receive DEA services. A convenience sample was recruited from states currently using AEPS in this study. Depending on how participants are recruited, programs include home visits, parent/early childhood groups, childcare centers, publicly funded preschool programs, and Head Start programs. When children were assessed, they were assessed by only one assessor, but most assessors did complete the AEPS-3 with multiple children during their

participation in the study. More precisely, the AEPS-2 data came from the AEPSi, which is an online AEPS-2 data collection system. The AEPS-3 data come from the program that uses AEPS in the states that the publisher identified.

### 3.2.2 Item Calibration Procedure

In this section, we first present an overview of the Rasch model and the distinction between Rasch measurement theory and the Item Response Theory (IRT). We then briefly explore its historical application in Early Childhood Education (ECE) assessment.

#### *3.2.2.1 Rasch Model as the method of analyses*

The Rasch model, named after Danish statistician Georg Rasch, is a specific one-parameter logistic (1PL) model within the broader IRT, widely used in psychometrics and educational statistics (Sundberg, 2019). It stands out for its simplicity, strict assumptions, and focus on item difficulty as the sole parameter. Further details regarding the differences between Rasch Measurement Theory and IRT can be found in Section 3.2.2.2. In this study, the Rasch Model is used to analyze categorical responses, providing accurate information about a person's ability, item difficulty, rating scale, rater severity, and other traits. This information enables examining and improving the performance and quality of instruments, such as constructing a scale, analyzing item quality, monitoring instrument quality, and measuring changes in a person's ability or item difficulty (Andrich, 1988; Bond, Yan, & Heene, 2020; Lynch & McNamara, 1998; Rasch, 1993; Wright & Masters, 1982).Before using the Rasch Model as the primary method for analyzing the study's categorical data, it is essential to understand its small set of

assumptions. In this study, I am adopting the Rasch Model based on a set of assumptions outlined by Trevor Bond & Fox (2013, page 45). The model assumes the following: (a) the capability to recognize and arrange observations of behavior along a continuum of none/some/more/all, which depends on a guiding underlying theory; (b) attentiveness to the sequence in which the skills or abilities under investigation are acquired (i.e., the model is expected to disclose the order of developmental acquisitions); (c) the potential to calculate the distances between the hierarchically arranged developmental skills or individuals; and (d) the proficiency to ascertain the overall development pattern exhibited among items and individuals, which can be generalized across all items and individuals.

#### *3.2.2.2 The Brief History of the Application of Rasch Model in the Early Childhood Education Assessment Field*

The potential benefits of using the Rasch Model in ECE/ECSE assessment were recognized over 25 years ago (Garwood, 1982; Sheehan, 1982; Snyder & Sheehan, 1992). Various early childhood assessments have been developed using Rasch model (Berry, Bridges, & Zaslow, 2004; Meisels, 2007)

A growing trend in educational literature advocates for the use of Rasch model in creating early childhood assessments that describe sequences, growth patterns, and ability levels (Wright, 1999). The current literature emphasizes using Rasch model as a means to obtain information about a child's relative position in their ability on a specific developmental path ordered by difficulty (Meisels, 2007).

#### *3.2.2.3 Rasch Model with Category Data*

Rasch (1960) developed a measurement model for responses to dichotomous items. As the simplest model in the Rasch model family, this model excludes the score categories, and the items were scored either correct or incorrect. In this two-categories data format, the number of correct scores equals the number answered correctly. The following equation defines the dichotomous Rasch model:

$$P(x_{ni} = 1|\theta_n, \beta_i) = \frac{e(\theta_n - \beta_i)}{1 + e(\theta_n - \beta_i)}$$

Where $x_{ni}$ is the score of student *n* for the item *i*, $x_{ni}$ =1 present the correct response, and $x_{ni}$ =0 present incorrect response. $\beta_i$ present the item difficulty and the $\theta_n$ presents the person ability. In this dichotomous Rasch model, *e's* exponential form was used to raise the power of $(\theta_n - \beta_i)$. According to the equation, the probability of student n answering item i correctly ($P(x_{ni} = 1)$) was decided by the difference between student's ability ($\theta_n$) and item difficulty ($\beta_i$).

An alternative expression of the dichotomous Rasch model is in terms of log odds which can transfer the model to a simple linear function of the ability parameter and item difficulty parameter:

$$Ln\left[\frac{p_{nik}}{1 - p_{nik}}\right] = \theta_n - \beta_i$$

Where the log odds of the probability of response to the item equals the difference between the person ability and the item difficulty. When the person ability equals the item difficulty, the ratio of the probability of successes to the probability of failures is 1, which also means the chance to answer correctly and incorrectly is 50% vs. 50%.

Since the AEPS's score system includes a three-point rating scale (2,1,0), the polytomous Rasch model will be introduced next, one of the widely used extensions of the dichotomous Rasch model. The rating scale model adds threshold parameters to the basic Rasch model that describe the rating scale's function (Andrich, 2005b). Reflecting on the special situation of response probability that was mentioned before, the chance of correct response vs. incorrect response is 50% vs. 50% when the person's ability level equals the item difficulty level. In the rating scale model, the threshold parameter refers to the location where a person ability has an equal probability (50%) to respond to one of two adjacent categories. The transition point demonstrates the location of the highest uncertainty of a person's response between two adjacent categories. These transition points are called Rasch-Andrich thresholds (Bond & Fox, 2013; Linacre, 2006a, 2010b; see also Andrich, 1998, 2005b). Based on this rationale, the RMS model's function is to extend the dichotomous data format to the polytomous data format in the Rasch model family (J. M. Linacre, 2000). The log odds form of the RSM is defined as the following equation:

$$Ln\left[\frac{p_{nik}}{p_{nik-1}}\right] = \theta_n - \beta_i - \tau_k$$

Where the $p_{nik}$ refers to the probability of examinee n responding to item i with the category k, and the $p_{nik-1}$ refers to the probability that examinee n responds to item i with the category k-1. $\tau_k$ present the threshold parameter which is the difficulty of responding with category k (relative to k-1). The log odds of the probability of response to the item equals the person's ability, item difficulty, and threshold. The RSM assumes the same threshold parameter across all items. In the AEPS, all the items are scored with

three points rating scale (2,1,0). When the child's behavior consistently met the criterion that the experts described, the item was scored 2, which means the child can perform the functional skill independently across time, materials settings, and people. When a child's behavior was performed with the assistance of certain people or in a certain setting, the behavior was scored 1. A child is scored 0 when they lack functional skills or cannot be observed exhibiting the target behaviors across various times and settings.

The AEPS scoring system offers a consistent threshold structure for all items in each developmental component of the test. Consequently, in this research, the Rating Scale Rasch model will be employed to examine the construction of the scale.

*3.2.2.4 The Operational Definition for Each Domain*

The following section provides an operational definition for each domain. The AEPS test features three distinct methods for collecting assessment information. Observation is considered the primary method, enabling the collection of more comprehensive details about a child's behavior (e.g., form, frequency, environmental factors). The main form used is the child observation data recording form. AEPS-2 comprises two age-level recording forms: birth to three years level and three to six years level. The content of six developmental areas for both age levels are presented in Table 3.2 as follows:

Table 4 The content of six developmental areas

| Area | **AEPS-2 (Birth to three)** | **AEPS-3 (Birth to Six)** |
|---|---|---|
| **Fine Motor Area** | In this domain, the assessment emphasizes evaluating the essential skills associated with grasping, reaching, and manipulation. | This area assesses a range of fine motor skills, including hand-eye coordination, finger dexterity, and manual dexterity. |

| | | |
|---|---|---|
| **Gross Motor Area** | Items in this area are designed to examine abilities of walking, running, jumping, climbing, and maintaining balance during dynamic movement. | The test assesses a range of gross motor skills, including balance, coordination, and strength. |
| **Adaptive Area** | Items in this area focus on assessing the skills related to feeding, hygiene, and undressing, which are essential for a child's self-care and independence. | The evaluation will examine a range of skills, such as using utensils for feeding, chewing and swallowing food, washing hands, brushing teeth, and maintaining overall cleanliness. Additionally, it will assess the child's ability to undress, remove shoes, and handle different types of clothing fasteners, such as buttons and zippers. |
| **Cognitive Area** | This area of assessment provides a comprehensive evaluation of an individual's cognitive and problem-solving abilities, as well as their understanding of key indexing concepts. | The items within this assessment area focus on evaluating an individual's responses to environmental simulations, problem-solving, and concepts related to indexing, such as object permanence, causality, imitation, and object differentiation. |
| **Social-Communication Area** | The items in this area evaluate an individual's social-communication interactions, as well as their comprehension and word production skills. | This area of assessment focuses on evaluating an individual's receptive, expressive, and social communication skills. The assessments aim to measure an individual's ability to understand and respond appropriately to social cues, engage in effective communication with others, and produce language effectively. |
| **Social / Social-Emotion Area** | The items within this assessment evaluate an individual's skills related to interacting with adults and peers, as well as their ability to respond appropriately to social conventions. | The items in this assessment area evaluate an individual's skills related to interacting with both adults and peers, responding to social conventions, understanding and responding to the environment, knowledge of self and others, and group participation skills. |

The AEPS-2 birth to three-year test is a 249-item observation instrument that assesses children's crucial developmental skills across six specific developmental domains: fine motor area, gross motor area, adaptive area, cognitive area, social-communication area, and social-emotional area (Diane Bricker & Waddell, 2002). The

whole scale was not assessed due to the multi-dimensionality and the specific interests of score equating in each area. Instead, each domain within the survey was analyzed as a unidimensional component and contributing construct. All items on the AEPS level I scale have possible responses on a three-point scores scale format (2, 1, or 0). After examining the wording of the item descriptions, no negatively worded items were found to exist on the scale. The whole scale (249 items) has been divided into six domains, as shown in Table 3.

Studies of the AEPS-2 have consistently demonstrated acceptable psychometric properties for both age-level scales. Utilizing classical test theory, two studies (Gao & Grisham-Brown, 2011; Grisham-Brown, Hallam, & Pretti-Frontczak, 2008) indicated that AEPS-2 is a reliable measure of child development. At the same time, other efforts have resulted in the construction of AEPS-2, which aims to provide a precise measurement on the individual level and predict various outcomes in child performance (Diane Bricker et al., 2008; Diane Bricker, Yovanoff, Capt, & Allen, 2003; Castaneda-Villa & James, 2007; Gao & Grisham-Brown, 2011; Wang, Sandall, Davis, & Thomas, 2011).

According to dynamical systems theory, child development includes physical, social, emotional, cognitive, and language development. Different domains are heavily mediated by each other. Motor development adheres to predetermined genetic programming, thereby following theoretical milestones (Bertenthal & Clifton, 1998). Gesell also points out that physical growth is a transformation of "…architectonics of the actions system." According to the biological rules of human growth, child development follows a specific hierarchical sequence. The body generally develops sequentially from the head and neck, then to the trunk, and finally to the lower limbs. Motor development

follows a sequence, which starts from the trunk to the limbs, then to the hands and feet, and finally to the fingers and toes. The AEPS evaluation test is deliberately divided into gross motor development and fine motor development areas. Complementing the specific definitions of child development milestones (CDC, 2021), in the AEPS test level I (birth to three years), the gross motor area includes four strands: movement and locomotion in supine and prone position; balance in sitting; balance and mobility; and play skills. Simultaneously, the fine motor includes two strands: reach, grasp, and release; and functional use of fine motor skills.

Child development is a multifaceted process, with psychological development occurring concurrently with physical development. According to the CDC's definition of children's health and development, "children of all abilities, including those with special education needs, can grow up in environments where their social, emotional, and educational needs are met." (CDC,2023). The AEPS Level I (birth to three years) encompass adaptive, cognitive, social-emotional, and social-communication domains. The adaptive domain includes feeding, personal hygiene, and undressing. The cognitive domain covers sensory stimuli, object permanence, causality, imitation, problem-solving, interaction with objects, and early concepts. The social-communication domain features four strands: prelinguistic communicative interactions; transition to words; comprehension of words and sentences; and production of social-communicative signals, words, and sentences. Lastly, the social domain consists of interaction with familiar adults, interaction with the environment, and interaction with peers.

The AEPS scale was analyzed within the Rasch model in this study. Compared to many alternative item response models (Boone et al., 2013), the Rasch model requires all

items to be equally sensitive to participants' person ability and responses involved in no guessing behavior. In the AEPS test, the scores were collected through observation by professionals. Child guessing behavior is rarely involved in the AEPS Test scoring process; therefore, the guessing parameter is zero. As a model assumption, the item discrimination is one. Besides, one of the Rasch model assumptions is unidimensionality. The model fit statistics evidence of unidimensionality will be provided to prove the validation of the analysis process.

### 3.2.3 Scale Validation

To explore a scale's validation and reliability under the Rasch model, it is important to consider several criteria, including unidimensionality, item fit, person fit, targeting, and reliability. These criteria are critical for ensuring that the scale is accurately measuring the underlying construct in a consistent and meaningful manner. Here are some additional details about each of these criteria:

#### *3.2.3.1Unidimensionality*

The dimensionality analysis of AEPS was discussed by Winchell (2011) and Toland, Grisham, Waddell, Crawford, and Dueber (2021). In Winchell's study, with confirmatory factor analysis and exploratory factor analysis, the six components paralleled the six developmental areas in the AEPS-2. The model data fit statistics also determined each of the six developmental areas (Diane Bricker et al., 2008). Furthermore, the Rasch Rating scale model was applied in each developmental area. Principal component analysis of the standardized residuals (PCAR) was further conducted to evaluate the unidimensionality assumption in each area. Below is the

operational definition of fundamental unidimensionality with the three criteria defined by Linacre (1998, 2021). Toland summarized these three criteria as 1) the variance explained by the measure should reach 50%, 2) the eigenvalue of the first contrast of the standardized residuals should be less than 2.0 (Arrindell & Van der Ende, 1985), 3) the ratio of the variance explained by the Rasch dimension to the variance explained by the first contrast of the residuals should be high (2021). If any issues were detected with the three criteria about unidimensionality, then, the next step was to inspect if the eigenvalue of contrast of the standardized residuals is higher than two. Researchers can identify items clustering at high or low loadings. When no clustering of the first contrast is present, unidimensionality can be considered plausible. Otherwise, a professional group must conduct an item content analysis to determine the meaning of the construct.

#### *3.2.3.2Data-model Fit*

Infit and outfit are two types of fit statistics used in the Rasch model to evaluate the degree to which the observed data match the model's expectations. Both infit and outfit statistics help identify problematic items that do not fit the model well, but they focus on different aspects of the data:

Infit (Information-weighted fit): The infit statistic is sensitive to the pattern of responses for items that are targeted towards an individual's ability level (Linacre, 2003). It is more concerned with the unexpected behavior of respondents on items that should be informative for their ability level. Infit gives more weight to the responses of individuals who are close to the item's difficulty level, as it is calculated using the squared residuals weighted by the information function.

Outfit (Outlying fit): The outfit statistic is sensitive to unexpected response patterns on items that are either too easy or too difficult for the individual's ability level (Linacre, 2003). The outfit gives equal weight to all residuals, regardless of the individual's ability level relative to the item's difficulty level. This statistic helps identify outliers that may be affecting the overall fit of the data to the model.

Both infit and outfit statistics are reported as mean square values (MNSQ) and the standardized fit (ZSTD). MNSQ with a value close to 1 indicates a good fit to the Rasch model. Values significantly greater than 1 suggest that the item exhibits more noise or randomness than expected, while values significantly less than 1 indicate that the item is overly predictable and may not be contributing useful information to the measurement of the latent trait (Linacre, 2003).

Infit and Outfit item indices were assessed to determine if items follow the consistent pattern with the model (i.e., data-model fit). If one item's Infit mean-square residual values or Outfit mean-square residual values are in the range of 0.5 to 1.5, this item was kept as an item that fit the model (Smith, 1995). If an item's two indices fall outside the range of 0.5 to 1.5, it is recommended to consider removing that item. This process continues until all remaining items exhibit acceptable fit statistics.

ZSTD is typically based on the standardized residual, which is calculated as the difference between observed and expected responses, divided by the standard error of that difference. A ZSTD value of 0 suggests a perfect fit, implying that the observed data perfectly match the model's predictions. Positive values indicate that the item is more unpredictable than the model predicts, a situation known as overfitting, whereas negative values suggest the item is less unpredictable than the model predicts, known as

underfitting. A ZSTD value within the range of -2 to +2 is often considered acceptable, as this indicates the data are not deviating from the model's predictions excessively.

#### *3.2.3.3Separation & Reliability*

The last index that needs to be checked is separation and reliability. In the Rasch model, separation measures how well the scale items distinguish between individuals with varying levels of the construct. A separation of 2 or higher indicates that the scale can differentiate at least two groups (Linacre, 2023). The separation value is a noise-to-information ratio that represents the proportion of the true score to the error in the observed score. Higher separation values indicate a higher portion of the true score within the observed score. Separation tells you how many statistically distinct strata (groups) of person ability the test can identify.

Reliability is a measure of the consistency and stability of the scale scores over time or across different samples of individuals. In the Rasch model, reliability is calculated as the ratio of the true score variance to the total score variance. The true score variance is the variance of the underlying construct being measured, while the total score variance is the sum of the true score variance and the error variance. Reliability values range from 0 to 1, with higher values indicating greater consistency and stability of scale scores. A scale is considered consistent and stable when the value of person reliability measures exceeds 0.8 and the value of item reliability measures exceeds 0.9. (Linacre, 2023). Reliability measure in the Rasch model is equivalent to the conventional Kuder-Richardson Formula 20 and Cronbach's alpha indices of measurement reproducibility.

## 3.3 Equating Procedure

Given the history of AEPS development, the data were found to naturally align with the common item nonequivalent groups design. To select the common items, I used cosine similarity calculations and applied two criteria based on the items' descriptions: exact matching and functional matching. The results of both matching processes were then evaluated by subject matter experts through a qualitative review. The fixed parameter calibration linking method was then utilized, wherein the item parameters were calibrated separately for each assessment during the equating process. The anchor item parameters from the new scale were transformed to the old scale, and the final step involved creating a true-score and observed-score conversion table based on the results of equating.

### 3.3.1 Selection of Anchor Items

In the second step of the data analysis, the focus was on selecting high-quality common items. To achieve this, cosine similarity was employed to match the content descriptions of items between AEPS-2 and AEPS-3. Two criteria were used during this process. The first criterion was "identical match," which involved selecting items whose descriptions were exactly the same, excluding the replacement of synonyms or sentences. For example, "Indicates need to use toilet" and "Indicates toileting" were considered not an exact match. The second criterion was "functional matching," which involved selecting items whose descriptions were very close but not identical, with a threshold of 90% match. For instance, the similarity score between "Locates object in second of two hiding places" and "Locates object in latter of two successive hiding places" is 0.97. This

pair of items was considered part of the anchor set after the functional matching process. Table 16 and Table 17 offer additional information regarding the anchor item content, while the complete list of item content can be found in appendix 3 and 4. Further discussion on functional matching is available in Chapter Four, Section 4.3.2 - Functional Matching (page 64).

It is important to note that the set of common items/anchor test ideally needs to be a mini test that does not lose any critical statistical features during the equating process (Liu, Sinharay, Holland, Curley, & Feigenbaum, 2011). However, Sinharay and Holland (2007) suggest that the set of common items/anchor test can be more flexible. Therefore, the selection criteria for common items in this study were tolerant and allowed for greater flexibility in the location selection process.

The second step, which involves selecting the functional anchor item, will be a data-driven process since the number of the item is higher than 30% in the six developmental areas. This step introduces a new concept: item parameter drift. Within the Rasch framework, item parameter drift signifies changes in a test item's difficulty level either over time or across different groups. Anchoring necessitates fixing item parameters to ensure test consistency. However, if these parameters significantly alter, it denotes item parameter drift, posing a challenge to maintaining test comparability. In this study, the item parameter drift was checked using the displacement parameter in Rasch analysis for AEPS-2 after anchoring to further determine the selection of common items.

Based on the previous study of the anchor test length (Budescu, 1985; Ricker & von Davier, 2007; Yang & Houang, 1996), a rule of thumb is that a 20% threshold can reach the level of relative efficiency. Additionally, Ricker and von Davier (2007)

demonstrated that the longer length of the anchor test reduced the standard errors of equating in the common item nonequivalent groups design. Since the percentage of identical common items in all six areas of the birth to three level AEPS-2 varies (i.e., 3% to 49%), to ensure that the ratio reaches at least 20% in all six areas, two sets of common items (one including exact match common items, and another excluding functional matching common items) will be applied in the equating process to examine the level of efficiency.

### 3.3.2 Equating Method: Fixed Parameter Calibration (Anchoring)

Kolen and Brennan (2004, p.430) suggested using calibration to equate the observed-score or true scores when the two assessments have the same frameworks, or the framework is viewed as sharing common features and/or use. The same frameworks or common features provide a set of items in the tests for applying calibration in the common-item nonequivalent groups anchor test design. The three most commonly used methods of calibration in equating are 1). Concurrent calibrations, 2). Separate calibration with transformation, 3) and fixed parameter calibration. This study conducted the parameter calibration (Hanson and Béguin 2002; Kang and Petersen 2009; Kim 2006) to estimate the item parameters with the item level data to link assessment A to assessment B, the first step is to establish the scale for assessment B, just as would be done under the separate calibration method. Next, the items from assessment A are projected onto the established scale for assessment B by calibrating the items from assessments A and B together but keeping assessment B's item parameters fixed. Compared to separate calibration, fixed-parameter calibration does not require an item transformation method to place items from one assessment onto the scale of the other.

The major steps involved in anchoring in this study include: 1) calibrating the AEPS-2 items and the AEPS-3 scale separately; 2) using the common item set to estimate the AEPS-2 item difficulty parameter on the AEPS-3 scale; and 3) estimating population proficiencies using the AEPS-2 item parameter after anchoring for the AEPS-2 sample.

#### *3.3.2.1 The Process of Converting True Score*

The ultimate goal of implementing equating in this study is to create a conversion table for children's developmental scores, which includes person ability measure and observed score conversion tables. These two tables allow for the exchange of the person ability measure or observed score between two AEPS test versions. This process ensures that the scores obtained on different versions of an assessment can be accurately compared and interpreted.

In the case of the AEPS-2 and AEPS-3 assessments, the conversion of person ability measures is necessary to equate an individual's measure on the AEPS-2 scale to the transformed measure on the AEPS-3 scale. The equating process helps to transform the items of AEPS-2 onto the AEPS-3 scale, and then the person ability measures on the same scale (i.e., AEPS-3), which means they are comparable. The conversion allows for a more accurate and reliable evaluation of an individual's progress over time and provides a comprehensive analysis of their skills and abilities across different assessment periods.

#### *3.3.2.2 The process of converting observed score*

In the process of converting a person's observed score, I first used the Rasch model to calibrate the item parameters for each of the tests separately. Once the item parameters have been calibrated separately for each test, the anchor item parameters from

the new test are transformed to the scale of the old test using the item parameter estimates.

Next, I calculated the observed score for each individual on the new test using the transformed item parameter estimates from the previous step. This involves summing the observed scores for each individual and converting them to a score on the old test scale using the conversion table of the person ability measure equating process.

Overall, the observed score calculation process after anchoring in the Rasch model involves: a) transforming the item parameters from the new test to the scale of the old test, b) calculating the observed scores for each individual on the new test using the transformed item parameters, and c) comparing the observed scores on the old and new tests to create the observed score conversion table between the two versions.

### 3.3.3 Evaluation of the Equating

In this section, the general rule of evaluation of the results of equating under the Rasch model framework and the special situation in the anchoring were introduced.

#### *3.3.3.1 General rule of evaluation in the equating*

The first and second-order properties of equating can be used to evaluate the equating process. The goal of equivalence is that they will get similar results no matter which tests the examinees take. The equity manifested in two aspects: the similarity of test takers' scores, the fairness of the first order, the variance of examinees' scores (measurement error). The second-order fairness is similar.

Inaccurate test score equivalence results can mislead high-stakes decision-making. The gold standard for evaluating the variability (or inaccuracy) of equating results is the measure of standard error. Two types of standard errors can occur: systematic and random errors (Kolen, 1988). Systematic errors exist when the correct application criteria are not followed, bias is present in the equating method, or assumptions of the method or model are violated. Incorrect implementation of equating designs or different alternative forms can also cause systematic error (Kolen & Brennan, 2004). The standard error indicates the number of random errors that exist in the equating process; based on the mathematical equation, the random error in equating is the standard deviation between the equivalence of test X and test Y. Therefore, the standard error of the sample equating error is the SEE. RMSE estimates the total error in equating and is the square root of the sum of bias and SEE. Random errors arise when the sample size is limited, including the participants' sample size and the anchor item selection. Due to sampling, some uncertainty exists between the estimation and the true value of item parameter equating. The standard error of equating is inversely related to the participants' sample size. Meanwhile, Michaelides and Haertel (2014) noted that the selection of common items shows the dominant influence on the standard deviation of equated scores over hypothetical replication after the sample size increases.

#### *3.3.3.2 Evaluation in this study*

The Rasch model is considered invariant because it assumes that a person's response to an item is only influenced by their level of the underlying construct and the item's difficulty, without being affected by external factors such as time, context, or

sample characteristics. Therefore, the model is invariant across different contexts, time periods, and samples as long as the construct being measured is consistent.

Fixed parameter calibration equating, which is a critical process in large-scale testing, involves identifying item functions in the anchor set over time using screening statistics, such as item-level displacement. The success of equating relies on employing an appropriate set of anchor items to preserve the existing scale's integrity.

In cases where both the old and new forms of the scale examined by the Rasch model are of high quality and no additional steps are needed for transforming the scales during the anchoring process, the evaluation step would involve visually checking the displacement parameter of the anchor items, or items' displacement parameter. This is because the Rasch model provides a fixed set of parameters that remain consistent across different samples. If the underlying construct being measured is the same, the model is invariant. If not, the estimation of displacement would reveal the inconsistency between the two samples.

### 3.4 Chapter Summary

Chapter Three presents the methodology employed in this study. The Common-Item Non-equivalent Group equating design is applied, considering the similarities and differences in the assessment structure, which are closely tied to the data structure. The psychometric properties of both assessment versions are thoroughly evaluated using the rating scale Rasch model. To establish anchor item sets, the cosine similarity coefficient and expert content check are employed, resulting in the selection of identical and functional matching items. Subsequently, a fixed parameter calibration (anchoring)

process is conducted to estimate both item and person parameters. Lastly, the equating process is carefully evaluated, and person measure and observed score conversion tables between AEPS-2 and AEPS-3 are generated, providing valuable insights into the relationship between the two assessment versions.

# CHAPTER 4 RESULTS

This chapter contains five sections presenting the results of the analysis. In the data sample section, I provided an overview of the data sample for AEPS-2 and AEPS-3 tests. Section two provides information regarding the dimensionality (e.g., research question 1) of the AEPS-2 and AEPS-3 tests. The item structure of both AEPS tests (e.g., research question 1) is analyzed in the third and fourth sections. The fourth section presents the results of scale (e.g., research question 1.) analysis. Finally, the results of the scaling equating investigation (e.g., research question 2, 3, 4) are presented.

## 4.1 Data Sample

The AEPS-2 data is collected from the online data collection system called AEPSi, which is supported by Paul H. Brookes Publishing Company. The online system originally consisted of two separate data files, each with one level. Data preparation was conducted at both levels, including identifying duplicate cases, calculating ages, locating missing data, and deleting cases with missing ages or ages of less than 0 months. Cases with ages greater than 96 months in the AEPS-3 dataset or greater than 36 months in the AEPS-2 dataset were also deleted, and invalid scoring entries were excluded. Cases with missing data were examined, and those with all missing data were discarded. A total of 83 to 117 cases were excluded for one or more of the above reasons in six areas of the AEPS-3, respectively. For Level I of the AEPS-2 Test, 558 cases were excluded in each area.

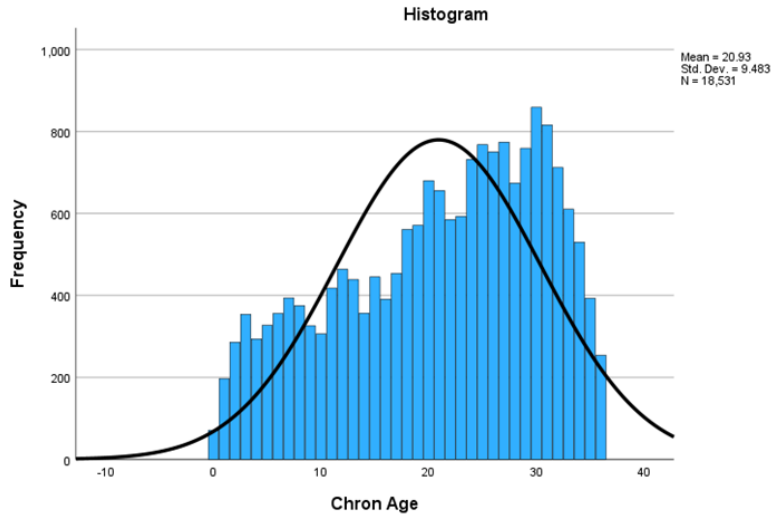
Descriptive analysis was conducted for Level I of the AEPS-2 Test. The means, standard deviations, and frequencies for the provided variables were determined.

Demographic variables, such as a child's gender, developmental status, state information, and age, were also reported.

### 4.1.1 Descriptive overview of AEPS-2 test level I Sample

The sample size for Level I of the AEPS-2 Test included 18,531 cases comprised of 6,675 females and 11,856 males. The children ‘s status included children who were at-risk (n = 789); children who were typically developing (n = 445); and children who had developmental delays or disabilities (n = 17,283). Children’s age ranged from 0 months to 36 months with a mean of 20.93 months (SD = 9.48). Refer to Figure 2 for the distribution of children by age in months. The data set included children from 21 states representing all geographic quadrants of the United States. The frequency distribution of the sample by age for Level I was skewed to the left indicating a predominance of child records for ages 36 months and younger.

Figure 2 The Histogram of the Age Distribution for Level I Cases

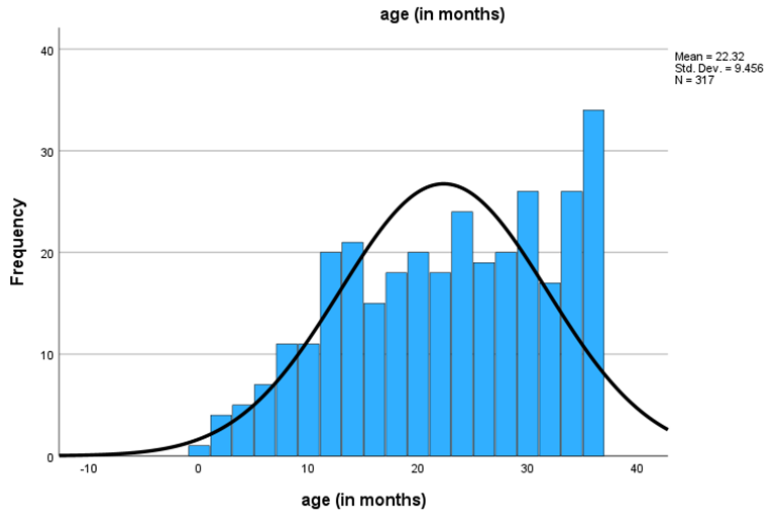


### 4.1.2 Descriptive Overview of AEPS-3 Test

Similar to the previous section, the AEPS-3 dataset included a descriptive analysis to determine the dataset's means, standard deviations, and frequencies. Demographic variables such as a child's gender, developmental status, numeric identification code, state label, and chronological age are reported next.

AEPS-3 Test sample size included 317 cases with 123 females and 194 males. 149 at-risk kids, 168 typically developing kids participated in the study. The children ranged in age from 0 months to 36 months, with a mean age of 22.32 months (SD = 9.46). The age distribution of children is shown in Figure 3. Children from 8 states of the country were included in the data set. The frequency distribution of the sample by age for AEPS-3 was skewed to the left indicating a predominance of child records for ages range from 0 to 36 months. For AEPS-3 cases, see Figure 3 for a histogram.

Figure 3 The Histogram of the Age Distribution for AEPS-3 Cases



### 4.1.3 Descriptive Statistics of the Data Related to the Anchor Design

Descriptive statistics for the two adjusted data sets, the final use AEPS-2 data and AEPS-3 data in the six developmental areas, are reported in Table 6, respectively. Table 6 displays the inter-item correlation statistics for both AEPS-2 and AEPS-3 across several developmental areas. The adaptive area shows an average inter-item correlation of .542 for both AEPS-2 and AEPS-3, with a range from .086 to .942. In the cognitive area, the average inter-item correlation is .414 for AEPS-2 and .533 for AEPS-3, with respective ranges of .076 to .902 and -.191 to .918. For the fine motor area, AEPS-2 has an average inter-item correlation of .512, ranging from .074 to .937, while AEPS-3 has an average of .482, ranging from 0.099 to .916. In the gross motor area, AEPS-2 has an average inter-item correlation of .583, ranging from .074 to .975, whereas AEPS-3 has an average of .460, ranging from 0.032 to .967. Regarding social communication, AEPS-2 exhibits an average inter-item correlation of .612, ranging from .144 to .923, whereas AEPS-3 has an average of .455, ranging from 0.093 to .931. Lastly, in the adaptive area, AEPS-2 has an average inter-item correlation of .501, ranging from .152 to .845, while AEPS-3 has an average of .509, ranging from -.285 to .921. As shown in Table 6, in the five of six developmental areas (e.g., Adaptive area, cognitive area, fine motor area, gross motor area, and social emotion area), the average total score is higher for the AEPS-3 compared to the AEPS-2 based on the number of the item. In addition, the average common item score is higher for AEPS-3 compared to the AEPS-2 in the adaptive area, fine motor area, gross motor area, and social emotion area. The difference between the common-item means indicates that the AEPS-3 is a higher achieving group than the AEPS-2 under the classical testing theory. The distributions of the total scores and common-item scores are

negatively skewed for six areas. With the exception of the AEPS-3 scores in fine motor and gross motor, the total scores and common-item scores in the other areas exhibited flatter distributions compared to a normal distribution (kurtosis = 0). Cronbach coefficient alphas are computed as reliability of the scores for both forms in six areas. The reliability of the scores of the both forms in the six developmental areas are higher than 0.95, and the correlations between the total scores and the common item scores for the AEPS-2 form and the AEPS-3 form are all higher than 0.8 (the range is 0.822 to 0.991.

Table 5 Descriptive Statistics of Raw Scores for Both Versions: Total Items and Common Items in Six Developmental Areas

| | **Total Item score** | | | | | | | | | | | | **Common item score** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AEPS-2 | | | | | | AEPS-3 | | | | | | AEPS-2 | | | | | | AEPS-3 | | | | | |
| | Adapt | GM | FM | Cog | Soc-C | Soc-E | Adapt | GM | FM | Cog | Soc-C | Soc-E | Adapt | GM | FM | Cog | Soc-C | Soc-E | Adapt | GM | FM | Cog | Soc-C | Soc-E |
| **Mean** | 33.04 | 73.98 | 43.03 | 57.95 | 86.7 | 28.60 | 70.01 | 103.34 | 51.96 | 60.67 | 57.01 | 81.72 | 12.18 | 40.15 | 2.74 | 12.18 | 8.93 | 4.42 | 17.97 | 48.21 | 4.85 | 10.33 | 5.81 | 6.28 |
| **SD** | 17.06 | 33.337 | 17.64 | 26.79 | 21.93 | 13.41 | 29.89 | 26.96 | 12.89 | 30.06 | 32.59 | 35.64 | 6.48 | 16.91 | 2.31 | 6.48 | 1.79 | 2.49 | 5.65 | 10.61 | 1.94 | 4.56 | 2.71 | 2.01 |
| **Skewness** | -3.07 | -.85 | -.78 | -.11 | .49 | -.30 | -5.92 | -1.74 | -1.75 | -.31 | -.36 | -1.08 | -.32 | -1.03 | 0.09 | -1.03 | -.38 | -.24 | -1.45 | -2.55 | -1.56 | -1.45 | -1.06 | -.57 |
| **Kurtosis** | -.921 | -.72 | -.49 | -.776 | -.39 | -.89 | -8.94 | 2.53 | 2.61 | -1.21 | -1.22 | -1.00 | -1.12 | -.472 | -1.53 | -1.27 | -.24 | -1.00 | 1.01 | 5.93 | 1.01 | -.28 | -.20 | -.17 |
| **Correlation** | 0.97 | 0.99 | 0.87 | 0.94 | 0.89 | 0.93 | 0.92 | 0.96 | 0.92 | 0.92 | 0.90 | 0.82 | | | | | | | | | | | | |

Note. 1). Adapt, GM, FM, Cog, Soc-C, Soc-E mean Adaptive Area, Gross Motor Area, Fine Motor Area, Cognitive Area, Social Communication Area, Social Emotional Area, respectively. 2). Correlation means the correlation between common item scores and total item scores.

## 4.2 Result of Scale Calibration

To support measurement validity and transform the raw scores into meaningful metric values, scale calibrations are the first step, providing the foundation for equating different test forms and enable meaningful score comparisons across forms or administrations. The results of the separate calibrations in this section provide information for each scale, including statistics of dimensionality, Wright maps, model-data fit, separation, and reliability. All these details are presented below.

### 4.2.1 Dimensionality

To address the first research question one, I investigated the evidence for dimensionality, model-data fit, and scale reliability. Toland (2021) identifies three criteria for unidimensionality in the Rasch model, which were used to guide my investigation. Firstly, the ratio of raw variance explained by items compared to the unexplained variance in the first contrast should be high. Second, the variance explained by the measures should be >50% (Linacre, 2020a). Third, the eigenvalue of 1st contrast should be under than 2 and accounts for less than 5% of the unexplained variance. The area that satisfies two of the three criteria is considered functional unidimensionality.

For six developmental areas (e.g., fine motor area, gross motor area, cognitive area, social/social-emotional area, and social communication area.) of the AEPS-2 and the AEPS-3, PCARs (see Table 7) were conducted to compare the raw variance explained by items and the unexplained variance. Based on the first contrast index, the ratios do not raise concerns for Fine Motor, Gross Motor, Adaptive, Cognitive, social communication, and social-emotional skills. In other words, the primary Rasch dimension within FM,

GM, Adaptive, Cognitive, social-emotional, and social-communication areas tended to dominate about 8.0 (Social-Emotional) to 32.1 (Gross Motor) times the secondary dimension. Similarly, results show that the data for all six areas fulfill the criteria of Linacre (2020a) where variance explained by the measures should be >50% (i.e., 69.3% for social-emotional to 84.7% for GM). Although the eigenvalue of the first component of the residuals is greater than 2, it should be situated at the bottom of a standardized residual contrast 1 plot (not reported). With no evident item clustering for each area, the results indicate that each developmental area of the AEPS-2 can be regarded as fundamentally unidimensional.

Table 6 Dimensionality of AEPS-2 and AEPS-3

| **AEPS-2 Level I** | | | | | | |
|---|---|---|---|---|---|---|
| | Fine Motor | Gross Motor | Adaptive | Cognitive | Social | Social-Communication |
| **The ratios of primary dimension to the secondary dimension** | 17.5 | 32.1 | 12.2 | 19.7 | 8.0 | 15.7 |
| **variance explained by the measures** | 78.1% | 84.7% | 72.9% | 76.0% | 69.3% | 78.7% |
| **Eigenvalue of 1st contrast** | 3.1 | 4.4 | 3.6 | 5.0 | 3.2 | 6.0 |
| **AEPS-3** | | | | | | |
| | FM | GM | Adaptive | Cognitive | Social-Emotional | Social-Communication |
| **The ratios of primary dimension to the secondary dimension** | 13 | 24.3 | 17.8 | 6.7 | 4.6 | 11.7 |
| **variance explained by the measures** | 75.9% | 79.4% | 77.0% | 71.4% | 65.2% | 76.2% |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Eigenvalue of 1st contrast** | 2.9 | 5.1 | 3.9 | 7.8 | 7.4 | 5.8 |

### 4.2.2 Wright map

For this study, I utilized the Wright map to determine if the AEPS Test items target developmentally appropriate skills. The Wright maps are analyzed through visual examination, and they can be found in Appendices C and D. Figures C2 (AEPS-2 Level I: Wright Map: GROSS Motor Area) and D3 (AEPS-3: Wright Map: Fine Motor Area) provide examples of these maps.

A Wright map is a visual representation of Rasch results, simultaneously displaying both items and persons. The logit scale is employed to estimate the value of each item and person, with values expressed in logit units on an interval scale. The map positions the highest values at the top and the lowest values at the bottom. More challenging items have positive (higher) values, while more capable individuals also have positive (higher) values. The AEPS Test items are organized according to strands, which should correspond with the developmental order described in the literature.

Each Wright map presents all items within the corresponding developmental area of the AEPS Test (e.g., the Gross Motor area for AEPS-2 Level I include 58 items, all of which are displayed on the Wright map). Refer to figure 2 in appendix 1. In this example, multiple items share a single difficulty level and are clustered at the positive one logit location (see the upper "S" marker). A gap exists at the negative two logits location (see the bottom "T" marker), indicating a lack of items to measure children's gross motor skills at this level.

### 4.2.3 Data fit

In the Rasch model, data fit offers insight into how well the observed response patterns correspond to the model's predictions. Overfitting occurs when the data too closely matches the model, while underfitting transpires when data demonstrates more variability or unpredictability than predicted by the model. In these analyses, both the mean-square infit or outfit statistic (MSNQ) and the standardized z score (ZSTD) are derived to evaluate fit. Appendices E and F present more detailed information on the fit statistics for AEPS-2 and AEPS-3, respectively, according to developmental areas. Both MSNQ and ZSTD are reported. The expected value for MSNQ is 1, and fit statistics are interpreted within an accepted judgment range from 0.5 to 1.5. In terms of ZSTD, a value of 0 denotes a perfect fit, indicating that the observed data align perfectly with model predictions. Overfit in ZSTD, marked by a value greater than 2, occurs when data is more predictable than the model predicts, while underfit, signaled by a value less than 2, arises when the data is less predictable than anticipated by the model.

Table 8 indicates that some children with high abilities received a score of 0 or 1 instead of 2 for some items that were easy to approximate average in difficulty. Despite this, Outfit is known to be sensitive to unexpected responses (e.g., a response from a person whose location is well beyond or below where the item is measured; Linacre, 2002). To assess the impact of misfit, I replaced suspect responses with a missing value and ran a sensitivity analysis. Based on these sensitivity analyses, Outfit indices were acceptable. It was decided to retain the misfitting items as quality items.

Table 7 reveals that infit indices for all items in the adaptive motor area of AEPS-2 fall within the range of 0.74 to 1.53. Two items exceed 1.5 (infit = 1.53), but this is still

considered clinically acceptable. However, seven items' outfit indices fall outside the accepted range (0.5-1.5). In the social-communication area, two items (infit = 1.52, 1.54) exceed 1.5, which are still considered clinically acceptable. Twelve items' outfit indices (1.87-8.87) fall outside the acceptable range (0.5-1.5). In the social area, all item infit estimations (0.67-1.35) are located within the range of 0.5 to 1.5, and except for item B1.2, all other items' outfit are within the acceptable range. In the cognitive area, the infit values (0.81-1.39) fall within the acceptable range; however, ten items' outfit indices fall outside the acceptable range. In the fine motor area, the infit of item A5.4 is 1.56, and four item's outfit value (1.51-2.97) located outside of the acceptable range. In the gross motor area, 20 items' outfit indices (1.89-9.90) are located outside the acceptable range (0.5-1.5). The infit indices of items D4.3 (1.72), B2.2 (1.83) and D2.2(2.20) are also ill-fitting (more details see appendix 5).

In the AEPS-3 assessment, all item infit indices in the adaptive area fell within the acceptable range (0.5-1.5), with the exception of five items (A5.2, D1.1, A5.1, C1.6, C1.7). The infit z-value for this area spanned from -3.62 to 6.84. Within this area, 24 items displayed ill-fitting outfit MSNQ, and 21 items had outfit z-values outside of the acceptable range. Moving to the cognitive area, five item infit indices were recorded between 1.54 and 1.87, exceeding the accepted range (0.5 -1.5). A total of 27 items' infit z-values in the adaptive area were identified outside the acceptable range. In the fine motor area, three items fell outside the acceptable range, and the outfit indices for eleven items exceeded the acceptable range. Furthermore, ten and nine items had z-values that were located outside of the acceptable range. In the gross motor area, the infit parameters for eight items were identified outside the acceptable range, and a significant proportion

of the items' outfit indices (49/67) exceeded the accepted limit. In both the social and social-communication areas, three items in each were found with infit indices outside the acceptable range. Additionally, 13 and 28 items in these areas respectively had outfit indices that exceeded the acceptable range. A high percentage of items' z-values were also found exceeding the acceptable range across these areas (for more details, refer to appendix 6).

Table 7 Summary of item infit and outfit indices for AEPS-2 and AEPS-3

**AEPS-2 Level I**

| **Area** | Infit (MSNQ/Z-value) | | | Outfit (MSNQ/Z-value) | | |
|---|---|---|---|---|---|---|
| | Underfit | Fit | Overfit | Underfit | fit | Overfit |
| **Adaptive** | 2/1 | 30/13 | 0/18 | 6/10 | 25/3 | 1/19 |
| **Cognitive** | 0/20 | 58/10 | 0/28 | 10/17 | 48/10 | 0/31 |
| **Fine Motor** | 1/14 | 32/5 | 0/14 | 4/13 | 29/4 | 0/16 |
| **Gross Motor** | 3/12 | 52/13 | 0/30 | 20/23 | 23/4 | 12/28 |
| **Social-Communication** | 2/16 | 44/3 | 0/27 | 12/22 | 22/3 | 12/21 |
| **Social** | 0/12 | 25/2 | 0/11 | 1/6 | 24/2 | 0/16 |
| **Total Items** | 8 | 241 | 0 | 53 | 171 | 25 |

**AEPS-3**

| **Area** | Infit (MSNQ/Z-value) | | | Outfit (MSNQ/Z-value) | | |
|---|---|---|---|---|---|---|
| | Underfit | fit | Overfit | Underfit | fit | Overfit |
| **Adaptive** | 0/10 | 48/29 | 5/14 | 15/12 | 28/22 | 9/9 |
| **Cognitive** | 5/10 | 45/23 | 0/17 | 7/16 | 36/23 | 7/11 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Fine Motor** | 0/5 | 28/21 | 3/5 | 6/5 | 20/22 | 5/4 |
| **Gross Motor** | 7/7 | 59/37 | 1/21 | 34/27 | 15/24 | 15/14 |
| **Social** | 3/6 | 59/31 | 0/25 | 7/20 | 49/36 | 6/6 |
| **Social-Communication** | 2/4 | 46/33 | 1/12 | 15/3 | 21/34 | 13/12 |
| **Total Items** | 29 | 281 | 0 | 88 | 212 | 32 |

### 4.2.4 Separation and Reliability

Rasch model reliability is typically calculated using the Person Separation Index (PSI) or Person Reliability (PR), which are analogous to the concept of Cronbach's alpha in classical test theory. The PSI and PR are estimates of the consistency with which individuals can be separated into distinct performance levels based on their responses to the items in the assessment. A high PSI or PR value (typically above 0.7) indicates that the test is effectively distinguishing between individuals with different levels of the underlying construct.

As a result of the separation and reliability measures (see Table 8 and Table 9), all the item separation indices are more than 2 which means the item effectively separates the persons' different preference ability levels on the scale. Person reliability measures range from .94 (Social) to .98 (Cognitive), all exceeding the desired threshold of .80. A 1.00 item reliability measure is present for all AEPS-3 developmental areas, demonstrating that each test has items suitably dispersed along a developmental area continuum. Additionally, this implies that the likelihood of replicating item positions on the scale is high (with the recommended criteria value being 0.9).

Table 8 Summary of separation and reliability of AEPS-2

| AEPS-2 Level I | | | | |
|---|---|---|---|---|
| **AREA** | Separation | | Reliability | |
| | person | item | person | Item |
| **Adaptive** | 4.83 | 122.7 | 0.96 | 1.00 |
| **Cognitive** | 6.57 | 141.0 | 0.98 | 1.00 |
| **Fine Motor** | 5.59 | 128.1 | 0.97 | 1.00 |
| **Gross Motor** | 4.46 | 15.48 | 0.95 | 1.00 |
| **Social** | 3.97 | 109.9 | 0.94 | 1.00 |
| **Social-Communication** | 5.71 | 142.4 | 0.97 | 1.00 |

Table 9 Summary of separation and reliability of AEPS-3

| AEPS-3 | | | | |
|---|---|---|---|---|
| **AREA** | Separation | | Reliability | |
| | person | item | person | Item |
| **Adaptive** | 5.80 | 20.13 | 0.97 | 1.00 |
| **Cognitive** | 4.36 | 18.92 | 0.95 | 1.00 |
| **Fine Motor** | 2.79 | 14.96 | 0.89 | 1.00 |
| **Gross Motor** | 5.38 | 17.77 | 0.97 | 1.00 |
| **Social** | 4.46 | 15.48 | 0.95 | 1.00 |
| **Social-Communication** | 4.63 | 20.26 | 0.96 | 1.00 |

## 4.3 Equating

### 4.3.1 Common Item (anchor) Selection: Identical Matching

To select the common item set in each area, there are two steps. First, identical items are selected based on item descriptions. Second, items that are not invariant are removed based on calibration information (e.g., model-data fit). Third, the displacement parameter is used to detect the item functionality over time as the final anchor item.

Using cosine similarity matching, between AEPS-2 and AEPS-3, 1 In the gross motor area, I find twenty-seven identical common items and three in the fine motor area. In both the social communication and social emotion areas, there are four common items. The adaptive area holds eight identical common items, while the cognitive area contains a single identical common item. (See Table 10). The following is a list of item groups used to equate scores between AEPS-2 and AEPS-3: the fine motor area has three items; the social and social-communication areas both have four items; the adaptive area contains eight identical items; the gross motor area has 27 identical items, and the cognitive area has one final common item.

Table 10 Identical matching Anchor items were used in Fixed calibration equating (anchoring)

| Area | Nu. of item | AEPS-2 | AEPS-3 | Content | displace |
|---|---|---|---|---|---|
| **Fine Motor (32)** | 3 | A5.1 | B.3.2 | Grasps pea-size object using fingers in raking or scratching movement | -0.15 |
| | | B2.1 | B.3.3 | Aligns objects | 0.02 |
| | | B5.2 | C1.4 | Scribbles | 0.27 |
| **Gross Motor (55)** | 27 | A1. | A1 | Turns head, moves arms, and kicks legs independently of each other | -1.72 |
| | | A1.2 | A1.1 | Kicks legs | -2.11 |
| | | A1.3 | A1.2 | Waves arms | -0.29 |
| | | B1. | A4 | Assumes balanced sitting position | 0.14 |
| | | B1.4 | A4.4 | Sits balanced without support | -0.07 |
| | | B1.5 | A4.5 | Sits balanced using hands for support | 0.06 |

| | | | | | |
|---|---|---|---|---|---|
| | | B2.1 | A5.1 | Sits down in chair | 0.17 |
| | | A3.2 | B1.2 | Assumes creeping position | -0.26 |
| | | A3.3 | B1.3 | Crawls forward on stomach | -0.23 |
| | | A3.4 | B1.4 | Pivots on stomach | -0.51 |
| | | C1.4 | B2.2 | Stands unsupported | 0.63 |
| | | C2.2 | B2.3 | Pulls to standing position | 0.19 |
| | | C2.3 | B2.4 | Pulls to kneeling position | 0.14 |
| | | C1.1 | B3.1 | Walks without support | -0.08 |
| | | C1.2 | B3.2 | Walks with one-hand support | -0.16 |
| | | C1.3 | B3.3 | Walks with two-hand support | -0.06 |
| | | C1.5 | B3.4 | Cruises | -0.07 |
| | | C4.2 | B4.2 | Moves up and down stairs | 0.58 |
| | | C4.3 | B4.3 | Gets up and down from low structure | 0.64 |
| | | C3.1 | B5.1 | Runs | 0.08 |
| | | C3.2 | B5.2 | Walks fast | 0.02 |
| | | D1 | B6 | Jumps forward | 0.47 |
| | | D2 | C3.2 | Pedals and steers tricycle | 0.4 |
| | | D2.1 | C3.3 | Pushes riding toy with feet while steering | -0.45 |
| | | A3. | B1 | Creeps forward using alternating arm and leg movements | 0.08 |
| | | A2.1 | A3.1 | Rolls from stomach to back | -0.51 |
| | | A2.2 | A3 | Rolls from back to stomach | -0.35 |
| **Cognitive (58)** | 1 | D1.2 | E1.1 | Imitates familiar simple motor action | 1.27 |
| **Social Communication (46)** | 4 | A1. | A1 | Turns and looks toward person speaking | -1.56 |
| | | C1.5 | A1.1 | Quiets to familiar voice | -0.72 |
| | | D2 | C1.1 | Uses two-word utterances | 2.43 |
| | | B2.1 | C1.4 | Uses consistent consonant–vowel combinations | -0.29 |
| **Social (25)** | 4 | A1.2 | A1.1 | Responds appropriately to familiar adult's affective tone | 0.28 |
| | | B1 | E1. | Meets observable physical needs in socially appropriate ways | 0.6 |
| | | C1.1 | C1.1 | Initiates social behavior toward peer | 0.07 |
| | | C1.3 | C1.3 | Plays near one or two peers | -0.6 |
| **Adaptive (32)** | 8 | A1.4 | A1.2 | Swallows liquids | 2.78 |
| | | A4.2 | A3.2 | Eats with fingers | 1.26 |
| | | B2.0 | B2.2 | Washes and dries hands | -0.31 |
| | | C1.3 | C1.4 | Takes off pants | -0.48 |
| | | C1.5 | C1.5 | Takes off shoes | -0.41 |
| | | C1.4 | C1.6 | Takes off socks | 0.08 |
| | | C1.6 | C1.7 | Takes off hat | 0.82 |
| | | C1.2 | C1.3 | Takes off front-fastened coat, jacket, or shirt | 0.04 |

### 4.3.2 Common Item (anchor) Selection: Functional Matching

Since only one identical common item existed in the cognitive area, the functional match strategy is employed to expand the common item pool. Functional matching is utilized to identify common items when the items on two forms are similar in nature but not identical. This matching process involves comparing the content and level of difficulty in the current scenario to identify common items. This process also involves a certain degree of uncertainty and approximation in judging item similarity. Serving the same function as traditional common item selection, the functional match process helps ensure that the common items are similar enough to be considered as measuring the same construct on both forms.

Functional matching can be performed using a variety of similarity metrics, such as cosine similarity, Jaccard Similarity, Euclidean Distance, etc. Cosine similarity is a widely used measure in numerous tasks, including semantic similarity, document representation, and document distance calculation (Kusner, Sun, Kolkin, & Weinberger, 2015; Le & Mikolov, 2014; Mihalcea, Corley, & Strapparava, 2006). In this study, cosine similarity was applied. The SpaCy package was specifically used to compute the similarity. SpaCy is an open-source NLP library that offers a range of features, such as tokenization, part-of-speech tagging, and similarity scoring, among others. In this case, SpaCy was employed to determine the cosine similarity between AEPS-2 and AEPS-3 item descriptions. The script (refer to the figure below) calculates the similarity between the two documents using the built-in word embeddings available in the medium-sized English language model (en_core_web_md). The similarity score ranges from 0 to 1, with higher values indicating greater similarity.

After functional matching, items with similarity scores exceeding 0.9 were included in the item pool. Even though a common approach is to set a threshold value, such as 70% or 80%, as the minimum similarity score for two items to be considered a match, the criteria for items with similarity scores when using cosine similarity for functional matching often depend on the specific use case and the desired level of similarity. (Schütze, Manning, & Raghavan, 2008, page 419). Once an adequate number of common items was established, a forward deletion process was implemented to select high-quality anchor items with a displacement less than 0.5. The final version of the anchor items can be found in Table 11. The functional matching process for item anchors was not employed in the gross motor domain as the proportion of identical anchor items exceeded 30%.

Table 11 Functional matching Anchor items were used in Fixed calibration equating (anchoring)

| Area | Nu. of item | AEPS-2+ | AEPS-3+ | Similarity | displace |
|---|---|---|---|---|---|
| **Fine Motor (32)** | 6 | A1.1 | A1 | 0.96 | 0.41 |
| | | A5.1 | B3.2 | 1 | -0.36 |
| | | B2.1 | B3.3 | 1 | -0.08 |
| | | B5.2 | C1.4 | 1 | 0.16 |
| | | A3.3 | A2.4 | 0.94 | 0.22 |
| | | A4.3 | A2.3 | 0.93 | 0.25 |
| **Gross Motor (55)** | 30 | A1.3 | A1.2 | 1 | -0.29 |
| | | A2.1 | A3.1 | 1 | -0.48 |
| | | A3.1 | B1.1 | 0.98 | -0.05 |
| | | A3.2 | B1.2 | 1 | -0.02 |
| | | A3.3 | B1.3 | 1 | -0.06 |
| | | A3.5 | A2.0 | 0.96 | 0.06 |
| | | B1. | A4 | 1 | 0.19 |
| | | B1.2 | A4.2 | 0.95 | -0.03 |
| | | B1.4 | A4.4 | 1 | -0.05 |
| | | B1.5 | A4.5 | 1 | -0.22 |
| | | B2.1 | A5.1 | 1 | -0.23 |
| | | B2.2 | A5.2 | 0.98 | 0.50 |
| | | C1.4 | B2.2 | 1 | -0.07 |
| | | C1.5 | B3.4 | 1 | -0.11 |
| | | C2.0 | B2 | 0.97 | 0.35 |

| | | | | | |
|---|---|---|---|---|---|
| | | C2.1 | B2.1 | 0.99 | 0 |
| | | C2.2 | B2.3 | 1 | 0.08 |
| | | C2.3 | B2.4 | 1 | 0.12 |
| | | C1.1 | B3.1 | 1 | -0.01 |
| | | C1.2 | B3.2 | 1 | -0.14 |
| | | C1.3 | B3.3 | 1 | -0.03 |
| | | C1.5 | B3.4 | 1 | -0.13 |
| | | C3.1 | B5.1 | 1 | -0.19 |
| | | C3.2 | B5.2 | 1 | -0.07 |
| | | C4.2 | B2.1 | 1 | .42 |
| | | D2.0 | C3.2 | 1 | -0.18 |
| | | D2.2 | C3.4 | 0.94 | -0.10 |
| | | D1.0 | B6 | 1 | 0.14 |
| | | D1.2 | B6.2 | 0.97 | 0.01 |
| | | A3.0 | B1 | 1 | -0.16 |
| **Cognitive*** **(58)** | 5 | D1.0 | B1.1 | 0.94 | 0.21 |
| | | B2.0 | C1.1 | 0.97 | -0.43 |
| | | B3.0 | C1.0 | 0.99 | -0.03 |
| | | E2.0 | D1.0 | 0.98 | 0.19 |
| | | E2.1 | D1.1 | 0.97 | 0.16 |
| **Social Communication (46)** | 5 | B2.1 | C1.4 | 1 | 0.12 |
| | | C2.3 | B3.3 | 0.93 | -0.01 |
| | | A2.0 | B1.0 | 0.97 | 0.17 |
| | | C1.5 | A1.1 | 1 | 0.16 |
| | | C2.2 | B3.2 | 0.94 | -0.15 |
| **Social (25)** | 4 | A1.2 | A1.1 | 1 | 0.01 |
| | | C1.0 | C1 | 0.94 | -0.06 |
| | | C1.1 | C1.1 | 1 | 0.01 |
| | | C1.2 | C1.2 | 0.95 | 0.11 |
| | | B1.1 | | 1 | 0.37 |
| **Adaptive (32)** | 8 | A4.1 | A3.1 | 0.92 | 0.22 |
| | | A2.0 | A2.1 | 0.91 | 0.19 |
| | | C1.2 | C1.3 | 1 | 0.28 |
| | | B1.2 | B1.3 | 0.96 | -0.03 |
| | | B2.0 | B2.2 | 1 | 0 |
| | | C1.3 | C1.4 | 1 | -0.24 |
| | | C1.5 | C1.5 | 1 | -0.43 |
| | | C1.4 | C1.6 | 1 | 0.17 |

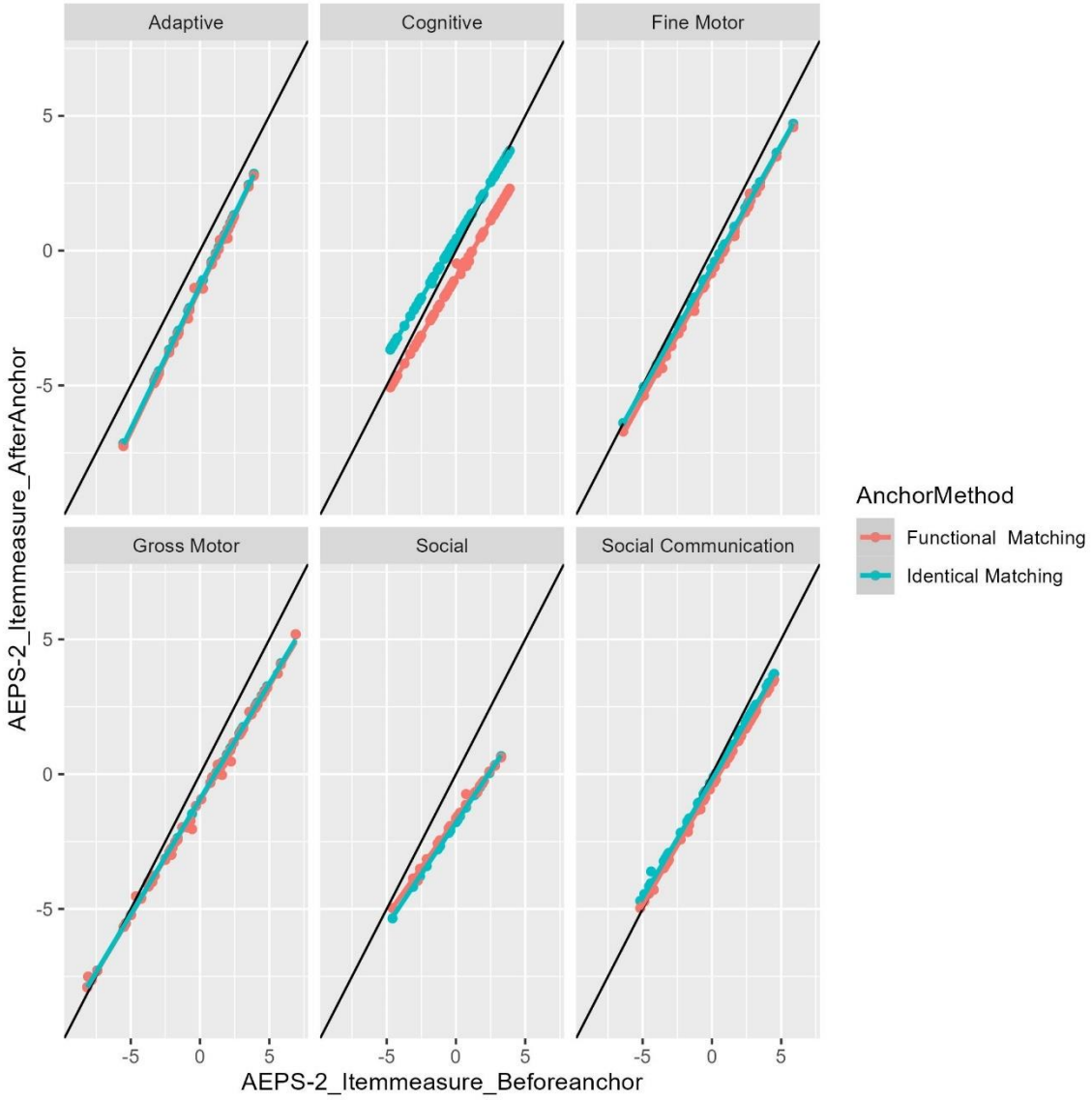
Note: * In the cognitive domain, only functional matching anchor item designs were conducted. + Descriptors of items can be found in Appendix 3 and 4.

### 4.3.3 Item Difficult Estimation Comparison before and after Anchoring

After establishing the two sets of anchor items, we conducted a fixed parameter calibration. This provided estimations of item difficulty and person ability. To compare the item difficulty parameters before and after anchoring, please refer to Figure 4. As I

employed the Rasch framework for equating, the model remains invariant regardless of the context, time period, or sample, as long as the measured construct is consistent. The ideal correspondence of the measure scores between AEPS-2 and AEPS-3 should align with the solid black line, indicating that the item parameters scale after anchoring without drifting, and the item difficulty measure before and after anchoring exhibits a linear relationship. If the line lies above the identity line, it implies that the item difficulty measure before anchoring is lower than it is after anchoring. Conversely, if the line falls below the ideal line, it indicates that the item difficulty measure before anchoring is higher than after anchoring. In the fine motor, social communication, and social emotion domains, the item difficulty measures before and after anchoring exhibit similar trends for both nearly identical and functional anchor designs. In the social communication domain, the item difficulty measures before and after anchoring align ideally, indicating no adjustment to measures after anchoring. In the adaptive and social emotion domains, the measures for items with lower difficulties increased after anchoring, while those for items with higher difficulties decreased. In the cognitive domain, since there was only one identical item in the anchor set, there are noticeable differences between the item difficulty measures from the two anchor designs.

Figure 4. the comparison of item difficult measures (identical matching anchoring vs. functional matching anchoring)




### 4.3.4 Conversion Relationship between AEPS-2 and AEPS-3

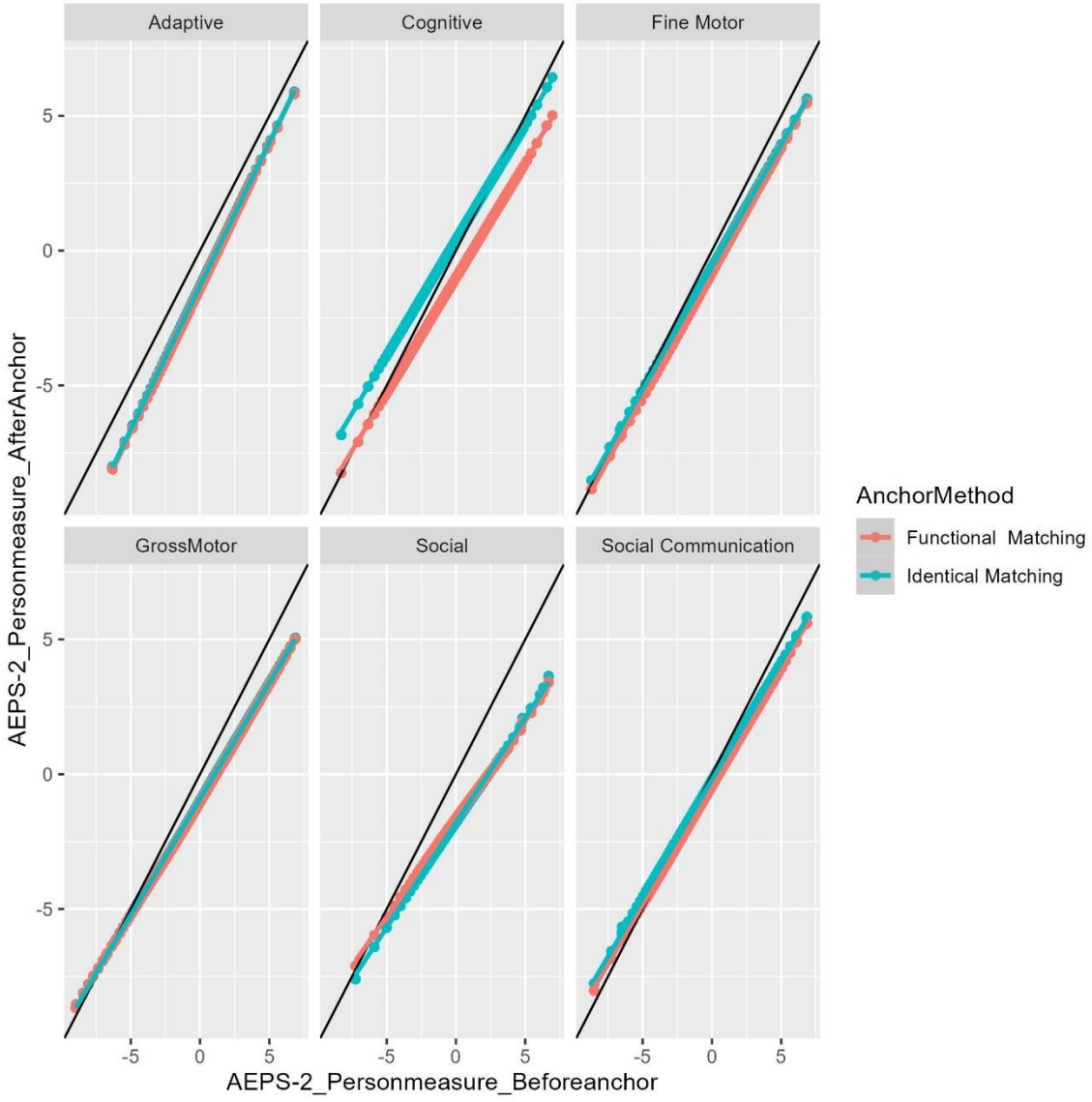
Upon completing the study, both the person ability measure conversion table and the observed score conversion table for person ability were made available. Equating took place in each developmental domain. Initially, the person's estimated ability measure was calculated using the AEPS-2 form. Subsequently, the estimated true score for the AEPS-2

items on the AEPS-3 scale was computed for the same person’s ability, establishing a person ability measure conversion relationship between the two versions. Due to different matching methods, two sets of anchor items were chosen for all domains. This led to the creation of two conversion tables for six domains. The children's person ability measures before and after anchoring are summarized and shown in Figure 5. The solid black line represents the line of perfect alignment, whereas the blue line illustrates the best-fitting regression line for the data. For more information about the person ability measure score and conversion table, please refer to Appendix 7.

The conversion relationship between person ability measures across the two versions mirrors the trend seen in item difficulty conversion. The ideal correspondence of the scores between AEPS-2 and AEPS-3 should align with the solid black line, illustrating that item parameters scale consistently after anchoring without drift, and the person's ability measure maintains a linear relationship before and after anchoring. If the line situates above the identity line, it suggests that the measure of person's ability was lower before anchoring than after. Conversely, if the line is below the ideal line, it indicates a higher measure of person's ability before anchoring than after. For the domains of fine motor, social communication, and social emotion, the person's measures display similar trends before and after anchoring across both nearly identical and functional anchor designs. Within the social communication domain, the measures of person's ability before and after anchoring align perfectly, signifying no adjustments to measures after anchoring. For the adaptive and social emotion domains, the measures for children with lower abilities increased post-anchoring, while those for children with higher abilities decreased. In the cognitive domain, owing to the presence of only one identical item in

the anchor set, noticeable differences emerge between the measures of person's ability from the two anchor designs.

Figure 5 The comparison of person ability measures (identical matching anchoring vs. functional matching anchoring)



Following the procedure for true score transformation, person ability measure scores from AEPS-2 were obtained on the AEPS-3 scale using fixed parameter calibration across all six areas: fine motor, gross motor, social communication, social emotion, cognitive, and adaptive. Previous results showed that both identical matching common item anchoring and functional matching common item anchoring yielded similar results in the fine motor, adaptive, social emotion, and social communication areas. Therefore, identical matching common item anchoring was used for fixed calibration in these four areas. In the gross motor area, only identical matching common item anchoring was conducted due to the high number of identical common items. In the cognitive domain, where only one identical common item existed, the functional matching anchor design was employed for fixed parameter calibration. After this, a conversion table for score equivalence was developed by comparing the values of matching measures between the observed scores of AEPS-2 and AEPS-3. For instance, after anchoring, if the observed score "2" from AEPS-2 has the same value of measure as the observed score "3" from AEPS-3, the conversion table will indicate that the observed score "2" from AEPS-2 is equivalent to the observed score "3" from AEPS-3. For further details, please refer to Table 12 provided below.

Table 12 Recommended Convert Table in Six Developmental Areas for Implementation

| **AREA** | **ADAPTIVE** | | **FINE MOTOR** | | **SOCIAL COMMUNICATION** | | **SOCIAL EMOTION** | | **COGNITIVE** | | **GROSS MOTOR** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **VERSION** | AEPS-2 | AEPS-3 | AEPS-2 | AEPS-3 | AEPS-2 | AEPS-3 | AEPS-2 | AEPS-3 | AEPS-2 | AEPS-3 | AEPS-2 | AEPS-3 |
| **SCORE** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 1 | 1 | 1 | 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | 2 | 2 | 2 | 7 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 2 |
| | 3 | 5 | 3 | 9 | 3 | 1 | 3 | 2 | 3 | 1 | 3 | 3 |
| | 4 | 6 | 4 | 11 | 4 | 2 | 4 | 3 | 4 | 2 | 4 | 4 |
| | 5 | 7 | 5 | 12 | 5 | 2 | 5 | 4 | 5 | 2 | 5 | 5 |
| | 6 | 8 | 6 | 13 | 6 | 3 | 6 | 5 | 6 | 3 | 6 | 6 |
| | 7 | 9 | 7 | 15 | 7 | 3 | 7 | 6 | 7 | 3 | 7 | 7 |
| | 8 | 9 | 8 | 16 | 8 | 4 | 8 | 7 | 8 | 4 | 8 | 8 |
| | 9 | 10 | 9 | 17 | 9 | 5 | 9 | 8 | 9 | 4 | 9 | 9 |
| | 10 | 11 | 12 | 21 | 10 | 5 | 10 | 9 | 10 | 4 | 10 | 10 |
| | 11 | 12 | 13 | 22 | 11 | 6 | 11 | 10 | 11 | 5 | 11 | 11 |
| | 12 | 13 | 14 | 24 | 12 | 7 | 12 | 11 | 12 | 5 | 12 | 12 |
| | 13 | 14 | 15 | 25 | 13 | 8 | 13 | 12 | 13 | 6 | 13 | 13 |
| | 14 | 14 | 16 | 26 | 14 | 9 | 14 | 13 | 14 | 6 | 14 | 14 |
| | 15 | 15 | 17 | 27 | 15 | 9 | 15 | 14 | 15 | 7 | 15 | 15 |
| | 16 | 15 | 18 | 28 | 16 | 10 | 16 | 15 | 16 | 8 | 16 | 16 |
| | 17 | 16 | 19 | 30 | 17 | 12 | 17 | 16 | 17 | 8 | 17 | 17 |
| | 18 | 17 | 20 | 31 | 18 | 13 | 18 | 17 | 18 | 9 | 18 | 18 |
| | 19 | 18 | 21 | 32 | 19 | 14 | 19 | 19 | 19 | 9 | 19 | 19 |
| | 20 | 19 | 22 | 34 | 20 | 15 | 20 | 20 | 20 | 10 | 20 | 20 |
| | 21 | 20 | 23 | 36 | 21 | 16 | 21 | 21 | 21 | 11 | 21 | 21 |
| | 22 | 20 | 24 | 37 | 22 | 17 | 22 | 22 | 22 | 12 | 22 | 22 |
| | 23 | 21 | 25 | 38 | 23 | 19 | 23 | 24 | 23 | 13 | 23 | 23 |
| | 24 | 22 | 26 | 40 | 24 | 20 | 24 | 25 | 24 | 14 | 24 | 24 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 25 | 23 | 27 | 41 | 25 | 21 | 25 | 26 | 25 | 15 | 25 | 25 |
| 26 | 24 | 28 | 43 | 26 | 22 | 26 | 28 | 26 | 16 | 26 | 26 |
| 27 | 25 | 29 | 45 | 27 | 24 | 27 | 29 | 27 | 16 | 27 | 27 |
| 28 | 26 | 30 | 46 | 28 | 25 | 28 | 31 | 28 | 17 | 28 | 28 |
| 29 | 26 | 31 | 48 | 29 | 27 | 29 | 32 | 29 | 18 | 29 | 29 |
| 30 | 27 | 32 | 49 | 30 | 28 | 30 | 34 | 30 | 19 | 30 | 30 |
| 31 | 28 | 33 | 51 | 31 | 29 | 31 | 35 | 31 | 20 | 31 | 31 |
| 32 | 29 | 34 | 53 | 32 | 31 | 32 | 37 | 32 | 21 | 32 | 32 |
| 33 | 30 | 35 | 55 | 33 | 32 | 33 | 39 | 33 | 23 | 33 | 33 |
| 34 | 30 | 36 | 56 | 34 | 33 | 34 | 41 | 34 | 23 | 34 | 34 |
| 35 | 31 | 37 | 58 | 35 | 35 | 35 | 43 | 35 | 24 | 35 | 35 |
| 36 | 32 | 38 | 60 | 36 | 36 | 36 | 45 | 36 | 26 | 36 | 36 |
| 37 | 33 | 39 | 62 | 37 | 38 | 37 | 47 | 37 | 27 | 37 | 37 |
| 38 | 34 | 40 | 64 | 38 | 39 | 38 | 50 | 38 | 28 | 38 | 38 |
| 39 | 34 | 41 | 65 | 39 | 40 | 39 | 53 | 39 | 29 | 39 | 39 |
| 40 | 35 | 42 | 67 | 40 | 42 | 40 | 56 | 40 | 30 | 40 | 40 |
| 41 | 36 | 43 | 70 | 41 | 43 | 41 | 59 | 41 | 31 | 41 | 41 |
| 42 | 37 | 44 | 71 | 42 | 44 | 42 | 62 | 42 | 32 | 42 | 42 |
| 43 | 38 | 45 | 74 | 43 | 46 | 43 | 66 | 43 | 33 | 43 | 43 |
| 44 | 38 | 46 | 75 | 44 | 47 | 44 | 70 | 44 | 35 | 44 | 44 |
| 45 | 39 | 47 | 78 | 45 | 49 | 45 | 75 | 45 | 36 | 45 | 45 |
| 46 | 40 | 48 | 80 | 46 | 50 | 46 | 81 | 46 | 37 | 46 | 46 |
| 47 | 41 | 49 | 82 | 47 | 52 | 47 | 88 | 47 | 38 | 47 | 47 |
| 48 | 42 | 50 | 84 | 48 | 53 | 48 | 96 | 48 | 39 | 48 | 48 |
| 49 | 43 | 51 | 87 | 49 | 55 | 49 | 107 | 49 | 40 | 49 | 49 |
| 50 | 44 | 52 | 89 | 50 | 56 | 50 | 124 | 50 | 41 | 50 | 50 |
| 51 | 45 | 53 | 91 | 51 | 58 | | | 51 | 43 | 51 | 51 |
| 52 | 47 | 54 | 93 | 52 | 59 | | | 52 | 44 | 52 | 52 |
| 53 | 48 | 55 | 95 | 53 | 61 | | | 53 | 45 | 53 | 53 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 54 | 49 | 56 | 97 | 54 | 63 | 54 | 46 | 54 | 54 |
| 55 | 50 | 57 | 99 | 55 | 64 | 55 | 47 | 55 | 55 |
| 56 | 52 | 58 | 100 | 56 | 66 | 56 | 48 | 56 | 56 |
| 57 | 54 | 59 | 102 | 57 | 68 | 57 | 50 | 57 | 57 |
| 58 | 56 | 60 | 103 | 58 | 69 | 58 | 51 | 58 | 58 |
| 59 | 58 | 61 | 105 | 59 | 71 | 59 | 52 | 59 | 59 |
| 60 | 61 | 62 | 106 | 60 | 72 | 60 | 53 | 60 | 60 |
| 61 | 64 | | | 61 | 74 | 61 | 54 | 61 | 61 |
| 62 | 69 | | | 62 | 76 | 62 | 56 | 62 | 62 |
| 63 | 78 | | | 63 | 77 | 63 | 57 | 63 | 63 |
| 64 | 106 | | | 64 | 79 | 64 | 58 | 64 | 63 |
| | | | | 65 | 80 | 65 | 59 | 65 | 64 |
| | | | | 66 | 81 | 66 | 60 | 66 | 65 |
| | | | | 67 | 83 | 67 | 61 | 67 | 66 |
| | | | | 68 | 84 | 68 | 63 | 68 | 67 |
| | | | | 69 | 85 | 69 | 64 | 69 | 68 |
| | | | | 70 | 86 | 70 | 65 | 70 | 69 |
| | | | | 71 | 87 | 71 | 66 | 71 | 70 |
| | | | | 72 | 88 | 72 | 67 | 72 | 71 |
| | | | | 73 | 89 | 73 | 68 | 73 | 72 |
| | | | | 74 | 90 | 74 | 69 | 74 | 73 |
| | | | | 75 | 91 | 75 | 71 | 75 | 73 |
| | | | | 76 | 92 | 76 | 71 | 76 | 74 |
| | | | | 77 | 92 | 77 | 73 | 77 | 75 |
| | | | | 78 | 93 | 78 | 74 | 78 | 76 |
| | | | | 79 | 94 | 79 | 75 | 79 | 77 |
| | | | | 80 | 94 | 80 | 76 | 80 | 78 |
| | | | | 81 | 95 | 81 | 77 | 81 | 79 |
| | | | | 82 | 95 | 82 | 78 | 82 | 79 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 83 | 95 | 83 | 79 | 83 | 80 |
| 84 | 96 | 84 | 80 | 84 | 81 |
| 85 | 96 | 85 | 80 | 85 | 82 |
| 86 | 96 | 86 | 81 | 86 | 83 |
| 87 | 97 | 87 | 82 | 87 | 84 |
| 88 | 97 | 88 | 83 | 88 | 85 |
| 89 | 97 | 89 | 84 | 89 | 86 |
| 90 | 98 | 90 | 85 | 90 | 87 |
| 91 | 98 | 91 | 86 | 91 | 88 |
| 92 | 98 | 92 | 86 | 92 | 89 |
| | | 93 | 87 | 93 | 90 |
| | | 94 | 88 | 94 | 91 |
| | | 95 | 88 | 95 | 92 |
| | | 96 | 89 | 96 | 93 |
| | | 97 | 90 | 97 | 95 |
| | | 98 | 90 | 98 | 96 |
| | | 99 | 91 | 99 | 97 |
| | | 100 | 92 | 100 | 99 |
| | | 101 | 92 | 101 | 100 |
| | | 102 | 93 | 102 | 102 |
| | | 103 | 93 | 103 | 103 |
| | | 104 | 94 | 104 | 105 |
| | | 105 | 94 | 105 | 107 |
| | | 106 | 95 | 106 | 110 |
| | | 107 | 95 | 107 | 113 |
| | | 108 | 96 | 108 | 116 |
| | | 109 | 96 | 109 | 121 |
| | | 110 | 97 | 110 | 130 |
| | | 111 | 97 | | |

| | | |

## 4.4 Chapter Summary

Chapter four presents the results of the score equating study. Firstly, the assessment of scale quality is evaluated using a sample of 18,531 cases from the AEPS-2 Test Level I and 939 cases from the AEPS-3 Test. The psychometric properties of both assessment versions undergo meticulous evaluation, revealing a good fit between the model and the data. Through the utilization of the cosine similarity coefficient and expert content check, identical and functional matching anchor item sets are carefully selected, showcasing satisfactory quality. The investigation further explores the impact of different anchor sets on the estimation of person parameters during the anchoring process. As a result, person measure and observed score conversion tables between AEPS-2 and AEPS-3 are generated, providing valuable insights into the correlation between the older and updated versions of the assessment.

# CHAPTER 5 DISCUSSION AND CONCLUSION

## 5.1 Discussion

The results of this study are discussed in three distinct sections, each addressing one of the research questions. These sections cover the outcomes of the calibration process, the selection of common items, and the creation of a conversion table between AEPS-2 and AEPS-3.

### 5.1.1 Discussion of Research Question One

**1. To what extent do AEPS-2 and AEPS-3 instruments fit the Rasch Rating Scale Model?**

In the scale calibration, the study examined the dimensionality, model-data fit, and reliability of the scale. The evidence presented in Chapter Four supports the application of the Rasch model across various developmental areas, such as the fine motor, gross motor, cognitive, social, social communication, and adaptive areas.

Unidimensionality is a fundamental assumption of the Rasch model, but it is a debated issue in practical applications. According to Wright and Linacre (1989), unidimensionality is more of a conceptual rather than factual or quantitative concept, and no test can be perfectly unidimensional. In this study, a group of experts defined the latent trait in each area based on items that followed children's developmental sequence before conducting statistical analysis. measures considered more than adequate" to "measures, which is considered more than adequate. Furthermore, if a secondary

dimension has an eigenvalue of less than three and accounts for less than 5% of the unexplained variance, unidimensionality is plausible (Linacre, 2009).

The Wright map and infit/outfit statistics provided valuable insights into the effectiveness of the AEPS items in targeting children's developmental skills. The results revealed a gap in item difficulty at the highest and lowest ability levels across all six developmental areas, which is consistent with previous studies (Winchell, 2011; Toland et al., 2021). Additionally, the analysis of AEPS-3 data indicated the presence of outlier easy items, which may be explained by the sample size distribution. Specifically, the lack of extreme age groups in the sample may have limited the ability to detect data fit issues in the Rasch model. Therefore, future studies should consider including a more diverse sample with extreme ability levels to ensure the assessment quality based on the Rasch framework.

In the fit statistics, the Z-value brings the attention as the high value over 4. Specifically, when mean-square fit statistics are close to 1.0, the associated Z-values are very large (over 4 to 9.9) due to the large sample size (18,411 children in the AEPS-2 dataset and 317 children in the AEPS-3 dataset), which gives the study high statistical power to test the null hypothesis of exact model fit. The large sample size in the study raises questions about the interpretation of fit statistics. However, the Rasch model assumes that data fit is useful rather than perfect, and empirical observations may not perfectly align with the ideal Rasch model when the sample size is large. Therefore, the null hypothesis of exact model fit is typically rejected in these cases. The conclusion to research question one is the AEPS-2 and AEPS-3 instruments fit the Rasch Rating Scale

Model with minor model misfit, which relates to the third research question, discussed in section 5.1.3.

### 5.1.2 Discussion of Research Question Two

**2. What is the most efficient set of the common items in the six developmental areas, respectively, for the purpose of equating across two measures?**

Selecting anchor items is an important step in determining whether the item parameters from two independent Rasch calibrations are invariant (Smith, 1996). The Rasch model is theoretically considered an invariant measurement model, because it meets specific requirements that ensure stable and consistent comparisons across individuals and items, regardless of the particular sample being analyzed. Ideally, anchor items should have the same descriptions and invariant item parameters. However, in experimental settings, item drift may occur in anchor items over time due to factors such as context effects, overexposure, or curriculum changes. Item drift can compromise equating accuracy and undermine score interpretation in equating practice, as it is reflected in the displaced parameter. To address this issue, items that have a displacement parameter greater than 0.5 are typically removed from anchor item sets during the functional matching anchor item selection process before conducting equating (Donoghue & Isham, 1998; He et al., 2013; Hu et al., 2008; Huang & Shyu, 2003).

#### *5.1.2.1 Anchor item selection issue in identical matching*

Kolen and Brennan (2014, p. 287) recommend that the anchor item set should be constructed to have the same content and statistical specifications as the total test to ensure that the anchor items adequately reflect group differences. Moreover, Linacre

(2004) suggests that the number of anchor items should be at least 20% of the original items. In Wang's (2004) study, it was found that four anchor items are sufficient for detecting item drift. However, after exact matching in the six developmental areas, there are less than four invariant items in the fine motor and cognitive areas. The fine motor and cognitive areas have only three and one identical item, respectively, which may compromise the accuracy of equating and result in underestimation or overestimation of item parameters in subsequent equating steps.

#### *5.1.2.2 Reason to Conduct Functional Matching Method*

The functional matching on the common item gives more flexibility and the possibility of score equating between assessments when there are not enough identical items. Due to the difference between the assessment and the test, language modification has always been part of the different versions based on age changes or knowledge developments. Traditionally, equating studies use the same items in non-equivalent common item designs between different versions of tests. This study provides an example of employing functional matching common items in the scale transformation score equating study between different assessments. Developing an assessment is a long-term process. For instance, early childhood education experts spent almost two decades developing the AEPS-3, with concepts and focus changing according to times. Using functional matching to select common items is an additional applicable method. Moreover, all the items in the assessment were designed as functional items, which means the main goal is to assess children's functions. Even with slight differences in the description, two functionally identical items still measure the same skill.

The functional matching anchor item selection method may require adjustments to precalibrated item parameters to meet non-drift common item selection criteria. This adjustment cannot be done based on their precalibrated values and may impact metric conversion, especially if a large number of precalibrated items need to be adjusted. Ye and Xin (2014) found that as item parameter drift increases, recovery results decrease, and achievement estimates deviate significantly from true values at 0.5 logits of drift in the fixed parameter calibration with the common item design in the vertical linking. In this study, the functional anchor item selection was limited to items with drift amounts of 0.5 logits or less. However, it remains unclear how potential anchor selection issues such as selection order or displacement magnitudes impact the production of robust linking results using the fixed parameter calibration method, and further research is needed.

#### *5.1.2.3 Item parameter drift in the anchoring*

In the context of the Rasch model, item drift occurs when an item's difficulty parameter changes over different test administrations or across different groups. This can be problematic when using anchoring techniques, especially when selecting common items to link different test forms or administrations.

When selecting common items for anchoring (also known as anchor items), these items are presumed to function the same way across different forms or administrations of the test. They are chosen because of their stability in terms of difficulty and discrimination parameters.

However, if an anchor item experiences item drift, its performance characteristics change, undermining the assumption of stability and potentially distorting the linking or

equating process. Therefore, it's crucial to monitor for item drift when using anchor items to ensure the comparability of scores across different test forms or administrations. If item drift is detected, the item may need to be dropped as an anchor item, and the linking or equating process may need to be adjusted accordingly.

#### *5.1.2.4 Model Data Misfit & Limited Anchor Item Pool*

According to Fischer and his colleagues' (2021) findings, the anchor item pool in this study was limited and model data misfit existed. They suggest that the choice of linking method is not as crucial when linking Rasch modeled data, regardless of the presence or absence of (moderate) model misfit. Rather, it is important for practitioners to be aware that a combination of moderate model misfit and certain factors such as the empirical relation of anchor item difficulty parameters and anchor item discrimination parameters, composition of anchor items, person-item fit, and sample size may lead to a distorted parameter estimation. However, there are currently no applicable diagnostics or concrete guidelines for empirical data available.

Results from Zhao and Hambleton's (2017) large-scale assessment indicate that the consequences of model misfit varied depending on the choice of model and IRT scaling methods. When compared to mean/sigma (MS) and Stocking and Lord characteristic curve (SL) methods (2017), the separate calibration with linking and fixed common item parameter (FCIP) procedure was more sensitive to model misfit and more robust against various amounts of ability shifts between two adjacent administrations, regardless of model fit.

### 5.1.3 Discussion of Research Question Three

**3. How adequate was the fixed parameter calibration, in terms of the accuracy of equating?**

It's important to note that the complex nature of children's development makes accurate and comprehensive assessment challenging. A child's developmental and learning pace can be uneven, and children might demonstrate their knowledge and skills differently in different contexts. This underlines the need for assessment methods that are not only consistent and equatable but also sensitive to the child's developmental stage, cultural background, language proficiency, and individual characteristics. Effective assessment practices should be developmentally, culturally, and linguistically responsive to authentically assess children's learning.

Reflecting on the results of the equating in this study, the number of items in the social-emotion area has notably increased. This indicates society's intensifying focus on this domain, especially the nuanced aspects of young children's social emotions. In the AEPS-3's cognitive area, the target group is children aged 0-6. Although this study's sample comprises children aged 0-3, the revised scale design incorporates the developmental traits of children aged 3-6. These adaptations are vital, considering the swift changes observed in children's cognitive development within the cognitive and social-emotional domains. Nonetheless, it's essential to acknowledge that adjustments between the two versions of the assessment, driven by cultural shifts over time, could potentially influence these scales. Such transitions could lead to discrepancies in the outcomes, which emphasizes the need for continuous recalibration. This situation also

highlights the importance of applying a functional matching method when selecting anchor items.

Cultural considerations in the design of assessment item content require an understanding that culture, language, and societal norms are dynamic and ever evolving. It's imperative to acknowledge that societal attitudes and cultural norms are not stagnant but transform over time. Thus, regular updates to the assessment content are necessary to mirror these changes, ensuring relevance and the avoidance of perpetuating outdated stereotypes or norms. Parallelly, language usage undergoes constant evolution. Consequently, it's pivotal to guarantee that the language used in test items aligns with current usage, and obsolete terminology and phrases, or those that have morphed in meaning, are updated to maintain accuracy and clarity. Technological progression is another aspect of cultural change that needs to be taken into account. As technology integrates more deeply into various cultures, assessment items must reflect these modifications. This includes updating references to outdated technology to uphold the items' relevance and relatability.

Moreover, the societal landscape is often shaped by current events, which can significantly influence cultural attitudes and experiences. Being cognizant of these events and their impact when designing assessment items can help maintain the assessment's cultural sensitivity and accuracy. Additionally, as educational standards and curricula adapt and evolve, alignment of the assessment items with current teaching practices and learning goals becomes essential to maintain their educational validity.

Cultural diversity is another significant factor. Over time, societies often grow more diverse, with changes in demographics and increasing cultural intermingling. This

increased diversity may need to be mirrored in the assessment items to ensure cultural inclusivity and relevance.

In conclusion, the dynamism of societal and cultural changes necessitates ongoing review and revisions of assessment items. This continual refinement is vital to ensure the cultural relevance, fairness, and overall validity of the assessments over time.

### 5.1.4 Discussion of Research Question Four

**4. What score conversion table will be provided on the six developmental areas from AEPS-2 to AEPS -3?**

In this study, two score conversion tables were applied: the person ability measure score conversion table and the observed score conversion table. The person ability measure conversion offers a practical and accurate method for exchanging participants' scores from a psychometric perspective. However, interpreting the person ability measure conversion table requires early childhood education professionals to have a basic understanding of the Rasch model, which may lead to confusion or necessitate additional training. On the other hand, the observed score conversion table displays the relationship between the raw scores of the two versions and is easier to comprehend for individuals without psychometric training. Additionally, this study provided conversion tables with identical anchor items or functional matching anchor items in six child developmental areas, revealing informative true score equating results (further details are provided below).

#### *5.1.4.1Conversion table comparison*

In this study, two distinct anchor sets were utilized to obtain item parameters. Comparing the conversion tables between identical anchor items and functional matching anchor items in six child developmental areas revealed informative true score equating results that warrant further investigation into their underlying mathematical rationale. There are three different scale transformation scenarios:

Firstly, identical anchor items and functional matching anchor items yield nearly identical linear parameter transformations from the AEPS-2 scale to the AEPS-3 scale (e.g., social communication area, social emotion area, and fine motor area in Figure 4). In these three areas, both anchor sets have similar anchor item numbers and locations. Even with a few items exhibiting high displacement in the identical anchor item set, the impact on item parameters is minimal. Secondly, when the number of anchor items is considerably low, the results are more likely to underestimate a person's ability compared to identical anchoring. For instance, in the cognitive area, there is only one identical anchor item, which is insufficient for equating. In Figure 4, the item parameter estimated using one-item anchoring is lower than that estimated using five-item anchoring, which still falls short of the 20% ratio criteria. Expanding the anchor item pool to 80% similarity through functional matching or adding new identical items could be a future direction to explore for creating a more accurate score conversion table in the cognitive area. Thirdly, when item locations are primarily focused in the middle and upper parts of the scale, differences in the estimation of lower person abilities are noticeable between the two methods (e.g., adaptive area in Figure 4). As AEPS is an assessment also suitable for children with developmental delays, potentially inaccurate lower person abilities could cause issues in eligibility decision-making. This problem also emerges in the

observed score conversion table. There is a substantial gap between the highest score and the second-highest score in the conversion table of the adaptive area, which necessitates further evidence to evaluate the equating results.

These scenarios offer valuable insights into the effects of anchor item selection and location on the equating process, emphasizing the need for appropriate anchor item choices to achieve accurate scale transformation during score equating.

*5.1.4.2Impact of the Coding Scheme and Missing Data on Equating*

To address the third research question, the approach to handling missing data is discussed here. The strategy employed in this study to address missing data and zero values in the original dataset involved excluding participants with missing data. This was done to minimize potential biases in the equating results at the upper and lower ends of the item continuum. Based on Shin (2009), omitted responses in equating data should be treated with caution as they can be wrong or not administered. The findings suggest that one should leave omitted responses as missing and use a large sample size to ensure the accuracy of the screening tools during equating. AEPS as an assessment tool for the children with or without the developmental delay, the items on the upper end of the developmental continuum refer to higher item difficulty and advanced developmental skills for the children, the rater may skip the item when they observed children with low performance. This behavior creates systematical missing data in the original AEPS-2 and AEPS-3 dataset. Waterbury (2019) conducted simulation study to examine the impact of missing data mechanisms, sample size, test length, and proportions of missing data on the standard errors and biases of item parameters using the Rasch measurement model. Findings demonstrated that the item parameters were significantly biased when the

missing data existed as skip the item, particularly with higher proportions of missing responses (0.5). The results also indicated that standard errors were primarily affected by sample size, with larger sample sizes yielding smaller standard errors. Therefore, a large sample size is recommended when dealing with varying amounts of missing data in Rasch model-based analyses. The ample sample sizes of AEPS-2 and AEPS-3 in the present study help prevent the underestimation of item difficulty due to systematic missing data. Nevertheless, excluding all participants with missing data may overlook the possibility of randomized missing data or limit the representation of children with developmental delays. Therefore, further research is necessary to explore methods of handling missing data during equating.

*5.1.4.3Value of functional matching anchoring method in implication*

The Functional Matching Anchoring method expands the common item pool, increasing the chances of accurately transforming scores from different versions of an assessment onto a common scale. This process allows scores from diverse test forms to be directly compared or treated as though they originated from the same test form, even without identical item descriptions. This ensures fairness and validity in the assessment process.

Accommodating temporal changes requires a flexible approach, and this is where functionally similar items come into play. They are capable of incorporating shifts in focus and changes in constructs over time. For example, as the concept under evaluation in an assessment evolves or alters, these items can skillfully capture this progression. Consequently, this strategy enables successful equating across different test versions, regardless of the changes induced over time.(Heo & Squires, 2012) Common item

equating is a process that utilizes a set of items, known as the anchor test, that are shared between two different tests. When these common items share functional similarities, they can effectively align the mean item location of the common items. This procedure ultimately guarantees comparability across various test forms.

Thus, choosing functionally similar items for equating is an effective method to ensure that scores from different test forms are comparable, that changes over time are accommodated, and that fairness and validity are maintained in the assessment process.

## 5.2 Contribution and Implication

This study uses fixed parameter calibration (anchoring) for score equating to ensure the comparability of observed and true scores between AEPS-2 and AEPS-3 within each developmental domain. This allows for accurate comparisons of children's progress, helping educators, researchers, and other stakeholders make informed decisions and tailor interventions to support children's growth and development effectively. The AEPS, with its 30-year history, provides developmental information for thousands of children, emphasizing the need for long-term data to accurately evaluate children's abilities. The conversion table developed in this study ensures the generalizability of the assessment system and preserves longitudinal developmental information across different versions of the tool, allowing for consistent tracking of children's development.

This study addresses score accuracy during the AEPS transition period by offering a conversion table that ensures equitable decision-making regarding eligibility for services for children with developmental scores from different assessment versions. This highlights the importance of assessment accuracy in special education policy and serves as a model

for other evidence-based assessments in early childhood special education, emphasizing the need for consistency in evaluations and determining eligibility for appropriate services and interventions.

By closing the gap in score equating application in the early childhood assessment field, this study demonstrates how to implement score equating when assessment versions change over time. This contributes to the development of reliable and accurate assessment practices, ensuring consistent tracking and comparison of children's developmental progress across different assessment versions, ultimately enabling better decision-making and targeted interventions in ECE settings.

Nonetheless, the study also discovered that additional high-quality anchor items and data samples from both upper and lower developmental abilities are necessary. Further research is needed to enhance score accuracy and provide evidence for decision-making in the ECE and ECSE domains.

This study plays a vital role in addressing the assessment requirements in special education and linking educational policy with research goals. In terms of assessment requirements in ECSE, accurate and consistent assessment tools are essential. They ensure that children with special needs are properly identified and receive the appropriate interventions and supports. The AEPS score equating process allows for comparison of scores between different versions of the assessment (AEPS-2 and AEPS-3). This process maintains the accuracy of the assessment, crucial for evaluating children's progress and adjusting educational strategies as necessary.

To connect education policy with research goals, the equating process assists in implementing policies that advocate for consistent and valid assessment tools in early childhood special education. These policies often arise from research evidence suggesting the importance of early and accurate identification of developmental delays or disabilities. Also, the evidence derived from the AEPS score equating process can inform future education policies. By equating scores, the research is directly addressing a policy goal of maintaining high-quality, reliable assessments for children with special needs. This ensures that assessments remain useful tools for educators and specialists, ultimately helping achieve research goals of improving educational outcomes for this population.

In conclusion, The AEPS score equating process not only meets special education assessment requirements but also establishes a clear connection between educational policy and research goals. It fosters a cohesive approach where policy and research work in synergy to improve the quality of special education services.

### 5.3 Limitations & Future Research

This dissertation has several limitations. First, in the AEPS-2 assessment, each developmental area consists of multiple subsets of items. For instance, in the cognitive area for children aged 0-3 years old, there are seven subsets with 58 items. The item difficulty within each subset typically follows the order of child development. However, some items from different subsets are interwoven, suggesting that children's developmental processes in each area are not linearly independent. This finding supports the use of the Rasch model, rather than a multilevel model, for evaluating the six developmental areas. Nevertheless, the subset structure can be mathematically confusing,

and further research is required. Second, this study lacks concurrent child performance data, making it difficult to validate the impact of item drift on person ability estimation; further investigation on this topic is needed. Also, the content representation of the anchor item set was not taken into account when removing outlier anchor items in the functional matching anchor item selection process. Further research is needed to examine how the lack of content balance and the number of items in the anchor item set can significantly impact equating.

The final limitation of this study is that, while I am aware that early childhood educators collected the data from the AEPS, I lack information about these raters. As a result, I am currently unable to investigate the impact of rater severity on the scores. This limitation, however, opens up a promising avenue for future research. Delving into the influence of rater severity on scores could significantly enhance our understanding of the assessment process. Furthermore, incorporating score equating using the Many-Facet Rasch Model is an innovative addition to the family of score equating methods.

Future studies focusing on expanding the high-quality anchor item pool could enhance the accuracy of the conversion table. Further research could utilize fixed parameter calibration (anchoring) to obtain high-quality longitudinal data when developing new assessment versions. This may involve different content, age levels, and even target populations (e.g., parents, teachers) in ECE and ECSE.

## APPENDIX 1. WRIGHT MAP IN THE SIX DEVELOPMENTAL AREAS IN AEPS-2

Figure 1. Adaptive Area in AEPS-2

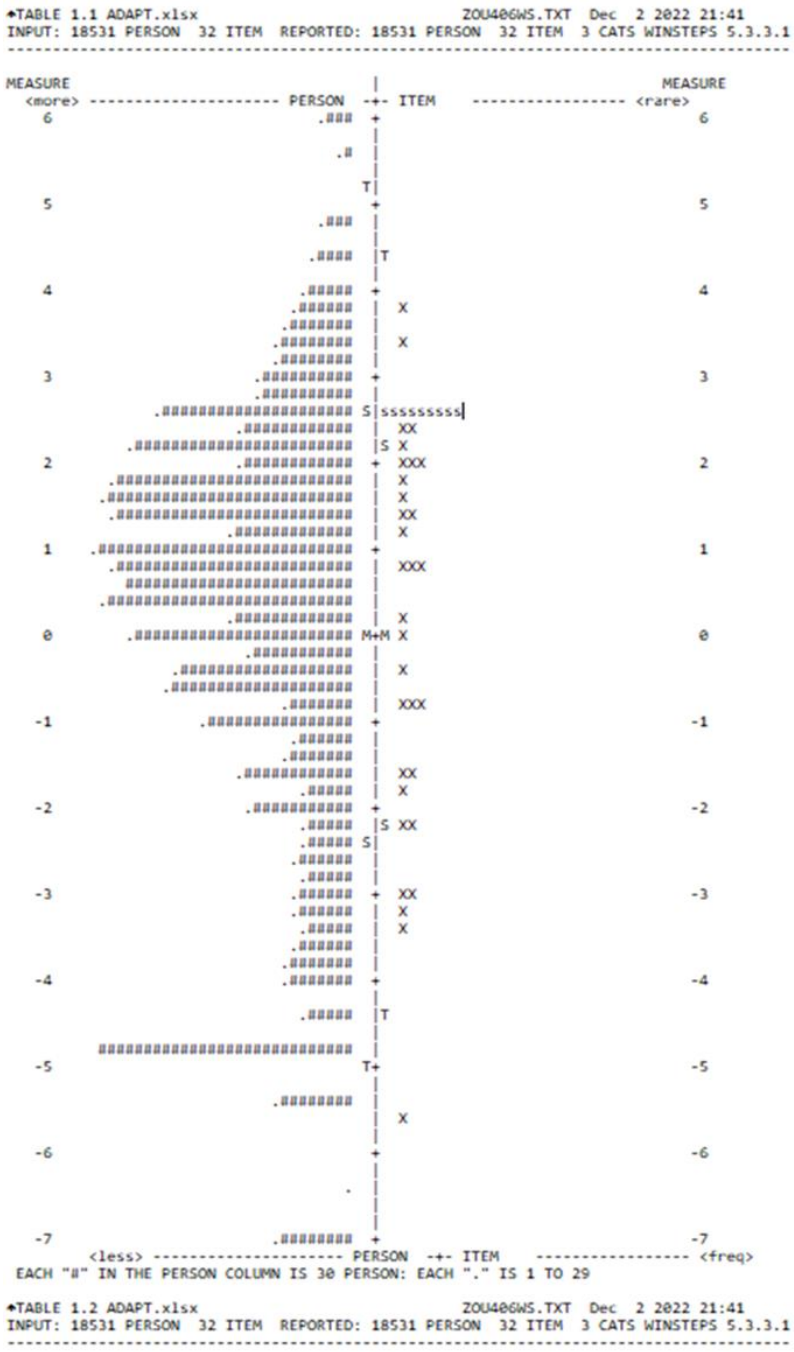


Figure 2. Cognitive Area in AEPS-2

```
TOO BIG FOR TABLE 1.0 - NEEDS LINELENGTH=2251

+TABLE 1.1 COG.xlsx                              ZOU796WS.TXT  Dec  2 2022 22: 8
INPUT: 18531 PERSON  58 ITEM  REPORTED: 18531 PERSON  58 ITEM  3 CATS WINSTEPS 5.3.3.1
-----------------------------------------------------------------------------------------

MEASURE                                 |                                MEASURE
  <more> -------------------- PERSON  -+- ITEM    ----------------- <rare>
    7                                .  +                                  7
                                     .  |
                                        |
    6                                   +                                  6
                                    .#  |
                                     .  |
                                        |T
    5                               .##  +                                  5
                                     .# T|
                                     .#  |
                                   .###  |
    4                              .###  +  X                              4
                                   .###  |  XXX
                                .######  |  X
                              .########  |  XXX
    3                        .##########  +  X                              3
                            .##########  |  XXXXX
                         .#############  |S X
                       .################ S|
    2                 .#################  +  XX                             2
                     .##################  |  XX
              .##########################  |
                   .######################  |  X
    1               .#####################  +  XXXX                         1
          .##############################  |  XX
                   .######################  |  XX
                      .####################  |  X
    0           .########################## M+M X                           0
                     .####################  |  XXX
                   .######################  |  X
                         .################  |  XXXX
   -1                    .################  +                                -1
                         .################  |  XX
                            .############  |  X
                         .#############  |  XXXX
   -2                       .############  +                                -2
                                .#######  S|
                            .###########  |S XX
                                .######  |  XX
   -3                       .###########  +  X                              -3
                               .#######  |  X
                              .########  |
                                .######  |  X
   -4                         .######  +                                   -4
                                   .##  |  X
                             .##### T|  XXX
                                .###  |  XX
   -5                          .####  +                                   -5
                                  .#  |T
                                   .  |
                                      |
   -6                              .  +                                   -6
                                   .  |
                                      |
                                      |
   -7                              .  +                                   -7
                                      |
                                      |
                                      |
   -8                              .  +                                   -8
         <less> -------------------- PERSON  -+- ITEM    ----------------- <freq>
 EACH "#" IN THE PERSON COLUMN IS 39 PERSON: EACH "." IS 1 TO 38

+TABLE 1.2 COG.xlsx                              ZOU796WS.TXT  Dec  2 2022 22: 8
INPUT: 18531 PERSON  58 ITEM  REPORTED: 18531 PERSON  58 ITEM  3 CATS WINSTEPS 5.3.3.1
-----------------------------------------------------------------------------------------
```

Figure 3. Fine Motor Area in AEPS-2

```
+TABLE 1.1 FM.xlsx                                ZOU437WS.TXT  Dec  2 2022 22:17
INPUT: 18531 PERSON  33 ITEM  REPORTED: 18531 PERSON  33 ITEM  3 CATS WINSTEPS 5.3.3.1
-----------------------------------------------------------------------------------

MEASURE                                  |                                MEASURE
  <more> -------------------- PERSON  -+- ITEM    ----------------- <rare>
    7                            .#####  +                                  7
                                .######  |
                                         |
                                         |
    6                   .############  +  X                                6
                                         |T
                         .############  |
                                         |
    5                .################  +                                  5
                      .################ S|  X
                                         |
                      .################  |
    4                  .###############  +                                  4
         .##############################  |
                      .################  |  X
                       .###############  |  X
    3        .########################  +S                                 3
              .########################  |  XXX
                           .###########  |  XX
                 .####################  |
    2              .###################  +                                  2
                    .##################M|  XX
                      .################  |  XX
                       .###############  |
    1                   .############  +  X                                1
                         .###########  |  X
                                #####  |  XX
                          .##########  |  X
    0                      .#########  +M X                                0
                            .########  |
                            .########  |  XX
                                .####  |
   -1                        .#######  +                                  -1
                                .####  |  XX
                            .######## S|  X
                                .####  |
   -2                        .#######  +                                  -2
                                 .###  |  X
                                 .###  |  X
                            .#######  |
   -3                            .###  +S X                               -3
                                  .##  |  X
                                  .##  |  X
                                 .###  |
   -4                            .###  +                                  -4
                                 .###  |  X
                                 .### T|
                               .####  |  X
   -5                                  +  X                               -5
                                 .###  |
                               .####  |
                                      |T
   -6                              .##  +                                  -6
                                      |  X
                                 .###  |
                                      |
   -7                                 +                                  -7
                                    .#  |
                                        |
                                        |
   -8                               .#  +                                  -8
         <less> -------------------- PERSON  -+- ITEM    ----------------- <freq>
 EACH "#" IN THE PERSON COLUMN IS 40 PERSON: EACH "." IS 1 TO 39

+TABLE 1.2 FM.xlsx                                ZOU437WS.TXT  Dec  2 2022 22:17
INPUT: 18531 PERSON  33 ITEM  REPORTED: 18531 PERSON  33 ITEM  3 CATS WINSTEPS 5.3.3.1
```

Figure 4. Gross Motor Area in AEPS-2

```
+TABLE 1.1 GM.xlsx                                ZOU125WS.TXT  Dec  2 2022 22:27
INPUT: 18531 PERSON  55 ITEM  REPORTED: 18531 PERSON  55 ITEM  3 CATS WINSTEPS 5.3.3.1
-----------------------------------------------------------------------------------------

MEASURE                                  |                                MEASURE
  <more> -------------------- PERSON  -+- ITEM    ----------------- <rare>
    9                          .#####  +                                  9
                                       |
                                       |
    8                       .#######  +                                  8
                                       |T
                          .##########  |
    7                   .#############  +  X                              7
                     .################ S|
                    .################  |
    6                .##############  +  X                              6
            ###########################  |  X
           .##########################  |
    5          .######################  +  X                              5
                .####################  |  X
        ############################  |  X
    4       .##########################  +  XXXX                           4
                 .####################  |S XX
               .######################  |
    3                   .############  +  XXXX                           3
                        .############  |
                             .#######  |  XXX
    2                      .#########  +  X                              2
                             .######  |  XXX
                            .########  |  XXXX
    1                        .######  +  X                              1
                            .#######  |  X
                            .#######  |
    0                        .######  +M X                              0
                             .######  |  X
                               .####  |  XX
   -1                             .##  +  X                             -1
                                 .###  |  X
                               .####  |  XX
   -2                         .##### S+  XXX                            -2
                                 .###  |  X
                             .######  |  X
   -3                            .###  +                                -3
                             .#####  |  XX
                               .####  |S XX
   -4                        .#####  +                                 -4
                             .######  |  X
                             .#####  |  X
   -5                        .######  +  X                             -5
                                 .###  |  X
                            .########  |  X
   -6                            .#### T+                              -6
                               .####  |
                            .#######  |
   -7                            .###  +                              -7
                                 .##  |  X
                                 .##  |T
   -8                            .##  +  XXX                          -8
                                  .#  |
                                      |
   -9                              .  +                              -9
                                      |
                                   .  |
  -10                              .  +                             -10
         <less> -------------------- PERSON  -+- ITEM    ----------------- <freq>
 EACH "#" IN THE PERSON COLUMN IS 39 PERSON: EACH "." IS 1 TO 38

+TABLE 1.2 GM.xlsx                                ZOU125WS.TXT  Dec  2 2022 22:27
INPUT: 18531 PERSON  55 ITEM  REPORTED: 18531 PERSON  55 ITEM  3 CATS WINSTEPS 5.3.3.1
-----------------------------------------------------------------------------------------
```

Figure 5. Social communication area in AEPS-2

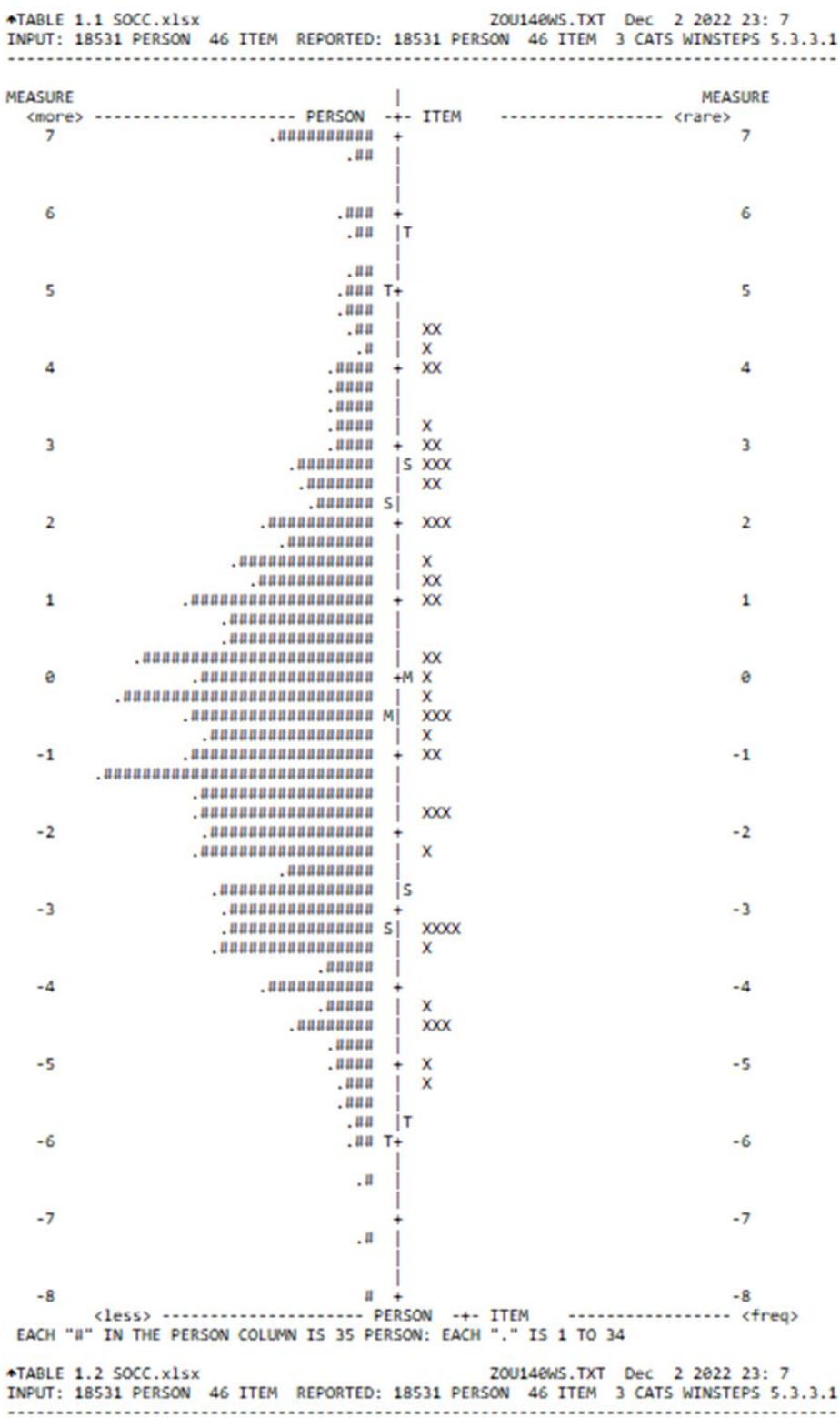


Figure 6. Social area in AEPS-2

```
+TABLE 1.1 SOCE.xlsx                                ZOU954WS.TXT  Dec  3 2022 10:15
INPUT: 18531 PERSON  25 ITEM  REPORTED: 18531 PERSON  25 ITEM  3 CATS WINSTEPS 5.3.3.1
--------------------------------------------------------------------------------------

MEASURE                                 |                                MEASURE
  <more> -------------------- PERSON  -+- ITEM    ----------------- <rare>
    6              .#################  +                                  6
                                      T|
                                       |
                            .#######   |
                                       |
    5                                  +                                  5
                                       |
                           .########   |
                                       |
                          .#########   |
    4                                 +T                                  4
                          .#########   |
                                       |
                         .##########   |
                        .########### S|  X
    3                   .############  +                                  3
                        .############  |  X
                       .#############  |
                      .##############  |  X
                     .###############  |
    2                ################  +S XX                              2
                      .##############  |  X
         .##############################  |  X
                        .############  |  XX
                       .#############  |
    1       .###########################  +                                  1
                        .############  |  XX
              .######################## M|
                        .############  |
              .########################  |  XX
    0                   .############  +M X                               0
                          .#########  |
               .#######################  |  XX
                          .#########  |  X
                          .#########  |
   -1              .##################  +                                 -1
                           .########  |  XX
                           .########  |  X
                             .######  |
                            .####### S|
   -2                        .######  +S                                 -2
                             .######  |  X
                             .######  |
                              .#####  |  X
                             .######  |  X
   -3                         .#####  +                                  -3
                                      |  X
                              .#####  |
                              .#####  |
                                      |
   -4                        .######  +T                                 -4
                                      |
                              .##### T|
                                      |  X
                                      |
   -5                        .######  +                                  -5
                                      |
                                      |
                                      |
                                      |
   -6                        .######  +                                  -6
         <less> -------------------- PERSON  -+- ITEM    ----------------- <freq>
 EACH "#" IN THE PERSON COLUMN IS 37 PERSON: EACH "." IS 1 TO 36

+TABLE 1.2 SOCE.xlsx                                ZOU954WS.TXT  Dec  3 2022 10:15
INPUT: 18531 PERSON  25 ITEM  REPORTED: 18531 PERSON  25 ITEM  3 CATS WINSTEPS 5.3.3.1
--------------------------------------------------------------------------------------
```

## APPENDIX 2. WRIGHT MAP IN THE SIX DEVELOPMENTAL AREAS IN AEPS-3

Figure. Adaptive area in AEPS-3

```
+TABLE 1.1 ADAPT.xlsx                          ZOU265WS.TXT  Dec  3 2022 15:22
INPUT: 710 PERSON  53 ITEM  REPORTED: 710 PERSON  53 ITEM  3 CATS WINSTEPS 5.3.3.1
--------------------------------------------------------------------------------

MEASURE                                |                              MEASURE
  <more> --------------------- PERSON  -+- ITEM    ----------------- <rare>
    7                   .#############  +                                 7
                              ########  |
                                        |
    6                          .######  +                                 6
                               .######  |T
                              .#######  |
    5                         ####### S+                                 5
                        .############  |
                   ###################  |
    4                    .###########  +                                 4
                 .####################  |
                 #####################  |  XX
    3                    .############  +S XXXXX                         3
                   ###################  |  XXX
                      .################  |  XX
    2                    .###########  +  XXX                            2
                    .################ M|  XXXXXXXX
                           .########  |  XXXX
    1                      ########  +  X                                1
                        ############  |  XXX
                       #############  |  X
    0                     ###########  +M XX                             0
                           .########  |
                        ############  |  XX
   -1                        #######  +  X                              -1
                           .#########  |  XX
                          .####### S|  X
   -2                          ####  +                                 -2
                           .########  |  XXX
                               #####  |
   -3                           ####  +S                                -3
                               .###  |  XX
                               .###  |  XX
   -4                            ##  +                                 -4
                                 .#  |  X
                                ###  |  XX
   -5                          .## T+                                  -5
                                ##  |
                                    |T X
   -6                            .  +  X                               -6
                                 .  |
                                    |
   -7                            .  +                                  -7
                                 .  |
                                    |
   -8                               +                                  -8
                          #######  |
                                    |
   -9                               +  X                               -9
                                    |
                                 .  |
  -10                               +                                 -10
                                    |
                                    |
  -11                            .  +                                 -11
         <less> --------------------- PERSON  -+- ITEM    ----------------- <freq>
 EACH "#" IN THE PERSON COLUMN IS 2 PERSON: EACH "." IS 1

+TABLE 1.2 ADAPT.xlsx                          ZOU265WS.TXT  Dec  3 2022 15:22
INPUT: 710 PERSON  53 ITEM  REPORTED: 710 PERSON  53 ITEM  3 CATS WINSTEPS 5.3.3.1
--------------------------------------------------------------------------------
```

Figure. Cognitive area in AEPS-3

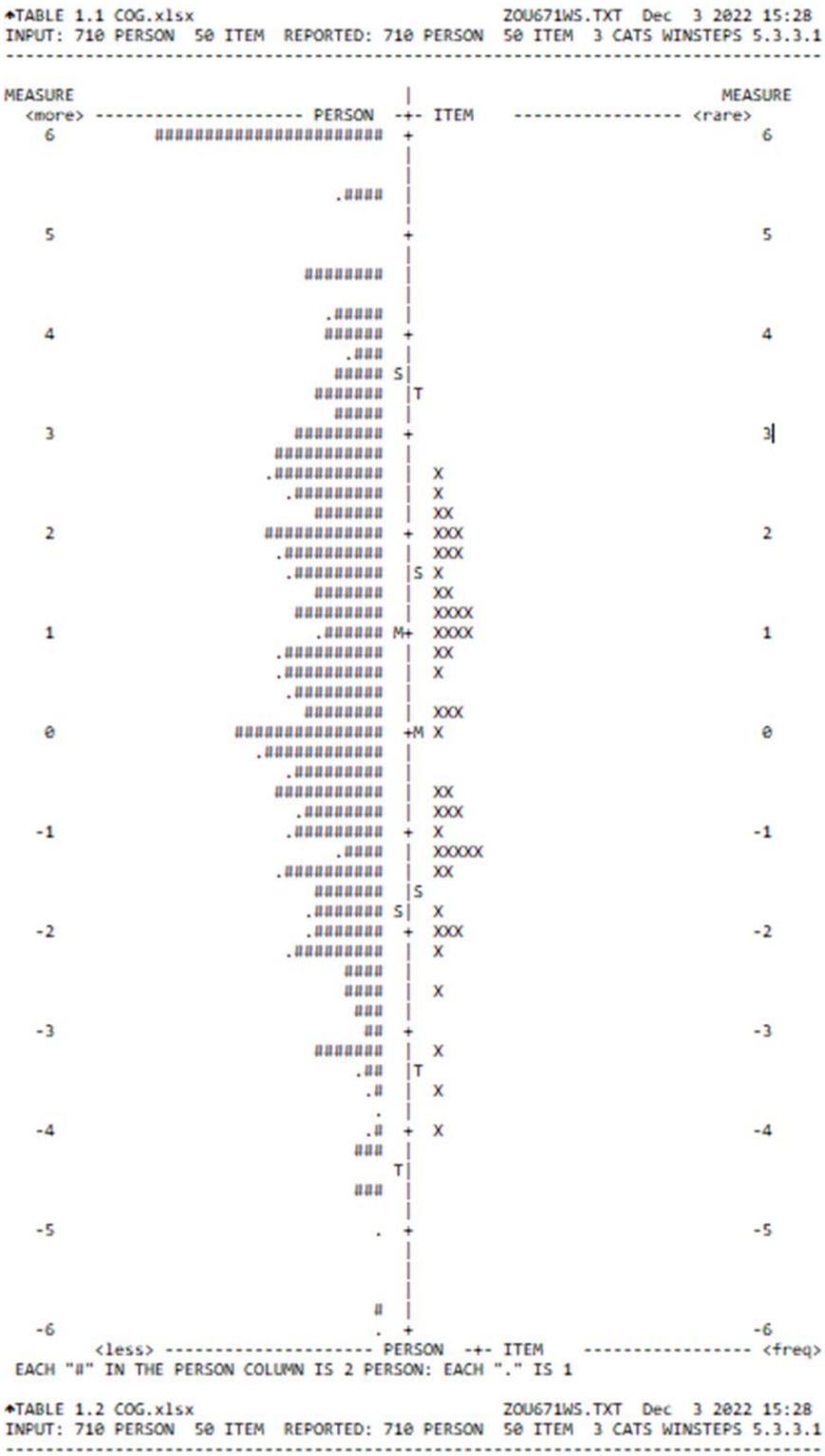


Figure. Gross motor area in AEPS-3

```
+TABLE 1.1 AEPS3 Data Export_07.22.22_revised_08. ZOU578WS.TXTs Dec  3 2022 15: 7
INPUT: 711 PERSON  65 ITEM  REPORTED: 711 PERSON  65 ITEM  3 CATS WINSTEPS 5.3.3.1
--------------------------------------------------------------------------------

MEASURE                                |                                MEASURE
  <more> --------------------- PERSON  -+- ITEM    ----------------- <rare>
    8                     ############  +                                 8
                                 #####  |
                                        |
                                        |
    7                        .######  +                                 7
                                        |
                           ########## S|T
                            #########  |
    6                            #####  +                                 6
                 .####################  |  X
                           ############  |  XX
                  .###################  |
    5         .########################  +  XX                            5
                .########################  |  XX
                  #######################  |
                        .############  |  X
    4         .#########################  +  XXX                           4
                           ##########  |  X
                .######################## M|
                  .###################  |S X
    3                        .######  +  XX                            3
                           ##########  |  X
                          ###########  |  XXX
                               .#####  |  XX
    2                         #######  +  XX                            2
                                  ###  |
                                  ###  |  X
                                   ##  |  X
    1                            #####  +                                 1
                                 .## S|  XX
                                  .#  |
                                  .#  |  X
    0                             ####  +M XXXX                          0
                                 .###  |  XXXX
                                 .###  |  X
                                  ###  |  X
   -1                              ##  +  X                            -1
                                #####  |  X
                                 .###  |
                                 ####  |  XX
   -2                               . T+  XX                            -2
                                    #  |  XXXX
                                  .#  |  XX
                                  .#  |  X
   -3                              ##  +  XX                            -3
                                  ###  |S X
                                 .##  |  X
                                  .#  |  X
   -4                              .  +  XXX                           -4
                                   #  |  X
                                      |  X
                                   #  |
   -5                              .  +  XX                            -5
                                   .  |
                                   #  |
                                   .  |
   -6                              .  +                                 -6
                                      |  X
                                   .  |T
                                      |
   -7                                 +                                 -7
                                      |  X
                                      |
                                      |
   -8                                 +                                 -8
                                   .  |
                                      |
                                      |
   -9                                 +                                 -9
         <less> --------------------- PERSON  -+- ITEM    ----------------- <freq>
EACH "#" IN THE PERSON COLUMN IS 2 PERSON: EACH "." IS 1
```

Figure. Fine Motor area in AEPS-3

```
↑TABLE 1.1 AEPS3 Data Export_07.22.22_revised_08. ZOU781WS.TXTs Dec  3 2022 14:57
INPUT: 710 PERSON  31 ITEM  REPORTED: 710 PERSON  31 ITEM  3 CATS WINSTEPS 5.3.3.1
--------------------------------------------------------------------------------

MEASURE                                |                                MEASURE
  <more> --------------------- PERSON  -+- ITEM    ----------------- <rare>
    7    ################################ S+                                 7
                                        |
                            ###########  |
                                        |
    6                                   +                                  6
                              .########  |
                                        |
                               .######  |T
    5                                   +                                  5
                             .#######  |  X
                              .######  |  X
                               .#####  |  X
    4                         .###### M+                                  4
                               ######  |  X
                                #####  |
                               .####  |
    3                            .##  +                                  3
                                .###  |S
                                .###  |  X
                                   #  |  X
    2                          .####  +  X                               2
                                .###  |  XXX
                                 .##  |
                                .###  |  X
    1                             .# S+  X                               1
                                   .  |  X
                                  .#  |  XX
                                  .#  |  X
    0                             .#  +M XX                              0
                                  .#  |  X
                                  .#  |  X
                                   #  |
   -1                             .#  +                                 -1
                                  .#  |
                                   #  |  X
                                   #  |  X
   -2                               . T+  XX                             -2
                                   #  |
                                   .  |  XX
                                   .  |S
   -3                               .  +  X                              -3
                                       |
                                   .  |
                                       |
   -4                               .  +  X                              -4
                                   .  |  X
                                   .  |  X
                                   .  |  X
   -5                                  +                                 -5
                                       |T
                                       |
                                       |
   -6                                  +                                 -6
                                       |
                                   .  |
                                       |
   -7                               .  +                                 -7
         <less> --------------------- PERSON  -+- ITEM    ----------------- <freq>
 EACH "#" IN THE PERSON COLUMN IS 5 PERSON: EACH "." IS 1 TO 4

↑TABLE 1.2 AEPS3 Data Export_07.22.22_revised_08. ZOU781WS.TXTs Dec  3 2022 14:57
INPUT: 710 PERSON  31 ITEM  REPORTED: 710 PERSON  31 ITEM  3 CATS WINSTEPS 5.3.3.1
--------------------------------------------------------------------------------
```

Figure. Social-Communication area in AEPS-3

```
+TABLE 1.1 SOCC.xlsx                          ZOU703WS.TXT  Dec  3 2022 15:34
INPUT: 710 PERSON  49 ITEM  REPORTED: 710 PERSON  49 ITEM  3 CATS WINSTEPS 5.3.3.1
--------------------------------------------------------------------------------

MEASURE                                |                                MEASURE
  <more> --------------------- PERSON  -+- ITEM    ----------------- <rare>
    6           .######################  +                                  6
                                        |
                            ##########  |
                                        |
                                        |
    5                    #############  +                                  5
                                        |
                             .######## S|
                                        |
                        ###############  |T
    4                          #######  +                                  4
                                 #####  |
                        .#############  |
                           ##########  |
                         #############  |
    3                          ######  +                                  3
                             .#######  |  X
                             ########  |  X
                       ###############  |  XXX
                             .######  |  XXXX
    2                         .#####  +S XXXX                             2
                           #########  |  XXX
                            .#######  |  XXX
                          .#########  |  X
                            .####### M|  XX
    1                       ########  +  X                                1
                            .#######  |  XX
                           #########  |
                          ###########  |  X
                          ##########  |  X
    0                           #####  +M                                 0
                               ######  |  X
                          ##########  |
                               #####  |
                                ####  |  X
   -1                          .####  +  XXX                            -1
                                 .##  |
                               #######  |
                              .#####  |  XXXX
                              ######  |  XX
   -2                         ####### S+S XX                            -2
                              ######  |  X
                                  ##  |
                                 .##  |  XX
                               #####  |  X
   -3                         .#####  +                                 -3
                                .###  |  X
                                 ###  |
                                  ##  |  X
                                 .##  |  XX
   -4                                 +  X                              -4
                                   #  |T
                                   .  |
                                      |
                                      |
   -5                           ###  +                                 -5
                                    T|
                                      |
                                 .##  |
                                      |
   -6                                 +                                 -6
                                  ##  |
                                      |
                                      |
                                      |
   -7                     ##########  +                                 -7
         <less> --------------------- PERSON  -+- ITEM    ----------------- <freq>
 EACH "#" IN THE PERSON COLUMN IS 2 PERSON: EACH "." IS 1
```

Figure. Social- Emotional area in AEPS-3

```
+TABLE 1.1 SOCE.xlsx                                ZOU265WS.TXT  Dec  3 2022 15:41
INPUT: 710 PERSON  62 ITEM  REPORTED: 710 PERSON  62 ITEM  3 CATS WINSTEPS 5.3.3.1
--------------------------------------------------------------------------------

MEASURE                                |                                MEASURE
  <more> -------------------- PERSON  -+- ITEM    ----------------- <rare>
    6                  .###############  +                                  6
                                       T|
                                        |
                              .#####    |
                                        |
    5                                   +                                  5
                                        |
                       .##############  |
                                        |
                             ########   |
    4                                   +                                  4
                                .####   |
                               .#####   |
                             ######## S|
                            .########   |
    3                           #####   +                                  3
                           ##########   |T
                        ##############  |
                              #######   |  XX
                            .########   |
    2                    .############  +                                  2
                          .###########  |  X
                         ##############  |  X
                        .##############  |S XX
                      .###############M|  XXXX
    1                  .##############  +  XX                              1
                     #################  |  XXXXX
                        .#############  |  XXXXXXXXX
                         ############  |  XXXXXXXXXXXX
                           .##########  |  XXX
    0                     .########   +M XX                               0
                         .##########  |  XXX
                              #######  |  X
                          ###########  |  X
                           .########   |
   -1                          ######  +                                 -1
                             ######## S|  XX
                              .######  |S XX
                              .######  |  X
                             ########  |
   -2                       .########  +  XX                             -2
                                  ###  |  XX
                                 .###  |
                               ######  |  X
                                  .##  |T XX
   -3                            .###  +                                 -3
                                    #  |
                                    .  |
                                   .# T|  X
                                   ##  |  X
   -4                                .  +                                 -4
                                     .  |
                                        |
                                        |
                                    #   |
   -5                                   +                                 -5
                                    #   |
                                        |
                                        |
                                        |
   -6                                #  +                                 -6
        <less> -------------------- PERSON  -+- ITEM    ----------------- <freq>
 EACH "#" IN THE PERSON COLUMN IS 2 PERSON: EACH "." IS 1
```

## APPENDIX 3. LIST OF ITEMS IN AEPS-2 ORDER BY ITEM DIFFICULTY PARAMETER

Adaptive Area Item (display by the item difficulty followed the dose from hard to easy)

| | | |
|---|---|---|
| **1** | **adapt_B1.0** | **Initiates toileting** |
| **2** | adapt_B1.1 | Demonstrates bowel and bladder control |
| **3** | adapt_C1.1 | Takes off pullover shirt/sweater |
| **4** | adapt_A5.0 | Transfers food and liquid between containers |
| **5** | adapt_A5.1 | Pours liquid between containers |
| **6** | adapt_B2.0 | Washes and dries hands |
| **7** | adapt_C1.2 | Takes off front-fastened coat, jacket, or shirt |
| **8** | adapt_B1.2 | Indicates awareness of soiled and wet pants and/or diapers |
| **9** | adapt_B3.0 | Brushes teeth |
| **10** | adapt_A5.2 | Transfers food between containers |
| **11** | adapt_C1.3 | Takes off pants |
| **12** | adapt_A3.0 | Drinks from cup and/or glass |
| **13** | adapt_B2.1 | Washes hands |
| **14** | adapt_A4.0 | Eats with fork and/or spoon |
| **15** | adapt_C1.0 | Undresses self |
| **16** | adapt_A3.1 | Drinks from cup and/or glass with some spilling |
| **17** | adapt_A4.1 | Brings food to mouth using utensil |
| **18** | adapt_B3.1 | Cooperates with teeth brushing |
| **19** | adapt_C1.5 | Takes off shoes |
| **20** | adapt_A3.2 | Drinks from cup and/or glass held by adult |
| **21** | adapt_A2.0 | Bites and chews hard and chewy foods |
| **22** | adapt_C1.4 | Takes off socks |
| **23** | adapt_C1.6 | Takes off hat |
| **24** | adapt_A2.1 | Bites and chews soft and crisp foods |
| **25** | adapt_A1.1 | Uses lips to take in liquids from a cup and/or glass |
| **26** | adapt_A4.2 | Eats with fingers |
| **27** | adapt_A2.2 | Munches soft and crisp foods |
| **28** | adapt_A1.0 | Uses tongue and lips to take in and swallow solid foods and liquids |
| **29** | adapt_A1.3 | Swallows solid and semi-solid foods |
| **30** | adapt_A1.2 | Uses lips to take food off spoon and/or fork |
| **31** | adapt_A4.3 | Accepts food presented on spoon |
| **32** | adapt_A1.4 | Swallows liquids |

Cognitive Area Item (display by the item difficulty followed the dose from hard to easy)

| No. | Item | Content |
|---|---|---|

| | | |
|---|---|---|
| **1** | cog_G5.0 | Demonstrates use of common opposite concepts |
| **2** | cog_G4.1 | Orally fills in or completes familiar text while looking at picture books |
| **3** | cog_G6.0 | Repeats simple nursery rhymes |
| **4** | cog_G5.1 | Demonstrates use of at least four pairs of common opposite concepts |
| **5** | cog_G6.1 | Fills in rhyming words in familiar rhymes |
| **6** | cog_G2.0 | Demonstrates functional use of one-to-one correspondence |
| **7** | cog_G1.0 | Categorizes like objects |
| **8** | cog_G1.1 | Groups functionally related objects |
| **9** | cog_G1.2 | Groups objects according to size, shape, and/or color |
| **10** | cog_G4.2 | Makes comments and asks questions while looking at picture books |
| **11** | cog_G4.0 | Demonstrates functional use of reading materials |
| **12** | cog_G5.2 | Demonstrates use of at least two pairs of common opposite concepts |
| **13** | cog_G3.0 | Recognizes environmental symbols (signs, logos, labels) |
| **14** | cog_G6.2 | Says nursery rhymes along with familiar adult |
| **15** | cog_G2.1 | Demonstrates concept of one |
| **16** | cog_F1.0 | Uses imaginary objects in play |
| **17** | cog_G1.3 | Matches pictures and/or objects |
| **18** | cog_D2.0 | Imitates words that are not frequently used |
| **19** | cog_G3.1 | Labels familiar people, actions, objects, and events in pictures |
| **20** | cog_E4.0 | Solves common problems |
| **21** | cog_D2.1 | Imitates speech sounds that are not frequently used |
| **22** | cog_F1.1 | Uses representational actions with objects |
| **23** | cog_G4.3 | Sits and attends to entire story during shared reading time |
| **24** | cog_E2.0 | Uses an object to obtain another object |
| **25** | cog_D1.0 | Imitates motor action that is not commonly used |
| **26** | cog_B3.0 | Maintains search for object that is not in its usual location |
| **27** | cog_E4.1 | Uses more than one strategy in attempt to solve common problem |
| **28** | cog_E2.1 | Uses part of object and/or support to obtain another object |
| **29** | cog_D2.2 | Imitates words that are frequently used |
| **30** | cog_B2.0 | Locates object in latter of two successive hiding places |
| **31** | cog_E3.0 | Navigates large object around barriers |
| **32** | cog_F1.2 | Uses functionally appropriate actions with objects |
| **33** | cog_C1.0 | Correctly activates mechanical toy |
| **34** | cog_C2.0 | Reproduces part of interactive game and/or action in order to continue game and/or action |
| **35** | cog_B3.1 | Looks for object in usual location |
| **36** | cog_E1.0 | Retains objects when new object is obtained |
| **37** | cog_E3.1 | Moves barrier or goes around barrier to obtain object |
| **38** | cog_D1.1 | Imitates motor action that is commonly used |
| **39** | cog_B2.1 | Locates object and/or person hidden while child is watching |
| **40** | cog_E3.2 | Moves around barrier to change location |
| **41** | cog_C2.1 | Indicates desire to continue familiar game and/or action |
| **42** | cog_C1.1 | Correctly activates simple toy |
| **43** | cog_B2.2 | Locates object and/or person who is partially hidden while child is watching |
| **44** | cog_E1.1 | Retains one object when second object is obtained |
| **45** | cog_F1.3 | Uses simple motor actions on different objects |
| **46** | cog_C1.2 | Acts on mechanical and/or simple toy in some way |

| | | |
|---|---|---|
| **47** | cog_B2.3 | Reacts when object and/or person hides from view |
| **48** | cog_F1.4 | Uses sensory examination with objects |
| **49** | cog_E1.2 | Retains object |
| **50** | cog_C1.3 | Indicates interest in simple and/or mechanical toy |
| **51** | cog_B1.0 | Visually follows object and/or person to point of disappearance |
| **52** | cog_B1.1 | Visually follows object moving in horizontal, vertical, and circular directions |
| **53** | cog_A1.0 | Orients to auditory, visual, and tactile events |
| **54** | cog_A1.1 | Orients to auditory events |
| **55** | cog_B1.2 | Focuses on object and/or person |
| **56** | cog_A1.2 | Orients to visual events |
| **57** | cog_A1.3 | Orients to tactile stimulation |
| **58** | cog_A1.4 | Responds to auditory, visual, and tactile events |

Fine Motor Area Item (display by the item difficulty followed the dose from hard to easy)

| No. | Item | Content |
|---|---|---|
| **1** | fm_B5 | Copies simple written shapes after demonstration |
| **2** | fm_B5.1 | Draws circles and lines |
| **3** | fm_B4 | Orients picture book correctly and turns pages one by one |
| **4** | fm_B2 | Assembles toy and/or object that require(s) putting pieces together |
| **5** | fm_A5.1 | Aligns objects |
| **6** | fm_B2.1 | Fits variety of shapes into corresponding spaces |
| **7** | fm_B4.2 | Turns/holds picture book right side up |
| **8** | fm_A5 | Aligns and stacks objects |
| **9** | fm_B4.1 | Turns pages of books |
| **10** | fm_A5.2 | Places and releases object balanced on top of another object with either hand |
| **11** | fm_B2.2 | Fits object into defined space |
| **12** | fm_B1 | Rotates either wrist on horizontal plane |
| **13** | fm_B5.2 | Scribbles |
| **14** | fm_B3 | Uses either index finger to activate objects |
| **15** | fm_A4. | Grasps pea-size object with either hand using tip of the index finger and thumb with hand and/or arm not resting on surface for support |
| **16** | fm_A4.1 | Grasps pea-size object with either hand using tip of the index finger and thumb with hand and/or arm resting on surface for support |
| **17** | fm_B1.1 | Turns object over using wrist and arm rotation with each hand |
| **18** | fm_A5.3 | Releases hand-held object onto and/or into a larger target with either hand |
| **19** | fm_A4.2 | Grasps pea-size object with either hand using side of the index finger and thumb |
| **20** | fm_B3.1 | Uses either hand to activate objects |
| **21** | fm_A3 | Grasps hand-size object with either hand using ends of thumb, index, and second fingers |
| **22** | fm_A3.1 | Grasps hand-size object with either hand using the palm, with object placed toward the thumb and index finger |
| **23** | fm_A5.4 | Releases hand-held object with each hand |
| **24** | fm_A4.3 | Grasps pea-size object with either hand using fingers in a raking and/or scratching movement |
| **25** | fm_A2 | Brings two objects together at or near midline |

| | | |
|---|---|---|
| **26** | fm_A2.1 | Transfers object from one hand to the other |
| **27** | fm_A3.2 | Grasps cylindrical object with either handby closing fingers around it |
| **28** | fm_A2.2 | Holds an object in each hand |
| **29** | fm_A3.3 | Grasps hand-size object with either hand using whole hand |
| **30** | fm_A2.3 | Reaches toward and touches object with each hand |
| **31** | fm_A1 | Simultaneously brings hands to midline |
| **32** | fm_A1.1 | Makes directed batting and/or swiping movements with each hand |
| **33** | fm_A1.2 | Makes nondirected movements with each arm |

Gross Motor Area Item (display by the item difficulty followed the dose from hard to easy)

| | | |
|---|---|---|
| **1** | **gm_D2** | **Pedals and steers tricycle** |
| **2** | gm_D3.1 | Catches ball or similar object |
| **3** | gm_D1 | Jumps forward |
| **4** | gm_D1.1 | Jumps up |
| **5** | gm_D1.2 | Jumps from low structure |
| **6** | gm_D3 | Catches, kicks, throws, and rolls ball or similar object |
| **7** | gm_D3.2 | Kicks ball or similar object |
| **8** | gm_C4. | Walks up and down stairs |
| **9** | gm_C3 | Runs avoiding obstacles |
| **10** | gm_D2.1 | Pushes riding toy with feet while steering |
| **11** | gm_D4. | Climbs up and down play equipment |
| **12** | gm_C3.1 | Runs |
| **13** | gm_D3.3 | Throws ball or similar object at target |
| **14** | gm_D4.1 | Moves up and down inclines |
| **15** | gm_C4.1 | Walks up and down stairs using two-hand support |
| **16** | gm_D3.4 | Rolls ball at target |
| **17** | gm_C3.2 | Walks fast |
| **18** | gm_C4.2 | Moves up and down stairs |
| **19** | gm_C1. | Walks avoiding obstacles |
| **20** | gm_D4.2 | Moves under, over, and through obstacles |
| **21** | gm_B2. | Sits down in and gets out of chair |
| **22** | gm_C2. | Stoops and regains balanced standing position without support |
| **23** | gm_D2.2 | Sits on riding toy or in wagon while adult pushes |
| **24** | gm_C1.1 | Walks without support |
| **25** | gm_C4.3 | Gets up and down from low structure |
| **26** | gm_B2.1 | Sits down in chair |
| **27** | gm_C2.1 | Rises from sitting position to standing position |
| **28** | gm_C1.2 | Walks with one-hand support |
| **29** | gm_C1.4 | Stands unsupported |
| **30** | gm_C1.3 | Walks with two-hand support |
| **31** | gm_C1.5 | Cruises |

| | | |
|---|---|---|
| **32** | gm_C2.2 | Pulls to standing position |
| **33** | gm_C2.3 | Pulls to kneeling position |
| **34** | gm_B2.2 | Maintains a sitting position in chair |
| **35** | gm_A3. | Creeps forward using alternating arm and leg movements |
| **36** | gm_B1.1 | Assumes hands and knees position from sitting |
| **37** | gm_A3.1 | Rocks while in a creeping position |
| **38** | gm_A3.2 | Assumes creeping position |
| **39** | gm_B1. | Assumes balanced sitting position |
| **40** | gm_A3.3 | Crawls forward on stomach |
| **41** | gm_B1.2 | Regains balanced, upright sitting position after reaching across the body to the right and to the left |
| **42** | gm_B1.3 | Regains balanced, upright sitting position after leaning to the left, to the right, and forward |
| **43** | gm_B1.4 | Sits balanced without support |
| **44** | gm_A3.5 | Bears weight on one hand and/or arm while reaching with opposite hand |
| **45** | gm_A3.4 | Pivots on stomach |
| **46** | gm_B1.5 | Sits balanced using hands for support |
| **47** | gm_A2. | Rolls by turning segmentally from stomach to back and from back to stomach |
| **48** | gm_A2.1 | Rolls from back to stomach |
| **49** | gm_A2.2 | Rolls from stomach to back |
| **50** | gm_A3.6 | Lifts head and chest off surface with weight on arms |
| **51** | gm_B1.6 | Holds head in midline when in supported sitting position |
| **52** | gm_A1. | Turns head, moves arms, and kicks legs independently of each other |
| **53** | gm_A1.1 | Turns head past 45° to the right and left from midline position |
| **54** | gm_A1.3 | Waves arms |
| **55** | gm_A1.2 | Kicks legs |

Social Communication Area Item (display by the item difficulty followed the dose from hard to easy)

| | | |
|---|---|---|
| **1** | **sc_D3.1** | **Uses three-word negative utterances** |
| **2** | sc_D3.3 | Uses three-word action–object–location utterances |
| **3** | sc_D3.4 | Uses three-word agent–action–object utterances |
| **4** | sc_D3.2 | Asks questions |
| **5** | sc_D3 | Uses three-word utterances |
| **6** | sc_D2.3 | Uses two-word utterances to express location |
| **7** | sc_D2.4 | Uses two-word utterances to describe objects, people, and/or events |
| **8** | sc_D2.1 | Uses two-word utterances to express agent–action, action–object, and agent–object |
| **9** | sc_D2.2 | Uses two-word utterances to express possession |
| **10** | sc_D2.5 | Uses two-word utterances to express recurrence |
| **11** | sc_D2.6 | Uses two-word utterances to express negation |
| **12** | sc_D2 | Uses two-word utterances |
| **13** | sc_D1.1 | Uses five descriptive words |

| | | |
|---|---|---|
| **14** | sc_C2 | Carries out two-step direction without contextual cues |
| **15** | sc_D1.3 | Uses two pronouns |
| **16** | sc_D1.2 | Uses five action words |
| **17** | sc_C2.1 | Carries out two-step direction with contextual cues |
| **18** | sc_C1.1 | Locates common objects, people, and/or events in unfamiliar pictures |
| **19** | sc_C1 | Locates objects, people, and/or events without contextual cues |
| **20** | sc_D1 | Uses 50 single words |
| **21** | sc_D1.4 | Uses 15 object and/or event labels |
| **22** | sc_C1.2 | Locates common objects, people, and/or events in familiar pictures |
| **23** | sc_D1.5 | Uses three proper names |
| **24** | sc_C2.2 | Carries out one-step direction without contextual cues |
| **25** | sc_B2 | Uses consistent word approximations |
| **26** | sc_B1.1 | Responds with a vocalization and gesture to simple questions |
| **27** | sc_B1 | Gains person's attention and refers to an object, person, and/or event |
| **28** | sc_C1.3 | Locates common objects, people, and/or events with contextual cues |
| **29** | sc_B2.1 | Uses consistent consonant–vowel combinations |
| **30** | sc_C2.3 | Carries out one-step direction with contextual cues |
| **31** | sc_B1.2 | Points to an object, person, and/or event |
| **32** | sc_B1.3 | Gestures and/or vocalizes to greet others |
| **33** | sc_A2 | Follows person's gaze to establish joint attention |
| **34** | sc_B2.2 | Uses nonspecific consonant–vowel combinations and/or jargon |
| **35** | sc_A2.1 | Follows person's pointing gesture to establish joint attention |
| **36** | sc_B1.4 | Uses gestures and/or vocalizations to protest actions and/or reject objects or people |
| **37** | sc_A3 | Engages in vocal exchanges by babbling |
| **38** | sc_C1.4 | Recognizes own name |
| **39** | sc_B2.3 | Vocalizes to express affective states |
| **40** | sc_B2.4 | Vocalizes open syllables |
| **41** | sc_C1.5 | Quiets to familiar voice |
| **42** | sc_A1. | Turns and looks toward person speaking |
| **43** | sc_A3.1 | Engages in vocal exchanges by cooing |
| **44** | sc_A1.1 | Turns and looks toward object and person speaking |
| **45** | sc_A2.2 | Looks toward an object |
| **46** | sc_A1.2 | Turns and looks toward noise-producing object |

Social Area Item (display by the item difficulty followed the dose from hard to easy)

| | | |
|---|---|---|
| **1** | **soc_C2** | **Initiates and maintains communicative exchange with peer** |
| **2** | soc_C2.1 | Initiates communication with peer |
| **3** | soc_C1 | Initiates and maintains interaction with peer |
| **4** | soc_C1.1 | Initiates social behavior toward peer |
| **5** | soc_B1 | Meets observable physical needs in socially appropriate ways |
| **6** | soc_C2.2 | Responds to communication from peer |

| | | |
|---|---|---|
| **7** | soc_C1.2 | Responds appropriately to peer's social behavior |
| **8** | soc_A3 | Initiates and maintains communicative exchange with familiar adult |
| **9** | soc_B2 | Participates in established social routines |
| **10** | soc_B1.1 | Meets internal physical needs of hunger, thirst, and rest |
| **11** | soc_A3.1 | Initiates communication with familiar adult |
| **12** | soc_B2.1 | Responds to established social routines |
| **13** | soc_A2.1 | Initiates simple social game with familiar adult |
| **14** | soc_A2. | Initiates and maintains interaction with familiar adult |
| **15** | soc_B1.2 | Uses appropriate strategies to self-soothe |
| **16** | soc_C1.3 | Plays near one or two peers |
| **17** | soc_A3.2 | Responds to communication from familiar adult |
| **18** | soc_A2.2 | Responds to familiar adult's social behavior |
| **19** | soc_C1.4 | Observes peers |
| **20** | soc_C1.5 | Entertains self by playing appropriately with toys |
| **21** | soc_A1 | Responds appropriately to familiar adult's affect |
| **22** | soc_A1.1 | Displays affection toward familiar adult |
| **23** | soc_A1.2 | Responds appropriately to familiar adult's affective tone |
| **24** | soc_A2.3 | Uses familiar adults for comfort, closeness, or physical contact |
| **25** | soc_A1.3 | Smiles in response to familiar adult |

## APPENDIX 4. LIST OF ITEMS IN AEPS-3 ORDER BY THE ITEM DIFFICULTY PARAMETER

Adaptive Area Item (display by the item difficulty followed the dose from hard to easy)

| | | |
|---|---|---|
| **1** | **AdaptC2.1** | **Fastens clothing** |
| **2** | AdaptD4.1 | States or produces personal information to promote/maintain personal safety |
| **3** | AdaptD4.0 | Recognizes and reports information regarding safety |
| **4** | AdaptB3.0 | Completes all steps for personal hygiene, including brushing teeth, combing hair, and wiping nose |
| **5** | AdaptB2.0 | Bathes and dries self |
| **6** | AdaptD2.1 | Complies with graphic or written warning signs and symbols |
| **7** | AdaptC2.0 | Selects appropriate clothing and dresses self |
| **8** | AdaptD4.2 | Reports inappropriate events, actions, or language by others |
| **9** | AdaptD3.0 | Takes independent action when faced with dangerous conditions or substances |
| **10** | AdaptB2.1 | Washes and dries face |
| **11** | AdaptA6.0 | Prepares food for eating |
| **12** | AdaptC2.3 | Puts on pullover clothing |
| **13** | AdaptA6.2 | Serves food with utensil |
| **14** | AdaptD2.0 | Complies with common home and community safety rules |

| | | |
|---|---|---|
| **15** | AdaptB1.0 | Carries out all toileting functions |
| **16** | AdaptD3.1 | Responds appropriately to warnings of dangerous conditions or substances |
| **17** | AdaptC2.2 | Puts on front-opening clothing |
| **18** | AdaptD1.0 | Takes independent action to alleviate distress, discomfort, and pain |
| **19** | AdaptC1.1 | Unfastens clothing |
| **20** | AdaptC1.2 | Takes off pullover clothing over head |
| **21** | AdaptA6.1 | Pours liquid into variety of containers |
| **22** | AdaptB1.1 | Indicates need to use toilet |
| **23** | AdaptC1.0 | Undresses self by removing all clothing |
| **24** | AdaptB3.1 | Completes some steps to brush teeth, comb hair, and wipe nose |
| **25** | AdaptB1.2 | Has bowel and bladder contro |
| **26** | AdaptC2.4 | Puts on pull-up clothing |
| **27** | AdaptC2.5 | Puts on socks |
| **28** | AdaptC2.6 | Puts on shoes |
| **29** | AdaptD1.1 | Communicates internal distress, discomfort, or pain to adult |
| **30** | AdaptA5.0 | Uses culturally appropriate social dining skills |
| **31** | AdaptB1.3 | Indicates awareness of soiled and wet pants or diapers |
| **32** | AdaptB2.2 | Washes and dries hands |
| **33** | AdaptC1.4 | Takes off pants |
| **34** | AdaptC1.3 | Takes off front-opening coat, jacket, or shirt |
| **35** | AdaptA5.1 | Puts appropriate amount of food in mouth, chews, and swallows before taking another bite |
| **36** | AdaptA3.0 | Eats with eating utensils |
| **37** | AdaptA4.0 | Drinks from open-mouth container |
| **38** | AdaptC1.5 | Takes off shoes |
| **39** | AdaptA5.2 | Takes in appropriate amount of liquid and returns cup to surface |
| **40** | AdaptA3.1 | Brings food to mouth with eating utensil |
| **41** | AdaptC1.6 | Takes off socks |
| **42** | AdaptA2.1 | Eats hard and chewy foods |
| **43** | AdaptA2.0 | Eats foods from variety of food groups with variety of textures |
| **44** | AdaptA2.2 | Eats crisp foods |
| **45** | AdaptA4.1 | Drinks from cup with spouted lid |
| **46** | AdaptC1.7 | Takes off hat |
| **47** | AdaptA4.2 | Drinks from container held by adult |
| **48** | AdaptA3.3 | Accepts food presented on eating utensils |
| **49** | AdaptA3.2 | Eats with fingers |
| **50** | AdaptA2.3 | Eats soft and dissolvable foods |
| **51** | AdaptA1.0 | Uses lips to take semisolid foods off eating utensil |
| **52** | AdaptA1.1 | Swallows semisolid foods |
| **53** | AdaptA1.2 | Swallows liquids |

Cognitive Area Item (display by the item difficulty followed the dose from hard to easy)

| No. | Item | Content |
|---|---|---|
| **1** | CogE4.0 | Transfers knowledge |
| **2** | CogE3.0 | Investigates to test hypotheses |
| **3** | CogE4.1 | Communicates results of investigations |
| **4** | CogE4.2 | Demonstrates knowledge of properties of change resulting from investigations |
| **5** | CogD4.0 | Draws plausible conclusions about events beyond personal experience |
| **6** | CogE2.1 | Generates specific questions for investigation |
| **7** | CogE3.1 | Draws on prior knowledge to guide investigations |
| **8** | CogE2.0 | Anticipates outcome of investigation |
| **9** | CogE4.3 | Shows awareness that manipulation of materials or processes prompted change in those materials or processes |
| **10** | CogE2.2 | Demonstrates knowledge about natural happenings |
| **11** | CogE3.2 | Manipulates materials to cause change |
| **12** | CogE1.0 | Expands simple observations and explorations into further inquiry |
| **13** | CogD4.1 | Draws conclusions about causes of events based on personal experience |
| **14** | CogE2.3 | Makes observations |
| **15** | CogD3.0 | Solves problems using multiple strategies |
| **16** | CogD3.1 | Evaluates common solutions to solve problems or reach goals |
| **17** | CogC3.0 | Classifies using multiple attributes |
| **18** | CogB3.0 | Relates past events |
| **19** | CogE1.1 | Uses simple tools to gather information |
| **20** | CogC3.1 | Classifies according to function |
| **21** | CogC4.0 | Uses early conceptual comparisons |
| **22** | CogB3.1 | Relates recent events without contextual cues |
| **23** | CogC4.1 | Identifies common concepts |
| **24** | CogB3.2 | Relates recent events with contextual cues |
| **25** | CogB3.3 | Relates events immediately after they occur |
| **26** | CogC4.2 | Identifies concrete concepts |
| **27** | CogC3.2 | Classifies according to physical attribute |
| **28** | CogC3.3 | Discriminates between objects or people using common attributes |
| **29** | CogC2.0 | Recognizes symbols |
| **30** | CogC2.1 | Uses object to represent another object |
| **31** | CogD1.0 | Uses object to obtain another object |
| **32** | CogC1.0 | Maintains search for object not in its usual location |
| **33** | CogB2.0 | Imitates novel words |
| **34** | CogC1.1 | Locates object in second of two hiding places |
| **35** | CogB1.1 | Imitates novel simple motor action not already in repertoire |
| **36** | CogD2.0 | Coordinates actions with objects to achieve new outcomes |
| **37** | CogB2.1 | Imitates novel vocalizations |
| **38** | CogB1.0 | Imitates novel coordinated motor actions |
| **39** | CogE1.2 | Uses senses to explore |
| **40** | CogD1.1 | Uses part of object or support to obtain another object |
| **41** | CogD2.1 | Tries different simple actions to achieve goal |

| | | |
|---|---|---|
| **42** | CogC1.2 | Locates hidden object |
| **43** | CogB2.2 | Imitates familiar vocalizations |
| **44** | CogB1.2 | Imitates familiar simple motor action |
| **45** | CogD1.2 | Retains one object when second object is obtained |
| **46** | CogD2.2 | Uses simple actions on objects |
| **47** | CogA2.0 | Combines simple actions to examine people, animals, and objects |
| **48** | CogA1.0 | Orients to events or stimulation |
| **49** | CogA1.1 | Reacts to events or stimulation |
| **50** | CogA2.1 | Uses sensory means to explore people, animals, and objects |

Fine Motor Area Item (display by the item difficulty followed the dose from hard to easy)

| No. | Item | Content |
|---|---|---|
| **1** | FMC1.0 | Holds writing tool using three-finger grasp to write or draw |
| **2** | FMC1.1 | Writes or draws using mixed strokes |
| **3** | FMC1.2 | Writes or draws using curved lines |
| **4** | FMC1.3 | Writes or draws using straight lines |
| **5** | FMB3.1 | Assembles toy |
| **6** | FMB3.0 | Manipulates object with two hands, each performing different action |
| **7** | FMD1.0 | Uses finger to interact with electronic device |
| **8** | FMB3.2 | Aligns objects |
| **9** | FMD1.1 | Uses finger to interact with simple electronic game |
| **10** | FMB3.3 | Fits variety of shapes into corresponding spaces |
| **11** | FMB3.4 | Holds object with one hand and manipulates object or produces action with other hand |
| **12** | FMD1.2 | Uses finger to interact with touch screen |
| **13** | FMA3.0 | Stacks objects |
| **14** | FMB2.0 | Rotates wrist to manipulate object |
| **15** | FMC1.4 | Scribbles |
| **16** | FMA3.1 | Releases object into targeted space |
| **17** | FMB2.1 | Turns object using either hand |
| **18** | FMB1.0 | Activates object with finger |
| **19** | FMA2.0 | Grasps pea-size object |
| **20** | FMB1.1 | Uses finger to point or touch |
| **21** | FMB3.5 | Transfers object from hand to hand |
| **22** | FMA2.1 | Grasps hand-size object |
| **23** | FMA3.2 | Releases object into nondefined space |
| **24** | FMA2.3 | Grasps pea-size object using fingers in raking or scratching movement |
| **25** | FMB1.3 | Uses fingers to explore object |
| **26** | FMB1.2 | Uses hand to activate object |
| **27** | FMA2.2 | Grasps small cylindrical object |
| **28** | FMA2.4 | Grasps hand-size object using whole hand |
| **29** | FMA1.1 | Brings hands together near midline |

| | | |
|---|---|---|
| **30** | FMA1.2 | Makes directed movements with arms |
| **31** | FMA1.0 | Makes directed batting or swiping movements with each hand |

Gross Motor Area Item (display by the item difficulty followed the dose from hard to easy)

| NO. | Item | Content |
|---|---|---|
| **1** | GMB7.0 | Skips |
| **2** | GMC3.0 | Rides and steers bicycle |
| **3** | GMC3.1 | Pedals and steers bicycle with training wheels |
| **4** | GMC2.1 | Moves swing back and forth |
| **5** | GMB7.1 | Gallops |
| **6** | GMB7.2 | Hops forward on one foot |
| **7** | GMC1.1 | Bounces ball with one hand |
| **8** | GMC2.0 | Uses hands to hang on play equipment with bars |
| **9** | GMC1.2 | Bounces ball with two hands |
| **10** | GMC3.2 | Pedals and steers tricycle |
| **11** | GMC1.0 | Swings bat, club, or stick to strike stationary object |
| **12** | GMC1.3 | Catches ball |
| **13** | GMB4.0 | Alternates feet going up and down stairs |
| **14** | GMC1.5 | Throws ball overhand at target with one hand |
| **15** | GMB6.0 | Jumps forward |
| **16** | GMC2.2 | Climbs play equipmen |
| **17** | GMB6.1 | Jumps up and down in place |
| **18** | GMC1.4 | Kicks ball |
| **19** | GMB6.2 | Jumps down from low structure |
| **20** | GMC3.3 | Pushes riding toy with feet while steering |
| **21** | GMB5.0 | Runs while avoiding people, furniture, or other objects |
| **22** | GMC1.6 | Throws or rolls ball at target with two hands |
| **23** | GMB6.3 | Jumps down with support |
| **24** | GMB5.1 | Runs |
| **25** | GMB4.1 | Walks up and down stairs using support |
| **26** | GMB5.2 | Walks fast |
| **27** | GMB3.0 | Walks avoiding people, furniture, or objects |
| **28** | GMC2.3 | Goes down small slide |
| **29** | GMB3.1 | Walks without support |
| **30** | GMC3.4 | Sits on riding toy or in wagon while in motion |
| **31** | GMB4.2 | Moves up and down stairs |
| **32** | GMA5.0 | Gets out of chair |
| **33** | GMB2.0 | Stoops and regains balanced standing position |
| **34** | GMB2.1 | Rises from sitting to standing position |
| **35** | GMB3.2 | Walks with one-hand support |
| **36** | GMA5.1 | Sits down in chair |

| | | |
|---|---|---|
| **37** | GMB2.2 | Stands unsupported |
| **38** | GMB4.3 | Gets up and down from low structure |
| **39** | GMB3.3 | Walks with two-hand support |
| **40** | GMB3.4 | Cruises |
| **41** | GMB2.3 | Pulls to standing position |
| **42** | GMB2.4 | Pulls to kneeling position |
| **43** | GMB1.0 | Creeps forward using alternating arm and leg movements |
| **44** | GMA5.2 | Maintains sitting position in chair |
| **45** | GMB1.1 | Rocks while in creeping position |
| **46** | GMA4.1 | Assumes hands-and-knees position from sitting |
| **47** | GMB1.2 | Assumes creeping position |
| **48** | GMB1.3 | Crawls forward on stomach |
| **49** | GMA4.2 | Regains balanced, upright sitting position after reaching across body |
| **50** | GMA4.0 | Assumes balanced sitting position |
| **51** | GMA4.3 | Regains balanced, upright sitting position after leaning left, right, and forward |
| **52** | GMA4.4 | Sits balanced without support |
| **53** | GMB1.4 | Pivots on stomach |
| **54** | GMA4.5 | Sits balanced using hands for support |
| **55** | GMA3.0 | Rolls from back to stomach |
| **56** | GMA2.0 | Puts weight on one hand or arm while reaching with opposite hand |
| **57** | GMA2.1 | Remains propped on extended arms with head lifted |
| **58** | GMA3.1 | Rolls from stomach to back |
| **59** | GMA3.2 | Rolls from back or stomach to side |
| **60** | GMA2.2 | Remains propped on nonextended forearms with head lifted |
| **61** | GMA4.6 | Holds head in midline when sitting supported |
| **62** | GMA1.0 | Turns head, moves arms, and kicks legs independently of each other |
| **63** | GMA1.1 | Kicks legs |
| **64** | GMA1.2 | Waves arms |
| **65** | GMA1.3 | Turns head side to side |

Social-Communication Area Item (display by the item difficulty followed the dose from hard to easy)

| | | |
|---|---|---|
| **1** | **SCC2.1** | **Uses irregular plural nouns in multiple-word sentences** |
| **2** | SCC3.1 | Uses irregular past tense of common verbs |
| **3** | SCB3.0 | Follows multistep directions without contextual cues |
| **4** | SCB4.0 | Responds to comprehension questions related to why, how, and when |
| **5** | SCC3.2 | Uses regular past tense of common verbs |
| **6** | SCC2.0 | Uses plural pronouns to indicate subjects, objects, and possession in multiple-word sentences |
| **7** | SCC4.0 | Asks questions using inverted auxiliary |
| **8** | SCD2.0 | Provides and seeks information while conversing using words, phrases, or sentences |
| **9** | SCC3.0 | Uses helping verbs |

| | | |
|---|---|---|
| **10** | SCD3.0 | Uses conversational rules when communicating with others |
| **11** | SCC3.3 | Uses to be verbs |
| **12** | SCD3.3 | Responds to topic initiations from others |
| **13** | SCC4.1 | Asks wh- questions |
| **14** | SCD2.2 | Describes objects, people, and events as part of social exchange |
| **15** | SCD2.1 | Asks questions to obtain information |
| **16** | SCD3.5 | Responds to contingent questions from others |
| **17** | SCC2.2 | Uses regular plural nouns |
| **18** | SCD3.4 | Alternates between speaker and listener roles during conversations with others |
| **19** | SCB3.1 | Follows multistep directions with contextual cues |
| **20** | SCD3.2 | Varies voice to impart meaning and recognize social or environmental conditions |
| **21** | SCD1.0 | Uses language to initiate and sustain social interaction |
| **22** | SCB4.1 | Answers who, what, and where questions |
| **23** | SCD1.1 | Follows social conventions of language |
| **24** | SCC1.0 | Produces multiple-word sentences to communicate |
| **25** | SCD3.1 | Uses socially appropriate physical orientation |
| **26** | SCB3.2 | Follows one-step direction without contextual cues |
| **27** | SCC1.1 | Uses two-word utterances |
| **28** | SCC1.2 | Uses 50 single words, signs, or symbols |
| **29** | SCC1.4 | Uses consistent consonant–vowel combinations |
| **30** | SCC1.3 | Uses consistent approximations for words or signs |
| **31** | SCB3.3 | Follows one-step direction with contextual cues |
| **32** | SCB2.0 | Locates common objects, people, or events |
| **33** | SCB2.2 | Responds to single-word directive |
| **34** | SCA4.1 | Makes requests of others |
| **35** | SCB1.0 | Follows gaze to establish joint attention |
| **36** | SCB1.1 | Follows pointing gestures with eyes |
| **37** | SCA4.2 | Makes choices to express preferences |
| **38** | SCA3.0 | Engages in vocal exchanges |
| **39** | SCA4.0 | Uses intentional gestures, vocalizations, and objects to communicate |
| **40** | SCB2.1 | Recognizes own and familiar names |
| **41** | SCA3.1 | Vocalizes to another person expressing positive affective state |
| **42** | SCA3.2 | Vocalizes to another person expressing negative affective state |
| **43** | SCA2.0 | Produces speech sounds |
| **44** | SCA4.3 | Expresses desire to continue activity |
| **45** | SCA1.0 | Turns and looks toward person speaking |
| **46** | SCA2.1 | Coos and gurgles |
| **47** | SCB1.2 | Looks toward object |
| **48** | SCA4.4 | Expresses negation or protests |
| **49** | SCA1.1 | Quiets to familiar voice |

Social-Emotional Area Item (display by the item difficulty followed the dose from hard to easy)

| | | |
|---|---|---|
| **1** | **E4.1** | **States birthday** |
| **2** | D4 | Resolves conflicts using negotiation |
| **3** | D4.1 | Uses strategies to resolve conflicts |
| **4** | E4 | Relates identifying information about self |
| **5** | C4 | Maintains engagement in games with rules |
| **6** | C4.1 | Knows and follows game rules |
| **7** | C2 | Plans and acts out recognizable event, theme, or storyline in imaginary play |
| **8** | C3.1 | Initiates cooperative activity |
| **9** | C3 | Maintains cooperative activity |
| **10** | B3.1 | Explains or shows others how to do tasks mastered |
| **11** | C2.1 | Enacts roles or identities in imaginary play |
| **12** | C4.2 | Participates in game |
| **13** | B3 | Makes positive statements about self or accomplishments |
| **14** | E3.1 | Seeks adult permission when appropriate |
| **16** | B1.2 | Identifies/labels own emotions |
| **17** | B1.1 | Identifies/labels emotions in others |
| **18** | C3.2 | Joins others in cooperative activity |
| **19** | E1. | Meets observable physical needs in socially appropriate ways |
| **20** | B3.2 | Shares accomplishment with familiar caregiver |
| **21** | C3.3 | Shares or exchanges objects |
| **22** | E4.2 | States age |
| **23** | E3. | Follows context-specific rules |
| **24** | D2.2 | Responds appropriately to directions during large-group activities |
| **25** | D3 | Initiates and completes independent activities |
| **26** | E4.3 | Provides given name or nickname of self and others |
| **27** | D2. | Interacts appropriately with others during large-group activities |
| **28** | D3.1 | Responds to request to finish activity |
| **29** | D1 | Interacts appropriately with others during small-group activities |
| **30** | B2 | Uses appropriate strategies to manage emotional states |
| **31** | E2.2 | Adjusts behavior based on feedback from others or environment |
| **32** | D1.2 | Responds appropriately to directions during small-group activities |
| **33** | B1 | Responds appropriately to others' emotions |
| **34** | B2.1 | Responds appropriately to soothing by peer |
| **35** | D2.1 | Interacts appropriately with materials during large-group activities |
| **36** | C2.2 | Uses imaginary props in play |
| **37** | E3.2 | Follows established social rules in familiar environments |
| **38** | C1 | Maintains interaction with peer |
| **39** | E2. | Meets accepted social norms in community settings |
| **40** | D1.1 | Interacts appropriately with materials during small-group activities |
| **41** | E2.1 | Meets behavioral expectations in familiar environments |
| **42** | C1.1 | Initiates social behavior toward peer |
| **43** | D3.2 | Responds to request to begin activity |
| **44** | D4.2 | Claims and defends possessions |

| | | |
|---|---|---|
| **45** | D2.3 | Remains with group during large-group activities |
| **46** | C1.2 | Responds appropriately to peer social behavior |
| **47** | D1.3 | Remains with group during small-group activities |
| **48** | E1.1 | Meets internal physical needs of hunger and thirst |
| **49** | A3.1 | Initiates next step of familiar social routine |
| **50** | D3.3 | Entertains self by playing with toys |
| **51** | A3. | Participates in familiar social routines with caregivers |
| **52** | A3.2 | Follows familiar social routines with familiar adults |
| **53** | C1.3 | Plays near one or two peers |
| **54** | A1 | Initiates positive social behavior toward familiar adult |
| **55** | A2.2 | Repeats part of interactive game or action in order to continue game or action |
| **56** | B2.3 | Responds appropriately to soothing by adult |
| **57** | B2.2 | Seeks comfort, closeness, or physical contact from familiar adult |
| **58** | A2 | Maintains social interaction with familiar adult |
| **59** | A2.1 | Initiates simple social interaction with familiar adult |
| **60** | A2.3 | Responds to familiar game or action |
| **61** | A1.1 | Responds appropriately to familiar adult's affective tone |
| **62** | A1.2 | Responds to familiar adult's positive social behavior |

## APPENDIX 5. ITEM CALIBRATION AND FIT RESULTS BY DEVELOPMENTAL AREA FOR THE AEPS-2

Adaptive area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **adapt_B1.0** | 3.89 | 0.02 | 0.9043 | -5.3491 | 1.0844 | 1.6911 |
| **adapt_B1.1** | 3.5 | 0.02 | 0.9312 | -4.3291 | 0.6212 | -9.4294 |
| **adapt_C1.1** | 2.45 | 0.02 | 1.0389 | 3.101 | 0.8278 | -4.4092 |
| **adapt_A5.0** | 2.33 | 0.01 | 0.9072 | -7.9091 | 0.9466 | -1.3191 |
| **adapt_A5.1** | 2.18 | 0.01 | 0.9326 | -5.8391 | 0.7241 | -7.5593 |
| **adapt_B2.0** | 2.04 | 0.01 | 0.7869 | -9.8992 | 0.8685 | -3.4491 |
| **adapt_C1.2** | 1.99 | 0.01 | 0.9608 | -3.449 | 0.713 | -8.0593 |
| **adapt_B1.2** | 1.97 | 0.01 | 1.1588 | 9.9012 | 3.0756 | 9.9031 |
| **adapt_B3.0** | 1.79 | 0.01 | 0.9056 | -8.6791 | 1.3521 | 8.0814 |
| **adapt_A5.2** | 1.68 | 0.01 | 1.0312 | 2.771 | 0.8081 | -5.2392 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **adapt_C1.3** | 1.43 | 0.01 | 1.0377 | 3.411 | 0.8372 | -4.3592 |
| **adapt_A3.0** | 1.34 | 0.01 | 1.0246 | 2.251 | 1.0594 | 1.5011 |
| **adapt_B2.1** | 1.13 | 0.01 | 0.8242 | -9.8992 | 0.8857 | -3.2391 |
| **adapt_A4.0** | 0.86 | 0.01 | 0.7632 | -9.8992 | 0.8619 | -4.2491 |
| **adapt_A3.1** | 0.83 | 0.01 | 0.8846 | -9.8991 | 0.738 | -8.5193 |
| **adapt_C1.0** | 0.83 | 0.01 | 0.8396 | -9.8992 | 2.32 | 9.9023 |
| **adapt_A4.1** | 0.21 | 0.01 | 0.7484 | -9.8993 | 0.6564 | -9.8993 |
| **adapt_B3.1** | 0.08 | 0.01 | 1.3436 | 9.9013 | 2.0228 | 9.902 |
| **adapt_C1.5** | -0.41 | 0.01 | 1.2908 | 9.9013 | 1.0776 | 2.8811 |
| **adapt_A3.2** | -0.76 | 0.02 | 0.993 | -0.539 | 0.7898 | -8.2992 |
| **adapt_A2.0** | -0.86 | 0.02 | 0.9239 | -5.9791 | 0.7876 | -8.2592 |
| **adapt_C1.4** | -0.87 | 0.02 | 1.535 | 9.9015 | 1.374 | 9.9014 |
| **adapt_C1.6** | -1.55 | 0.02 | 1.5337 | 9.9015 | 1.1754 | 4.9412 |
| **adapt_A2.1** | -1.62 | 0.02 | 0.74 | -9.8993 | 0.5979 | -9.8994 |
| **adapt_A1.1** | -1.9 | 0.02 | 1.2225 | 9.9012 | 0.9115 | -2.3891 |
| **adapt_A4.2** | -2.21 | 0.02 | 0.7399 | -9.8993 | 0.4594 | -9.8995 |
| **adapt_A2.2** | -2.23 | 0.02 | 0.7445 | -9.8993 | 0.543 | -9.8995 |
| **adapt_A1.0** | -2.97 | 0.02 | 1.2414 | 9.9012 | 1.9733 | 9.902 |
| **adapt_A1.3** | -3.1 | 0.02 | 0.8559 | -8.2891 | 0.8589 | -2.6391 |
| **adapt_A1.2** | -3.21 | 0.02 | 0.8644 | -7.6691 | 0.7306 | -5.1693 |
| **adapt_A4.3** | -3.3 | 0.02 | 1.1548 | 7.8812 | 1.8808 | 9.9019 |
| **adapt_A1.4** | -5.53 | 0.03 | 1.4757 | 9.9015 | 3.6168 | 9.9036 |

Social-communication area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **sc_D3.1** | 4.49 | 0.03 | 0.7746 | -9.5692 | 0.3916 | -9.8996 |
| **sc_D3.3** | 4.41 | 0.03 | 0.7333 | -9.8993 | 0.3052 | -9.8997 |
| **sc_D3.4** | 4.13 | 0.03 | 0.7528 | -9.8992 | 0.325 | -9.8997 |
| **sc_D3.2** | 4.09 | 0.03 | 0.8631 | -6.1491 | 0.3852 | -9.8996 |
| **sc_D3.0** | 3.95 | 0.03 | 0.6914 | -9.8993 | 0.4656 | -9.8995 |
| **sc_D2.3** | 3.19 | 0.02 | 0.6821 | -9.8993 | 0.3374 | -9.8997 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **sc_D2.4** | 3.07 | 0.02 | 0.692 | -9.8993 | 0.3819 | -9.8996 |
| **sc_D2.1** | 2.99 | 0.02 | 0.6741 | -9.8993 | 0.4594 | -9.8995 |
| **sc_D2.2** | 2.85 | 0.02 | 0.6952 | -9.8993 | 0.3812 | -9.8996 |
| **sc_D2.5** | 2.76 | 0.02 | 0.6891 | -9.8993 | 0.3726 | -9.8996 |
| **sc_D2.6** | 2.69 | 0.02 | 0.7878 | -9.8992 | 0.456 | -9.8995 |
| **sc_D2.0** | 2.55 | 0.02 | 0.5726 | -9.8994 | 0.4995 | -9.8995 |
| **sc_D1.1** | 2.42 | 0.02 | 0.8514 | -9.6891 | 0.6113 | -8.9594 |
| **sc_C2.0** | 2.11 | 0.02 | 1.5742 | 9.9016 | 1.9824 | 9.902 |
| **sc_D1.3** | 1.97 | 0.02 | 0.9102 | -6.2491 | 0.5927 | -9.8994 |
| **sc_D1.2** | 1.9 | 0.02 | 0.8204 | -9.8992 | 0.6405 | -9.1894 |
| **sc_C2.1** | 1.49 | 0.02 | 1.2785 | 9.9013 | 1.2033 | 4.6712 |
| **sc_C1.1** | 1.29 | 0.02 | 1.0271 | 2.041 | 0.9963 | -0.079 |
| **sc_C1.0** | 1.18 | 0.02 | 1.177 | 9.9012 | 1.5592 | 9.9016 |
| **sc_D1.0** | 0.95 | 0.02 | 0.8092 | -9.8992 | 1.1817 | 4.7712 |
| **sc_D1.4** | 0.94 | 0.02 | 0.8147 | -9.8992 | 0.6499 | -9.8994 |
| **sc_C1.2** | 0.28 | 0.01 | 0.8383 | -9.8992 | 0.7805 | -7.5592 |
| **sc_D1.5** | 0.15 | 0.01 | 1.1522 | 9.9012 | 0.905 | -3.1991 |
| **sc_C2.2** | 0.12 | 0.01 | 0.9568 | -3.749 | 1.0318 | 1.041 |
| **sc_B2.0** | -0.15 | 0.01 | 0.8843 | -9.8991 | 0.943 | -1.9591 |
| **sc_B1.1** | -0.48 | 0.01 | 0.8291 | -9.8992 | 0.8683 | -4.7991 |
| **sc_B1.0** | -0.59 | 0.01 | 1.0026 | 0.231 | 1.3337 | 9.9013 |
| **sc_C1.3** | -0.6 | 0.01 | 0.828 | -9.8992 | 0.8258 | -6.4692 |
| **sc_B2.1** | -0.83 | 0.01 | 0.9718 | -2.499 | 1.1023 | 3.4411 |
| **sc_C2.3** | -0.98 | 0.01 | 0.8611 | -9.8991 | 1.0968 | 3.2411 |
| **sc_B1.2** | -1.01 | 0.01 | 0.8302 | -9.8992 | 0.7866 | -7.9092 |
| **sc_B1.3** | -1.64 | 0.01 | 0.8208 | -9.8992 | 1.0919 | 2.9111 |
| **sc_A2.0** | -1.73 | 0.01 | 1.5962 | 9.9016 | 3.657 | 9.9037 |
| **sc_B2.2** | -1.79 | 0.01 | 1.0723 | 6.0211 | 1.2615 | 7.6713 |
| **sc_A2.1** | -2.25 | 0.02 | 1.0511 | 4.1211 | 1.988 | 9.902 |
| **sc_B1.4** | -3.11 | 0.02 | 1.0847 | 6.1311 | 1.4612 | 9.9015 |
| **sc_A3.0** | -3.15 | 0.02 | 0.984 | -1.189 | 1.5172 | 9.9015 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **sc_C1.4** | -3.23 | 0.02 | 0.9859 | -1.029 | 4.2817 | 9.9043 |
| **sc_B2.3** | -3.36 | 0.02 | 1.3488 | 9.9013 | 2.3752 | 9.9024 |
| **sc_B2.4** | -3.48 | 0.02 | 1.423 | 9.9014 | 2.7481 | 9.9027 |
| **sc_C1.5** | -4.21 | 0.02 | 1.3505 | 9.9014 | 9.9 | 9.9099 |
| **sc_A1.0** | -4.39 | 0.02 | 1.1457 | 8.1111 | 5.9719 | 9.906 |
| **sc_A3.1** | -4.41 | 0.02 | 1.1191 | 6.6511 | 1.5513 | 8.9616 |
| **sc_A1.1** | -4.54 | 0.02 | 1.0579 | 3.1911 | 5.8057 | 9.9058 |
| **sc_A2.2** | -4.88 | 0.02 | 1.0408 | 2.071 | 2.5758 | 9.9026 |
| **sc_A1.2** | -5.17 | 0.02 | 0.8949 | -5.1491 | 2.2946 | 9.9023 |

Social Area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **soc_C2.0** | 3.26 | 0.02 | 0.8493 | -9.8992 | 0.7406 | -7.1393 |
| **soc_C2.1** | 2.82 | 0.02 | 0.8023 | -9.8992 | 0.6715 | -9.8993 |
| **soc_C1.0** | 2.42 | 0.02 | 0.8085 | -9.8992 | 0.7941 | -6.6392 |
| **soc_C1.1** | 2.01 | 0.01 | 0.7887 | -9.8992 | 0.682 | -9.8993 |
| **soc_B1.0** | 1.9 | 0.01 | 1.1426 | 9.9011 | 1.2612 | 8.1813 |
| **soc_C2.2** | 1.72 | 0.01 | 0.9249 | -7.0691 | 0.8931 | -3.8591 |
| **soc_C1.2** | 1.53 | 0.01 | 0.7705 | -9.8992 | 0.6748 | -9.8993 |
| **soc_A3.0** | 1.4 | 0.01 | 0.9287 | -6.8191 | 0.8366 | -6.6792 |
| **soc_B2.0** | 1.33 | 0.01 | 1.0837 | 7.6311 | 0.9994 | -0.019 |
| **soc_B1.1** | 0.73 | 0.01 | 1.0999 | 9.1211 | 1.1967 | 8.1112 |
| **soc_A3.1** | 0.71 | 0.01 | 0.8382 | -9.8992 | 0.7488 | -9.8993 |
| **soc_B2.1** | 0.29 | 0.01 | 0.946 | -5.1191 | 0.8577 | -7.1091 |
| **soc_A2.1** | 0.11 | 0.01 | 0.9509 | -4.599 | 0.8491 | -7.6292 |
| **soc_A2.0** | 0 | 0.01 | 0.9848 | -1.399 | 0.8797 | -5.9891 |
| **soc_B1.2** | -0.41 | 0.01 | 1.5232 | 9.9015 | 2.3528 | 9.9024 |
| **soc_C1.3** | -0.49 | 0.01 | 1.0581 | 5.0311 | 0.9152 | -3.8391 |
| **soc_A3.2** | -0.53 | 0.02 | 0.9823 | -1.559 | 0.9386 | -2.7191 |
| **soc_A2.2** | -1.15 | 0.02 | 0.8427 | -9.8992 | 0.7097 | -9.8993 |
| **soc_C1.4** | -1.17 | 0.02 | 1.1413 | 9.9011 | 0.9208 | -2.7691 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **soc_C1.5** | -1.32 | 0.02 | 1.1861 | 9.9012 | 1.1519 | 4.7112 |
| **soc_A1.0** | -2.11 | 0.02 | 1.1015 | 6.9611 | 1.0811 | 2.0211 |
| **soc_A1.1** | -2.6 | 0.02 | 0.9241 | -5.0391 | 0.6862 | -7.7793 |
| **soc_A1.2** | -2.77 | 0.02 | 1.1293 | 7.8011 | 1.2766 | 5.3913 |
| **soc_A2.3** | -3.1 | 0.02 | 1.2213 | 9.9012 | 1.0116 | 0.241 |
| **soc_A1.3** | -4.58 | 0.03 | 1.3948 | 9.9014 | 1.4151 | 5.3414 |

Cognitive area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **cog_G5.0** | 3.88 | 0.02 | 0.9748 | -1.279 | 0.6121 | -8.3094 |
| **cog_G4.1** | 3.77 | 0.02 | 0.8701 | -7.1091 | 0.5377 | -9.8995 |
| **cog_G6.0** | 3.7 | 0.02 | 0.9968 | -0.169 | 1.0665 | 1.2511 |
| **cog_G5.1** | 3.67 | 0.02 | 1.1153 | 6.0311 | 0.6406 | -7.8294 |
| **cog_G6.1** | 3.5 | 0.02 | 0.9876 | -0.709 | 0.7504 | -5.2992 |
| **cog_G2.0** | 3.33 | 0.02 | 0.9626 | -2.269 | 0.7008 | -6.6593 |
| **cog_G1.0** | 3.3 | 0.02 | 0.9293 | -4.3791 | 0.7662 | -5.0892 |
| **cog_G1.1** | 3.15 | 0.02 | 0.9415 | -3.7491 | 0.8818 | -2.4991 |
| **cog_G1.2** | 3.02 | 0.02 | 0.9054 | -6.3791 | 0.7124 | -6.6593 |
| **cog_G4.2** | 2.84 | 0.02 | 0.8085 | -9.8992 | 0.599 | -9.8994 |
| **cog_G4.0** | 2.81 | 0.02 | 1.0931 | 6.2411 | 1.4731 | 8.9215 |
| **cog_G5.2** | 2.77 | 0.02 | 1.302 | 9.9013 | 0.945 | -1.1991 |
| **cog_G3.0** | 2.72 | 0.02 | 0.9896 | -0.729 | 0.8443 | -3.5692 |
| **cog_G6.2** | 2.71 | 0.02 | 1.0344 | 2.401 | 0.9478 | -1.1391 |
| **cog_G2.1** | 2.5 | 0.02 | 0.9667 | -2.479 | 0.7805 | -5.3692 |
| **cog_F1.0** | 1.99 | 0.01 | 1.1677 | 9.9012 | 1.3141 | 7.2313 |
| **cog_G1.3** | 1.93 | 0.01 | 0.9302 | -5.8891 | 0.7057 | -8.3993 |
| **cog_D2.0** | 1.81 | 0.01 | 0.8793 | -9.8991 | 0.8516 | -4.0891 |
| **cog_G3.1** | 1.76 | 0.01 | 0.8151 | -9.8992 | 0.6749 | -9.7593 |
| **cog_E4.0** | 1.15 | 0.01 | 0.8763 | -9.8991 | 0.8375 | -5.3592 |
| **cog_D2.1** | 1.12 | 0.01 | 0.9867 | -1.199 | 1.0838 | 2.5711 |
| **cog_F1.1** | 1.11 | 0.01 | 0.9013 | -9.2291 | 0.7938 | -6.9892 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **cog_E2.0** | 0.94 | 0.01 | 1.147 | 9.9011 | 1.1333 | 4.2311 |
| **cog_G4.3** | 0.9 | 0.01 | 1.2271 | 9.9012 | 1.6414 | 9.9016 |
| **cog_D1.0** | 0.76 | 0.01 | 0.8652 | -9.8991 | 0.842 | -5.7992 |
| **cog_B3.0** | 0.64 | 0.01 | 1.0521 | 4.7011 | 1.0384 | 1.371 |
| **cog_E4.1** | 0.55 | 0.01 | 0.7998 | -9.8992 | 0.7765 | -8.8592 |
| **cog_D2.2** | 0.4 | 0.01 | 0.9858 | -1.299 | 0.9746 | -0.949 |
| **cog_E2.1** | 0.34 | 0.01 | 1.0572 | 5.1311 | 0.9996 | -0.009 |
| **cog_B2.0** | 0.05 | 0.01 | 1.0965 | 8.4211 | 1.1039 | 3.9911 |
| **cog_E3.0** | -0.16 | 0.01 | 0.9739 | -2.329 | 0.8585 | -5.9291 |
| **cog_F1.2** | -0.31 | 0.01 | 0.736 | -9.8993 | 0.6309 | -9.8994 |
| **cog_C1.0** | -0.35 | 0.01 | 0.9829 | -1.499 | 0.9622 | -1.509 |
| **cog_C2.0** | -0.46 | 0.01 | 0.952 | -4.209 | 0.9049 | -3.8591 |
| **cog_B3.1** | -0.63 | 0.01 | 0.9194 | -7.0091 | 0.7634 | -9.8992 |
| **cog_D1.1** | -0.73 | 0.01 | 0.9215 | -6.7491 | 0.8567 | -5.6991 |
| **cog_E1.0** | -0.79 | 0.01 | 1.209 | 9.9012 | 1.0996 | 3.6211 |
| **cog_E3.1** | -0.85 | 0.01 | 0.8565 | -9.8991 | 0.7338 | -9.8993 |
| **cog_B2.1** | -1.17 | 0.01 | 0.8287 | -9.8992 | 0.7365 | -9.8993 |
| **cog_E3.2** | -1.31 | 0.02 | 0.9042 | -7.7191 | 0.7474 | -9.1593 |
| **cog_C2.1** | -1.59 | 0.02 | 1.0714 | 5.2411 | 1.0781 | 2.3311 |
| **cog_C1.1** | -1.68 | 0.02 | 0.8101 | -9.8992 | 0.6817 | -9.8993 |
| **cog_B2.2** | -1.71 | 0.02 | 0.812 | -9.8992 | 0.6662 | -9.8993 |
| **cog_E1.1** | -1.74 | 0.02 | 0.9284 | -5.3991 | 0.8356 | -5.0692 |
| **cog_F1.3** | -1.86 | 0.02 | 1.1047 | 7.3311 | 1.0676 | 1.8711 |
| **cog_C1.2** | -2.51 | 0.02 | 0.8196 | -9.8992 | 0.6699 | -8.5993 |
| **cog_B2.3** | -2.62 | 0.02 | 0.9738 | -1.709 | 0.9415 | -1.3091 |
| **cog_F1.4** | -2.65 | 0.02 | 1.3925 | 9.9014 | 2.1169 | 9.9021 |
| **cog_E1.2** | -2.85 | 0.02 | 1.0798 | 4.8811 | 1.0295 | 0.621 |
| **cog_C1.3** | -3.03 | 0.02 | 0.9495 | -3.1191 | 0.8832 | -2.3891 |
| **cog_B1.0** | -3.3 | 0.02 | 1.1614 | 8.9812 | 1.5457 | 8.6915 |
| **cog_B1.1** | -3.72 | 0.02 | 1.0597 | 3.1911 | 1.8232 | 9.9018 |
| **cog_A1.0** | -4.24 | 0.02 | 1.0031 | 0.161 | 3.4812 | 9.9035 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **cog_A1.1** | -4.4 | 0.02 | 1.0935 | 4.2411 | 3.3558 | 9.9034 |
| **cog_B1.2** | -4.54 | 0.03 | 1.2642 | 9.9013 | 3.423 | 9.9034 |
| **cog_A1.2** | -4.57 | 0.03 | 0.9489 | -2.3091 | 2.3842 | 9.9024 |
| **cog_A1.3** | -4.63 | 0.03 | 0.986 | -0.609 | 2.426 | 9.9024 |
| **cog_A1.4** | -4.74 | 0.03 | 1.0627 | 2.6111 | 3.4792 | 9.9035 |

Fine motor area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **fm_B5.0** | 5.87 | 0.02 | 0.9689 | -1.709 | 1.3206 | 5.9213 |
| **fm_B5.1** | 4.67 | 0.02 | 0.9503 | -3.679 | 0.8519 | -3.6491 |
| **fm_B4.0** | 3.46 | 0.01 | 0.9332 | -5.8091 | 0.936 | -1.7491 |
| **fm_B2.0** | 3.21 | 0.01 | 0.7519 | -9.8992 | 0.6592 | -9.8993 |
| **fm_B2.1** | 2.78 | 0.01 | 0.7405 | -9.8993 | 0.6445 | -9.8994 |
| **fm_A5.1** | 2.73 | 0.01 | 0.9027 | -8.7591 | 0.6546 | -9.8993 |
| **fm_B4.2** | 2.67 | 0.01 | 1.0419 | 3.601 | 0.9357 | -1.8791 |
| **fm_A5.0** | 2.6 | 0.01 | 0.77 | -9.8992 | 0.6985 | -9.7393 |
| **fm_B4.1** | 2.4 | 0.01 | 1.0783 | 6.6111 | 1.1722 | 4.7412 |
| **fm_A5.2** | 1.64 | 0.01 | 0.7705 | -9.8992 | 0.5487 | -9.8995 |
| **fm_B2.2** | 1.63 | 0.01 | 0.9627 | -3.099 | 0.804 | -6.0792 |
| **fm_B5.2** | 1.62 | 0.01 | 1.074 | 5.9511 | 0.8953 | -3.1291 |
| **fm_B1.0** | 1.61 | 0.01 | 1.4136 | 9.9014 | 1.9639 | 9.902 |
| **fm_B3.0** | 0.93 | 0.02 | 0.9142 | -6.7191 | 0.7965 | -5.9892 |
| **fm_A4.0** | 0.79 | 0.02 | 1.0441 | 3.261 | 0.8848 | -3.1991 |
| **fm_A4.1** | 0.53 | 0.02 | 0.99 | -0.719 | 0.7459 | -7.0993 |
| **fm_B1.1** | 0.51 | 0.02 | 1.216 | 9.9012 | 1.1753 | 4.2312 |
| **fm_A5.3** | 0.2 | 0.02 | 0.8512 | -9.8991 | 0.6126 | -9.8994 |
| **fm_A4.2** | -0.05 | 0.02 | 0.9144 | -5.8491 | 0.6998 | -7.7993 |
| **fm_B3.1** | -0.52 | 0.02 | 1.1135 | 6.8111 | 0.9724 | -0.569 |
| **fm_A3.0** | -0.62 | 0.02 | 1.3768 | 9.9014 | 1.2968 | 5.6613 |
| **fm_A3.1** | -1.26 | 0.02 | 1.2011 | 9.9012 | 1.0128 | 0.251 |
| **fm_A4.3** | -1.27 | 0.02 | 0.989 | -0.619 | 0.6599 | -7.0093 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **fm_A5.4** | -1.51 | 0.02 | 1.5641 | 9.9016 | 1.119 | 2.0211 |
| **fm_A2.0** | -2.18 | 0.02 | 1.0048 | 0.251 | 0.7735 | -3.8592 |
| **fm_A2.1** | -2.41 | 0.02 | 0.8914 | -5.4791 | 0.6101 | -6.9594 |
| **fm_A3.2** | -2.92 | 0.03 | 1.1135 | 5.0111 | 1.2164 | 2.9912 |
| **fm_A2.2** | -3.31 | 0.03 | 0.9396 | -2.6791 | 1.1915 | 2.6312 |
| **fm_A3.3** | -3.58 | 0.03 | 1.202 | 8.0412 | 1.53 | 6.6115 |
| **fm_A2.3** | -3.99 | 0.03 | 0.8336 | -7.1192 | 1.219 | 3.0012 |
| **fm_A1.0** | -4.91 | 0.03 | 1.1124 | 3.9111 | 1.4836 | 6.6215 |
| **fm_A1.1** | -4.91 | 0.03 | 0.9492 | -1.8491 | 1.5077 | 6.9115 |
| **fm_A1.2** | -6.4 | 0.04 | 1.2268 | 5.6012 | 2.9652 | 9.903 |

Gross motor area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **gm_D2.0** | 6.9 | 0.02 | 1.0041 | 0.251 | 2.8407 | 9.9028 |
| **gm_D3.1** | 5.84 | 0.02 | 0.9072 | -7.3391 | 8.3058 | 9.9083 |
| **gm_D1.0** | 5.61 | 0.02 | 0.8539 | -9.8991 | 0.8551 | -4.2291 |
| **gm_D1.1** | 4.85 | 0.01 | 0.8112 | -9.8992 | 1.4187 | 9.9014 |
| **gm_D1.2** | 4.65 | 0.01 | 0.8738 | -9.8991 | 0.5355 | -9.8995 |
| **gm_D3.0** | 4.45 | 0.01 | 0.9849 | -1.249 | 9.9 | 9.9099 |
| **gm_D3.2** | 4.15 | 0.01 | 0.8667 | -9.8991 | 2.4628 | 9.9025 |
| **gm_C4.0** | 4.11 | 0.01 | 0.789 | -9.8992 | 0.5814 | -9.8994 |
| **gm_C3.0** | 4.07 | 0.01 | 0.8956 | -8.7291 | 0.6232 | -9.8994 |
| **gm_D2.1** | 3.99 | 0.01 | 1.1686 | 9.9012 | 1.0997 | 2.7711 |
| **gm_D4.0** | 3.72 | 0.02 | 0.8449 | -9.8992 | 0.8077 | -5.7492 |
| **gm_C3.1** | 3.57 | 0.02 | 0.8122 | -9.8992 | 0.4775 | -9.8995 |
| **gm_D3.3** | 3.11 | 0.02 | 1.3428 | 9.9013 | 2.9375 | 9.9029 |
| **gm_D4.1** | 2.99 | 0.02 | 0.9658 | -2.369 | 0.7349 | -7.4793 |
| **gm_C4.1** | 2.95 | 0.02 | 0.8005 | -9.8992 | 0.5038 | -9.8995 |
| **gm_D3.4** | 2.85 | 0.02 | 1.7248 | 9.9017 | 3.6839 | 9.9037 |
| **gm_C3.2** | 2.43 | 0.02 | 0.6861 | -9.8993 | 0.3193 | -9.8997 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **gm_C4.2** | 2.24 | 0.02 | 0.8626 | -8.6391 | 0.5688 | -9.8994 |
| **gm_C1.0** | 2.19 | 0.02 | 0.7841 | -9.8992 | 0.59 | -9.8994 |
| **gm_D4.2** | 1.94 | 0.02 | 1.4896 | 9.9015 | 4.4807 | 9.9045 |
| **gm_B2.0** | 1.62 | 0.02 | 1.0297 | 1.611 | 0.8495 | -3.1192 |
| **gm_C2.0** | 1.59 | 0.02 | 0.7681 | -9.8992 | 0.3486 | -9.8997 |
| **gm_D2.2** | 1.51 | 0.02 | 2.2092 | 9.9022 | 4.1398 | 9.9041 |
| **gm_C1.1** | 1.44 | 0.02 | 0.7113 | -9.8993 | 0.252 | -9.8997 |
| **gm_C4.3** | 1.41 | 0.02 | 0.994 | -0.309 | 0.5587 | -9.8994 |
| **gm_B2.1** | 1.28 | 0.02 | 1.0364 | 1.901 | 0.79 | -4.1392 |
| **gm_C2.1** | 1.19 | 0.02 | 0.8337 | -9.1592 | 0.3396 | -9.8997 |
| **gm_C1.2** | 0.88 | 0.02 | 0.7197 | -9.8993 | 0.2149 | -9.8998 |
| **gm_C1.4** | 0.74 | 0.02 | 0.6872 | -9.8993 | 0.2273 | -9.8998 |
| **gm_C1.3** | 0.09 | 0.02 | 0.8424 | -7.7092 | 0.3046 | -9.8997 |
| **gm_C1.5** | -0.3 | 0.03 | 0.7469 | -9.8993 | 0.1986 | -9.8998 |
| **gm_B2.2** | -0.57 | 0.03 | 1.8327 | 9.9018 | 1.9168 | 9.9019 |
| **gm_C2.2** | -0.69 | 0.03 | 0.7614 | -9.8992 | 0.2162 | -9.8998 |
| **gm_C2.3** | -0.94 | 0.03 | 0.8183 | -7.9192 | 0.3991 | -9.8996 |
| **gm_A3.0** | -1.28 | 0.03 | 1.0876 | 3.4111 | 0.9275 | -1.0091 |
| **gm_B1.1** | -1.62 | 0.03 | 0.7586 | -9.8992 | 0.4005 | -9.8996 |
| **gm_A3.1** | -1.78 | 0.03 | 0.9915 | -0.329 | 0.8389 | -2.3892 |
| **gm_A3.2** | -1.99 | 0.03 | 0.9778 | -0.869 | 0.9641 | -0.499 |
| **gm_B1.0** | -2.07 | 0.03 | 0.9652 | -1.379 | 0.6517 | -5.7693 |
| **gm_A3.3** | -2.09 | 0.03 | 1.0876 | 3.3711 | 1.2924 | 3.9113 |
| **gm_B1.2** | -2.21 | 0.03 | 0.8174 | -7.6892 | 1.0732 | 1.0711 |
| **gm_B1.3** | -2.5 | 0.03 | 0.8426 | -6.5992 | 0.6637 | -5.8693 |
| **gm_B1.4** | -3.27 | 0.03 | 0.8719 | -5.3991 | 0.5987 | -8.2894 |
| **gm_A3.5** | -3.44 | 0.03 | 1.1553 | 6.0012 | 3.1236 | 9.9031 |
| **gm_A3.4** | -3.71 | 0.03 | 1.0778 | 3.0911 | 2.7314 | 9.9027 |
| **gm_B1.5** | -3.73 | 0.03 | 0.8329 | -7.2092 | 1.0242 | 0.481 |
| **gm_A2.0** | -4.25 | 0.03 | 0.9901 | -0.399 | 2.6698 | 9.9027 |
| **gm_A2.1** | -4.62 | 0.03 | 0.9844 | -0.629 | 4.3971 | 9.9044 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **gm_A2.2** | -4.98 | 0.03 | 0.9576 | -1.739 | 4.3641 | 9.9044 |
| **gm_A3.6** | -5.36 | 0.03 | 1.1112 | 4.3411 | 9.2547 | 9.9093 |
| **gm_B1.6** | -5.5 | 0.03 | 1.1365 | 5.2711 | 1.5727 | 9.9016 |
| **gm_A1.0** | -7.42 | 0.04 | 0.8187 | -6.1292 | 9.9 | 9.9099 |
| **gm_A1.1** | -7.82 | 0.04 | 0.72 | -8.7593 | 9.9 | 9.9099 |
| **gm_A1.3** | -8.07 | 0.04 | 0.9731 | -0.689 | 9.9 | 9.9099 |
| **gm_A1.2** | -8.14 | 0.04 | 0.9678 | -0.809 | 9.9 | 9.9099 |

## APPENDIX 6. ITEM CALIBRATION AND FIT RESULTS BY DEVELOPMENTAL AREA FOR THE AEPS-3

Adaptive area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **AdaptD4.0** | 3.98 | 0.2 | 0.7099 | -1.8493 | 1.0652 | 0.2911 |
| **AdaptD4.2** | 3.79 | 0.19 | 1.0187 | 0.171 | 0.3586 | -2.0296 |
| **AdaptD4.1** | 3.6 | 0.18 | 0.8982 | -0.6591 | 0.662 | -0.8693 |
| **AdaptC2.0** | 3.57 | 0.17 | 0.8008 | -1.3992 | 0.4678 | -1.6195 |
| **AdaptC2.1** | 3.51 | 0.17 | 0.5568 | -3.6194 | 0.3493 | -2.2197 |
| **AdaptD2.1** | 3.33 | 0.16 | 0.8547 | -1.0591 | 0.6737 | -0.8893 |
| **AdaptD3.0** | 3.21 | 0.16 | 0.7246 | -2.2593 | 0.788 | -0.5192 |
| **AdaptB1.0** | 2.97 | 0.15 | 0.7964 | -1.7192 | 0.367 | -2.3896 |
| **AdaptB2.0** | 2.9 | 0.15 | 0.7555 | -2.1492 | 0.482 | -1.8095 |
| **AdaptB3.0** | 2.84 | 0.15 | 0.7042 | -2.7093 | 0.5411 | -1.5595 |
| **AdaptC2.3** | 2.72 | 0.14 | 0.7052 | -2.7893 | 0.4679 | -1.9495 |
| **AdaptA6.0** | 2.58 | 0.14 | 0.7763 | -2.1192 | 0.4636 | -2.0295 |
| **AdaptB2.1** | 2.58 | 0.14 | 0.8518 | -1.3491 | 0.5928 | -1.4194 |
| **AdaptC2.2** | 2.41 | 0.13 | 0.6894 | -3.1793 | 0.4636 | -2.1095 |
| **AdaptB1.1** | 2.33 | 0.13 | 0.7833 | -2.1692 | 0.474 | -2.0995 |
| **AdaptB1.2** | 2.31 | 0.13 | 0.8575 | -1.3691 | 0.497 | -1.9695 |
| **AdaptA6.2** | 2.21 | 0.13 | 0.8747 | -1.2291 | 0.5131 | -1.9395 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **AdaptC1.1** | 1.97 | 0.12 | 0.8112 | -2.0092 | 0.5429 | -1.8995 |
| **AdaptD2.0** | 1.9 | 0.12 | 0.9286 | -0.7191 | 0.7614 | -0.8692 |
| **AdaptC2.4** | 1.89 | 0.12 | 0.6158 | -4.5794 | 0.3969 | -2.8396 |
| **AdaptA6.1** | 1.78 | 0.12 | 1.0294 | 0.341 | 1.2202 | 0.8612 |
| **AdaptC1.0** | 1.78 | 0.12 | 1.0758 | 0.8111 | 0.7884 | -0.7692 |
| **AdaptC1.2** | 1.71 | 0.12 | 1.1005 | 1.0611 | 0.908 | -0.2691 |
| **AdaptC2.5** | 1.63 | 0.12 | 0.8233 | -1.9992 | 0.5253 | -2.1395 |
| **AdaptD3.1** | 1.6 | 0.12 | 1.2355 | 2.4012 | 1.3042 | 1.1613 |
| **AdaptD1.0** | 1.42 | 0.11 | 1.2949 | 3.0313 | 1.6092 | 2.1416 |
| **AdaptB3.1** | 1.21 | 0.11 | 0.6226 | -4.9894 | 0.4601 | -2.7395 |
| **AdaptC2.6** | 1.05 | 0.11 | 0.7703 | -2.8792 | 0.5481 | -2.2495 |
| **AdaptB2.2** | 0.8 | 0.11 | 0.5815 | -5.8794 | 0.4902 | -2.7895 |
| **AdaptB1.3** | 0.75 | 0.11 | 1.1267 | 1.4811 | 0.9415 | -0.1991 |
| **AdaptA5.0** | 0.64 | 0.11 | 1.2692 | 3.0113 | 1.1708 | 0.8312 |
| **AdaptC1.3** | 0.46 | 0.11 | 0.734 | -3.5293 | 0.5719 | -2.3794 |
| **AdaptC1.4** | 0.39 | 0.11 | 0.9478 | -0.6191 | 0.7137 | -1.4793 |
| **AdaptD1.1** | 0.29 | 0.11 | 1.6551 | 6.6717 | 2.0198 | 3.972 |
| **AdaptA4.0** | -0.47 | 0.11 | 1.0499 | 0.611 | 0.8562 | -0.7691 |
| **AdaptA3.0** | -0.79 | 0.11 | 1.0043 | 0.081 | 1.6313 | 2.9916 |
| **AdaptA5.1** | -0.91 | 0.11 | 1.7199 | 6.8417 | 2.1927 | 4.9722 |
| **AdaptC1.5** | -1.39 | 0.12 | 1.0322 | 0.371 | 0.85 | -0.7192 |
| **AdaptA5.2** | -1.4 | 0.12 | 1.6442 | 5.8916 | 1.596 | 2.5916 |
| **AdaptA3.1** | -1.41 | 0.12 | 1.0139 | 0.181 | 1.0941 | 0.5211 |
| **AdaptC1.6** | -2.51 | 0.13 | 1.8273 | 6.1018 | 2.1184 | 2.9721 |
| **AdaptA2.1** | -2.52 | 0.13 | 1.292 | 2.4613 | 0.8725 | -0.3491 |
| **AdaptA2.0** | -3.35 | 0.15 | 1.1232 | 0.9611 | 1.2449 | 0.7012 |
| **AdaptA4.2** | -3.58 | 0.16 | 1.433 | 2.8914 | 1.8101 | 1.7218 |
| **AdaptA4.1** | -3.74 | 0.17 | 0.8987 | -0.7091 | 0.4481 | -1.5396 |
| **AdaptA2.2** | -3.82 | 0.17 | 1.2177 | 1.4912 | 1.0781 | 0.3211 |
| **AdaptC1.7** | -4.22 | 0.18 | 2.3467 | 6.4823 | 1.0123 | 0.181 |
| **AdaptA3.2** | -4.91 | 0.21 | 0.9729 | -0.099 | 2.7474 | 2.5127 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **AdaptA2.3** | -5.75 | 0.25 | 0.9219 | -0.2991 | 0.4066 | -1.1796 |
| **AdaptA3.3** | -6.17 | 0.28 | 1.0791 | 0.4011 | 4.0287 | 3.304 |
| **AdaptA1.1** | -6.59 | 0.3 | 1.0647 | 0.3311 | 1.1999 | 0.5212 |
| **AdaptA1.0** | -6.68 | 0.31 | 0.8981 | -0.2991 | 1.1619 | 0.4612 |
| **AdaptA1.2** | -11.88 | 0.57 | 1.1921 | 0.5312 | 9.9 | 9.9099 |

Cognitive Area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **CogE4.0** | 2.76 | 0.24 | 0.7851 | -0.8192 | 0.2057 | -2.4498 |
| **CogD4.0** | 1.83 | 0.16 | 0.6991 | -1.9093 | 0.4004 | -2.1296 |
| **CogE3.2** | 1.77 | 0.16 | 1.0659 | 0.4411 | 0.6692 | -0.9893 |
| **CogE3.0** | 1.57 | 0.14 | 0.7633 | -1.6392 | 0.5412 | -1.6695 |
| **CogE4.1** | 1.47 | 0.14 | 0.9938 | 0.011 | 0.4659 | -2.1395 |
| **CogD4.1** | 1.45 | 0.14 | 0.7558 | -1.7892 | 0.5681 | -1.6194 |
| **CogE4.3** | 1.4 | 0.13 | 0.943 | -0.3591 | 0.4475 | -2.3196 |
| **CogE4.2** | 1.3 | 0.13 | 1.0578 | 0.4611 | 0.7096 | -1.0793 |
| **CogC3.0** | 1.14 | 0.12 | 0.6203 | -3.4494 | 0.3833 | -3.0296 |
| **CogC3.1** | 1.1 | 0.12 | 0.5904 | -3.8594 | 0.3719 | -3.1696 |
| **CogE3.1** | 1.07 | 0.12 | 0.7483 | -2.2193 | 0.7277 | -1.0893 |
| **CogE2.2** | 0.96 | 0.11 | 0.9581 | -0.329 | 0.8691 | -0.4691 |
| **CogE2.3** | 0.95 | 0.11 | 0.6546 | -3.3893 | 0.8119 | -0.7392 |
| **CogE2.0** | 0.92 | 0.11 | 0.942 | -0.4791 | 0.8191 | -0.7192 |
| **CogB3.2** | 0.86 | 0.11 | 0.7081 | -2.9093 | 0.484 | -2.6695 |
| **CogE2.1** | 0.81 | 0.11 | 1.0699 | 0.6711 | 1.0919 | 0.4711 |
| **CogB3.3** | 0.64 | 0.1 | 0.7747 | -2.4292 | 0.509 | -2.7995 |
| **CogE1.1** | 0.64 | 0.1 | 0.8007 | -2.1192 | 0.893 | -0.4491 |
| **CogC4.1** | 0.55 | 0.1 | 0.7545 | -2.7992 | 0.5521 | -2.6094 |
| **CogC4.0** | 0.53 | 0.1 | 0.8842 | -1.2491 | 0.7658 | -1.2092 |
| **CogC3.2** | 0.51 | 0.1 | 0.7146 | -3.3793 | 0.5419 | -2.7395 |
| **CogE1.0** | 0.27 | 0.09 | 1.5387 | 5.5915 | 1.6481 | 3.1516 |
| **CogB3.0** | 0.19 | 0.09 | 1.0641 | 0.8011 | 1.3341 | 1.8413 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **CogC3.3** | 0.17 | 0.09 | 0.7156 | -3.9393 | 0.5762 | -2.9494 |
| **CogD3.0** | 0.16 | 0.09 | 0.9041 | -1.2191 | 0.8228 | -1.0792 |
| **CogC4.2** | 0.12 | 0.09 | 0.8341 | -2.1992 | 0.7137 | -1.8793 |
| **CogD3.1** | -0.25 | 0.09 | 1.7218 | 8.5117 | 2.3474 | 7.1123 |
| **CogC2.0** | -0.37 | 0.08 | 0.8888 | -1.6591 | 0.7924 | -1.6292 |
| **CogC2.1** | -0.37 | 0.08 | 0.8015 | -3.0792 | 0.6833 | -2.6293 |
| **CogD1.0** | -0.4 | 0.08 | 0.8429 | -2.4092 | 0.7359 | -2.1593 |
| **CogC1.0** | -0.42 | 0.08 | 0.8795 | -1.8191 | 0.7657 | -1.8992 |
| **CogB3.1** | -0.43 | 0.08 | 1.8652 | 9.9019 | 3.0002 | 9.903 |
| **CogB2.0** | -0.44 | 0.08 | 0.7415 | -4.1593 | 0.6209 | -3.3294 |
| **CogC1.1** | -0.48 | 0.08 | 0.9898 | -0.129 | 0.9541 | -0.319 |
| **CogB1.1** | -0.57 | 0.08 | 0.9374 | -0.9191 | 0.7882 | -1.7692 |
| **CogD2.0** | -0.8 | 0.08 | 0.8555 | -2.2091 | 0.7705 | -2.0192 |
| **CogB2.1** | -0.85 | 0.08 | 0.773 | -3.5892 | 0.8367 | -1.3992 |
| **CogD1.1** | -0.87 | 0.08 | 0.779 | -3.4792 | 0.7119 | -2.6293 |
| **CogC1.2** | -0.96 | 0.08 | 1.2157 | 2.9312 | 1.2662 | 2.1013 |
| **CogD2.1** | -1.02 | 0.08 | 1.0127 | 0.201 | 1.0154 | 0.171 |
| **CogB1.2** | -1.08 | 0.08 | 1.3755 | 4.8114 | 1.3855 | 2.9014 |
| **CogE1.2** | -1.18 | 0.08 | 1.2562 | 3.3313 | 1.3555 | 2.6714 |
| **CogD1.2** | -1.22 | 0.08 | 1.0138 | 0.211 | 0.9977 | 0.021 |
| **CogD2.2** | -1.36 | 0.09 | 1.2002 | 2.5512 | 1.1305 | 1.0311 |
| **CogB1.0** | -1.5 | 0.09 | 1.2725 | 3.2813 | 1.3502 | 2.4414 |
| **CogA2.0** | -1.56 | 0.09 | 1.1194 | 1.4911 | 1.6753 | 4.2217 |
| **CogB2.2** | -1.58 | 0.09 | 0.8871 | -1.4491 | 0.9681 | -0.189 |
| **CogA1.0** | -1.75 | 0.09 | 1.7219 | 7.2117 | 2.1795 | 6.1822 |
| **CogA1.1** | -2.23 | 0.1 | 1.6609 | 5.7517 | 3.6453 | 9.0436 |
| **CogA2.1** | -3.45 | 0.14 | 0.8049 | -1.3792 | 1.8695 | 2.3519 |

Fine Motor Area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **FMC1.1** | 5.14 | 0.14 | 0.8525 | -1.1991 | 0.7131 | -0.7693 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **FMC1.2** | 4.92 | 0.13 | 0.6687 | -3.1993 | 0.445 | -1.9096 |
| **FMC1.0** | 4.88 | 0.13 | 0.9713 | -0.209 | 0.848 | -0.3392 |
| **FMC1.3** | 4.28 | 0.11 | 1.0062 | 0.101 | 0.6764 | -1.0593 |
| **FMB3.1** | 2.77 | 0.1 | 0.7519 | -3.1192 | 0.7308 | -1.2093 |
| **FMD1.0** | 2.46 | 0.1 | 1.0271 | 0.341 | 1.0085 | 0.111 |
| **FMB3.2** | 2.11 | 0.1 | 0.9232 | -0.8591 | 0.8434 | -0.6792 |
| **FMB3.0** | 2.1 | 0.1 | 1.5582 | 5.4416 | 1.5649 | 2.2916 |
| **FMD1.1** | 2.05 | 0.1 | 1.1109 | 1.2411 | 1.0044 | 0.091 |
| **FMB3.3** | 1.86 | 0.1 | 0.7021 | -3.6893 | 0.7307 | -1.3093 |
| **FMD1.2** | 1.31 | 0.11 | 1.3903 | 3.6614 | 1.0831 | 0.4311 |
| **FMA3.0** | 1.09 | 0.11 | 0.6203 | -4.3794 | 0.4777 | -2.6995 |
| **FMB3.4** | 1.05 | 0.11 | 0.8545 | -1.4891 | 1.0497 | 0.281 |
| **FMB2.0** | 0.74 | 0.12 | 1.3224 | 2.7613 | 2.033 | 3.252 |
| **FMC1.4** | 0.54 | 0.12 | 1.0597 | 0.5611 | 1.1931 | 0.7712 |
| **FMA3.1** | 0.32 | 0.13 | 0.8843 | -0.9891 | 0.627 | -1.4394 |
| **FMB1.0** | 0.03 | 0.14 | 0.7849 | -1.8292 | 0.6253 | -1.2894 |
| **FMB2.1** | -0.02 | 0.14 | 1.2588 | 1.9613 | 1.1075 | 0.4311 |
| **FMB1.1** | -0.38 | 0.15 | 0.8493 | -1.1492 | 0.4883 | -1.6695 |
| **FMA2.0** | -0.56 | 0.15 | 1.1161 | 0.8611 | 1.3247 | 0.9013 |
| **FMB3.5** | -1.86 | 0.19 | 1.254 | 1.4513 | 0.6127 | -0.6394 |
| **FMA2.1** | -2.05 | 0.2 | 1.6038 | 2.9716 | 1.5339 | 1.0015 |
| **FMA3.2** | -2.16 | 0.2 | 1.0166 | 0.151 | 0.8084 | -0.1692 |
| **FMA2.3** | -2.24 | 0.2 | 0.853 | -0.7791 | 0.759 | -0.2692 |
| **FMB1.2** | -2.51 | 0.22 | 0.5776 | -2.4994 | 0.3098 | -1.4697 |
| **FMB1.3** | -2.51 | 0.22 | 1.0314 | 0.221 | 1.2331 | 0.5712 |
| **FMA2.2** | -3.15 | 0.25 | 1.003 | 0.091 | 0.2026 | -1.9198 |
| **FMA2.4** | -4.36 | 0.33 | 0.9759 | 0.021 | 0.6141 | -0.6294 |
| **FMA1.1** | -4.83 | 0.37 | 1.5201 | 1.4515 | 9.4782 | 6.6595 |
| **FMA1.2** | -5.3 | 0.42 | 0.9919 | 0.111 | 4.2287 | 4.0842 |
| **FMA1.0** | -5.72 | 0.48 | 1.1293 | 0.4211 | 0.2906 | -2.1697 |

Gross Motor Area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **GMB7.0** | 6.68 | 0.23 | 0.9617 | -0.069 | 0.4915 | -1.4695 |
| **GMC3.1** | 6.64 | 0.22 | 1.093 | 0.4511 | 0.8822 | -0.1791 |
| **GMB7.1** | 5.72 | 0.15 | 0.9159 | -0.5091 | 0.4297 | -2.0196 |
| **GMB7.2** | 5.55 | 0.14 | 0.7016 | -2.2793 | 0.3405 | -2.5397 |
| **GMC2.1** | 5.26 | 0.13 | 1.0026 | 0.061 | 1.5363 | 1.4915 |
| **GMC3.2** | 5.19 | 0.12 | 0.8843 | -0.9591 | 3.4764 | 4.7635 |
| **GMC1.1** | 4.97 | 0.12 | 0.8286 | -1.6392 | 0.6759 | -1.0093 |
| **GMC2.0** | 4.49 | 0.1 | 1.0462 | 0.531 | 0.8633 | -0.3391 |
| **GMC1.2** | 4.43 | 0.1 | 0.8747 | -1.4391 | 1.2491 | 0.8312 |
| **GMC1.0** | 4.22 | 0.1 | 1.018 | 0.241 | 3.4712 | 4.9935 |
| **GMC1.3** | 3.96 | 0.1 | 0.7305 | -3.6793 | 6.9901 | 8.747 |
| **GMB6.0** | 3.73 | 0.09 | 0.6508 | -5.0993 | 0.4265 | -2.2996 |
| **GMC3.0** | 3.63 | 0.09 | 2.1732 | 9.9022 | 6.9371 | 9.0569 |
| **GMB4.0** | 3.55 | 0.09 | 0.9937 | -0.049 | 0.7595 | -0.7692 |
| **GMB6.1** | 3.31 | 0.09 | 0.6925 | -4.3193 | 0.4446 | -2.3196 |
| **GMC3.3** | 3.25 | 0.09 | 1.6068 | 6.4716 | 2.0187 | 2.812 |
| **GMC2.2** | 3.12 | 0.1 | 0.9946 | -0.039 | 0.9582 | -0.049 |
| **GMC1.5** | 3.09 | 0.1 | 1.2141 | 2.4812 | 9.9 | 9.9099 |
| **GMB6.2** | 3.03 | 0.1 | 0.6535 | -4.7593 | 0.4336 | -2.4096 |
| **GMC1.4** | 2.89 | 0.1 | 0.7942 | -2.5692 | 2.1403 | 3.1021 |
| **GMB5.0** | 2.84 | 0.1 | 0.682 | -4.1493 | 0.4727 | -2.2195 |
| **GMB6.3** | 2.38 | 0.11 | 0.7003 | -3.3893 | 0.4548 | -2.3395 |
| **GMB5.1** | 2.31 | 0.11 | 0.7417 | -2.7993 | 0.4525 | -2.3495 |
| **GMC1.6** | 2.01 | 0.11 | 1.5441 | 4.2915 | 4.908 | 7.3549 |
| **GMB4.1** | 1.92 | 0.12 | 0.7423 | -2.4493 | 0.6673 | -1.1893 |
| **GMB5.2** | 1.17 | 0.14 | 0.6791 | -2.5493 | 0.2648 | -3.2797 |
| **GMB3.0** | 0.99 | 0.14 | 0.7551 | -1.7792 | 1.0103 | 0.141 |
| **GMC2.3** | 0.78 | 0.15 | 1.2979 | 1.8213 | 1.0849 | 0.3611 |
| **GMB4.2** | 0.47 | 0.16 | 1.0898 | 0.5911 | 1.1744 | 0.5712 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **GMC3.4** | 0.4 | 0.16 | 2.5793 | 6.9226 | 3.0603 | 3.7631 |
| **GMA5.0** | 0.38 | 0.16 | 0.5961 | -2.8594 | 0.6145 | -1.0694 |
| **GMA5.1** | 0.35 | 0.16 | 0.5587 | -3.1894 | 0.2979 | -2.5097 |
| **GMB3.1** | 0.24 | 0.16 | 0.6548 | -2.3493 | 0.197 | -3.0698 |
| **GMB2.1** | 0 | 0.17 | 0.533 | -3.3795 | 0.2834 | -2.3297 |
| **GMB2.0** | -0.03 | 0.17 | 0.5891 | -2.8794 | 0.2629 | -2.4197 |
| **GMB3.2** | -0.11 | 0.17 | 0.7033 | -1.9593 | 0.164 | -2.9898 |
| **GMB2.2** | -0.32 | 0.17 | 0.5463 | -3.2595 | 0.2247 | -2.4198 |
| **GMB4.3** | -0.4 | 0.17 | 1.182 | 1.1012 | 0.9288 | 0.0009 |
| **GMB3.3** | -0.93 | 0.18 | 0.6658 | -2.2393 | 0.1781 | -2.2398 |
| **GMB3.4** | -1.18 | 0.18 | 0.547 | -3.2295 | 1.0329 | 0.251 |
| **GMB2.3** | -1.72 | 0.19 | 0.6517 | -2.2993 | 0.1338 | -2.1499 |
| **GMB2.4** | -1.97 | 0.19 | 0.8336 | -0.9592 | 0.1999 | -1.7698 |
| **GMB1.0** | -1.97 | 0.19 | 1.0938 | 0.5811 | 0.2373 | -1.6098 |
| **GMA5.2** | -2.04 | 0.19 | 1.4848 | 2.4715 | 0.5852 | -0.5694 |
| **GMB1.1** | -2.51 | 0.2 | 1.1279 | 0.7111 | 0.3145 | -1.3497 |
| **GMB1.2** | -2.72 | 0.21 | 1.3255 | 1.5713 | 0.44 | -0.9896 |
| **GMA4.1** | -2.72 | 0.21 | 0.7948 | -1.0592 | 0.187 | -1.9198 |
| **GMB1.3** | -2.77 | 0.21 | 1.3559 | 1.6814 | 0.3998 | -1.1196 |
| **GMA4.2** | -2.9 | 0.21 | 0.7321 | -1.3993 | 0.1257 | -2.3499 |
| **GMA4.0** | -2.99 | 0.22 | 0.578 | -2.3794 | 0.1017 | -2.5799 |
| **GMA4.3** | -3.18 | 0.22 | 0.748 | -1.2493 | 0.1099 | -2.6099 |
| **GMB1.4** | -3.59 | 0.23 | 1.7969 | 3.0418 | 7.54 | 5.8775 |
| **GMA4.4** | -3.76 | 0.24 | 0.6621 | -1.6593 | 3.8049 | 3.6238 |
| **GMA4.5** | -4 | 0.25 | 0.9231 | -0.2691 | 0.992 | 0.141 |
| **GMA2.0** | -4 | 0.25 | 0.8391 | -0.6592 | 0.1725 | -2.7498 |
| **GMA3.0** | -4.02 | 0.25 | 1.4231 | 1.6914 | 2.8513 | 2.9029 |
| **GMA3.1** | -4.53 | 0.27 | 0.9276 | -0.1991 | 0.2025 | -2.9098 |
| **GMA2.1** | -4.54 | 0.27 | 0.8724 | -0.4291 | 0.2533 | -2.5797 |
| **GMA3.2** | -4.77 | 0.29 | 0.9146 | -0.2291 | 0.2292 | -2.8998 |
| **GMA4.6** | -5.12 | 0.31 | 0.5131 | -1.8895 | 0.1099 | -4.2499 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **GMA2.2** | -5.44 | 0.33 | 1.1441 | 0.5311 | 3.6278 | 4.7736 |
| **GMA1.0** | -6.34 | 0.42 | 1.6908 | 1.5517 | 0.8454 | -0.3992 |
| **GMA1.1** | -6.58 | 0.43 | 1.7091 | 1.5617 | 1.9464 | 2.4419 |
| **GMA1.2** | -7.51 | 0.55 | 0.439 | -1.1396 | 0.0381 | -6.37 |
| **GMA1.3** | -8.28 | 0.69 | 0.7453 | -0.1393 | 0.0841 | -6.0599 |

Social Communication Area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **soc_C2.0** | 3.26 | 0.02 | 0.8493 | -9.8992 | 0.7406 | -7.1393 |
| **soc_C2.1** | 2.82 | 0.02 | 0.8023 | -9.8992 | 0.6715 | -9.8993 |
| **soc_C1.0** | 2.42 | 0.02 | 0.8085 | -9.8992 | 0.7941 | -6.6392 |
| **soc_C1.1** | 2.01 | 0.01 | 0.7887 | -9.8992 | 0.682 | -9.8993 |
| **soc_B1.0** | 1.9 | 0.01 | 1.1426 | 9.9011 | 1.2612 | 8.1813 |
| **soc_C2.2** | 1.72 | 0.01 | 0.9249 | -7.0691 | 0.8931 | -3.8591 |
| **soc_C1.2** | 1.53 | 0.01 | 0.7705 | -9.8992 | 0.6748 | -9.8993 |
| **soc_A3.0** | 1.4 | 0.01 | 0.9287 | -6.8191 | 0.8366 | -6.6792 |
| **soc_B2.0** | 1.33 | 0.01 | 1.0837 | 7.6311 | 0.9994 | -0.019 |
| **soc_B1.1** | 0.73 | 0.01 | 1.0999 | 9.1211 | 1.1967 | 8.1112 |
| **soc_A3.1** | 0.71 | 0.01 | 0.8382 | -9.8992 | 0.7488 | -9.8993 |
| **soc_B2.1** | 0.29 | 0.01 | 0.946 | -5.1191 | 0.8577 | -7.1091 |
| **soc_A2.1** | 0.11 | 0.01 | 0.9509 | -4.599 | 0.8491 | -7.6292 |
| **soc_A2.0** | 0 | 0.01 | 0.9848 | -1.399 | 0.8797 | -5.9891 |
| **soc_B1.2** | -0.41 | 0.01 | 1.5232 | 9.9015 | 2.3528 | 9.9024 |
| **soc_C1.3** | -0.49 | 0.01 | 1.0581 | 5.0311 | 0.9152 | -3.8391 |
| **soc_A3.2** | -0.53 | 0.02 | 0.9823 | -1.559 | 0.9386 | -2.7191 |
| **soc_A2.2** | -1.15 | 0.02 | 0.8427 | -9.8992 | 0.7097 | -9.8993 |
| **soc_C1.4** | -1.17 | 0.02 | 1.1413 | 9.9011 | 0.9208 | -2.7691 |
| **soc_C1.5** | -1.32 | 0.02 | 1.1861 | 9.9012 | 1.1519 | 4.7112 |
| **soc_A1.0** | -2.11 | 0.02 | 1.1015 | 6.9611 | 1.0811 | 2.0211 |
| **soc_A1.1** | -2.6 | 0.02 | 0.9241 | -5.0391 | 0.6862 | -7.7793 |
| **soc_A1.2** | -2.77 | 0.02 | 1.1293 | 7.8011 | 1.2766 | 5.3913 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **soc_A2.3** | -3.1 | 0.02 | 1.2213 | 9.9012 | 1.0116 | 0.241 |
| **soc_A1.3** | -4.58 | 0.03 | 1.3948 | 9.9014 | 1.4151 | 5.3414 |

Social- Emotional Area

| NAME | MEASURE | MODLSE | INFIT.MSQ | INFIT.ZSTD | OUTFIT.MSQ | OUTFIT.ZSTD |
|---|---|---|---|---|---|---|
| **SEC4.0** | 2.13 | 0.13 | 0.5655 | -3.6294 | 0.3721 | -2.5396 |
| **SED4.0** | 1.79 | 0.12 | 0.6735 | -3.0393 | 0.4283 | -2.5796 |
| **SED4.1** | 1.79 | 0.12 | 0.5529 | -4.4294 | 0.4631 | -2.3595 |
| **SEC4.1** | 1.65 | 0.11 | 0.66 | -3.3993 | 0.4497 | -2.5896 |
| **SEC4.2** | 1.64 | 0.11 | 0.6829 | -3.1493 | 0.5661 | -1.8994 |
| **SEC2.0** | 1.61 | 0.11 | 0.6655 | -3.3993 | 0.4582 | -2.5695 |
| **SEE4.4** | 1.5 | 0.11 | 0.708 | -3.0493 | 0.4307 | -2.8696 |
| **SEC2.1** | 1.44 | 0.11 | 0.6341 | -4.0894 | 0.3847 | -3.2896 |
| **SED1.0** | 1.13 | 0.1 | 1.1215 | 1.3311 | 0.9742 | -0.049 |
| **SED1.2** | 1.07 | 0.1 | 0.9675 | -0.339 | 0.8433 | -0.7092 |
| **SEE4.1** | 1.05 | 0.1 | 1.0273 | 0.341 | 0.6964 | -1.5293 |
| **SEE1.0** | 1.04 | 0.1 | 0.6987 | -3.8393 | 0.5912 | -2.2294 |
| **SED1.1** | 1.03 | 0.1 | 1.122 | 1.3711 | 1.0911 | 0.4911 |
| **SEB3.2** | 0.99 | 0.1 | 0.7846 | -2.6792 | 0.7944 | -0.9992 |
| **SED3.1** | 0.98 | 0.09 | 0.9783 | -0.229 | 0.8455 | -0.7292 |
| **SED2.0** | 0.91 | 0.09 | 0.9217 | -0.9291 | 0.8581 | -0.6691 |
| **SEE4.3** | 0.88 | 0.09 | 0.7947 | -2.6492 | 0.6063 | -2.2594 |
| **SED3.0** | 0.84 | 0.09 | 0.8449 | -1.9792 | 0.7423 | -1.3893 |
| **SEC3.3** | 0.75 | 0.09 | 1.0513 | 0.6611 | 1.0197 | 0.171 |
| **SEC3.0** | 0.7 | 0.09 | 0.7105 | -4.0893 | 0.5654 | -2.7894 |
| **SED2.1** | 0.69 | 0.09 | 0.6875 | -4.4793 | 0.6247 | -2.3494 |
| **SED2.2** | 0.68 | 0.09 | 0.5747 | -6.4394 | 0.5208 | -3.1795 |
| **SED1.3** | 0.66 | 0.09 | 1.4266 | 4.8914 | 1.4118 | 2.1114 |
| **SEE4.0** | 0.63 | 0.09 | 0.9386 | -0.7891 | 0.8427 | -0.8792 |
| **SED2.3** | 0.62 | 0.09 | 1.32 | 3.8113 | 1.2465 | 1.3612 |
| **SED3.2** | 0.61 | 0.09 | 1.0054 | 0.101 | 0.9147 | -0.4391 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **SEB3.1** | 0.51 | 0.09 | 1.0383 | 0.521 | 0.8493 | -0.8792 |
| **SEC3.1** | 0.44 | 0.09 | 0.8763 | -1.7191 | 0.6516 | -2.4093 |
| **SEC2.2** | 0.4 | 0.09 | 0.9165 | -1.1391 | 0.8677 | -0.8191 |
| **SEE4.2** | 0.35 | 0.09 | 0.8556 | -2.0391 | 0.6682 | -2.3493 |
| **SEE3.1** | 0.33 | 0.09 | 1.0348 | 0.491 | 0.9109 | -0.5391 |
| **SED4.2** | 0.31 | 0.09 | 1.0767 | 1.0511 | 0.9939 | 0.011 |
| **SEB3.0** | 0.2 | 0.09 | 1.0329 | 0.471 | 0.9332 | -0.4091 |
| **SEC3.2** | 0.19 | 0.09 | 0.7843 | -3.2192 | 0.6228 | -2.9294 |
| **SEC1.0** | 0.1 | 0.08 | 0.9445 | -0.7691 | 0.8882 | -0.7691 |
| **SEB1.1** | 0.06 | 0.09 | 0.9536 | -0.629 | 0.9728 | -0.139 |
| **SEE3.0** | -0.04 | 0.08 | 0.8466 | -2.2292 | 0.777 | -1.7092 |
| **SEB1.2** | -0.05 | 0.08 | 1.1489 | 2.0011 | 1.2638 | 1.8113 |
| **SEE3.2** | -0.15 | 0.08 | 0.8753 | -1.7791 | 0.8812 | -0.8691 |
| **SEE2.2** | -0.19 | 0.08 | 0.8063 | -2.8692 | 0.745 | -2.0593 |
| **SEB1.0** | -0.22 | 0.09 | 0.7679 | -3.4592 | 0.6899 | -2.5593 |
| **SEC1.1** | -0.25 | 0.08 | 0.7551 | -3.6892 | 0.6512 | -2.9793 |
| **SEB2.1** | -0.36 | 0.09 | 1.0158 | 0.241 | 1.0702 | 0.5611 |
| **SEE2.0** | -0.39 | 0.09 | 0.9383 | -0.8291 | 0.9682 | -0.199 |
| **SEC1.3** | -0.45 | 0.09 | 2.5466 | 9.9025 | 3.0838 | 9.9031 |
| **SEE2.1** | -0.45 | 0.09 | 1.006 | 0.111 | 1.1036 | 0.8011 |
| **SEB2.0** | -0.46 | 0.09 | 0.8798 | -1.6791 | 1.05 | 0.4211 |
| **SEA1.0** | -0.59 | 0.09 | 3.6718 | 9.9037 | 7.1819 | 9.9072 |
| **SEC1.2** | -0.68 | 0.09 | 0.6972 | -4.4293 | 0.5663 | -3.8294 |
| **SEE1.1** | -0.74 | 0.09 | 0.9716 | -0.349 | 1.1629 | 1.1912 |
| **SED3.3** | -0.84 | 0.09 | 1.5729 | 6.2316 | 1.9124 | 5.2819 |
| **SEA3.1** | -1.08 | 0.09 | 0.7775 | -2.8692 | 0.6783 | -2.3893 |
| **SEA3.2** | -1.2 | 0.09 | 0.9492 | -0.5791 | 0.7833 | -1.4492 |
| **SEA3.0** | -1.42 | 0.1 | 0.9854 | -0.129 | 1.8479 | 4.0318 |
| **SEA2.2** | -1.89 | 0.11 | 0.7326 | -2.7893 | 0.8251 | -0.8092 |
| **SEB2.3** | -1.9 | 0.11 | 1.6823 | 5.4517 | 1.6962 | 2.7817 |
| **SEA2.0** | -2.32 | 0.12 | 0.7765 | -1.9492 | 0.8987 | -0.3191 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **SEA2.1** | -2.58 | 0.13 | 0.6714 | -2.7393 | 0.5436 | -1.7995 |
| **SEB2.2** | -2.71 | 0.14 | 1.1274 | 0.9111 | 1.4667 | 1.4315 |
| **SEA2.3** | -2.93 | 0.15 | 0.8457 | -0.9992 | 0.7619 | -0.6492 |
| **SEA1.2** | -3.86 | 0.21 | 1.0952 | 0.4811 | 2.084 | 1.8621 |
| **SEA1.1** | -3.95 | 0.22 | 0.8582 | -0.5491 | 1.1402 | 0.4311 |

## APPENDIX 7. PERSON ABILITY MEASURE CONVERSION TABLE IN EACH AREA

Convert Table in the Adaptive Area

| V2_score | V2_measure | V3_measure | V3_F_measure |
|---|---|---|---|
| **1** | -4.88 | -6.65 | -7.65 |
| **2** | -7.64 | -10.69 | -6.6 |
| **3** | -4.44 | -6.09 | -5.93 |
| **4** | -5.45 | -8.05 | -5.48 |
| **5** | -2.92 | -4.74 | -5.15 |
| **6** | -4.09 | -5.74 | -4.89 |
| **7** | -2.4 | -4.29 | -4.67 |
| **8** | -6.33 | -9.55 | -4.48 |
| **9** | -3.8 | -5.48 | -4.31 |
| **10** | -3.55 | -5.27 | -4.16 |
| **11** | -2.57 | -4.44 | -4.01 |
| **12** | -1.5 | -3.43 | -3.88 |
| **13** | -3.32 | -5.08 | -3.75 |
| **14** | -3.12 | -4.91 | -3.63 |
| **15** | -2.74 | -4.59 | -3.51 |
| **16** | -1.94 | -3.85 | -3.4 |
| **17** | -2.24 | -4.14 | -3.29 |
| **18** | -1.64 | -3.57 | -3.18 |
| **19** | -0.69 | -2.64 | -3.07 |
| **20** | -1.36 | -3.29 | -2.96 |
| **21** | -1.22 | -3.16 | -2.86 |
| **22** | 0.42 | -1.66 | -2.75 |
| **23** | -0.56 | -2.52 | -2.65 |
| **24** | -0.18 | -2.18 | -2.54 |
| **25** | 0.18 | -1.86 | -2.44 |
| **26** | 0.87 | -1.29 | -2.33 |
| **27** | 0.99 | -1.2 | -2.23 |
| **28** | 0.3 | -1.76 | -2.13 |
| **29** | 1.65 | -0.68 | -2.03 |
| **30** | 2.24 | -0.23 | -1.93 |

| | | | |
|---|---|---|---|
| **31** | 1.1 | -1.11 | -1.83 |
| **32** | -1.79 | -3.71 | -1.73 |
| **33** | 1.32 | -0.94 | -1.63 |
| **34** | 0.06 | -1.96 | -1.53 |
| **35** | -2.09 | -4 | -1.44 |
| **36** | -0.95 | -2.89 | -1.35 |
| **37** | -1.09 | -3.03 | -1.25 |
| **38** | -0.82 | -2.77 | -1.16 |
| **39** | -0.06 | -2.07 | -1.08 |
| **40** | -0.31 | -2.29 | -0.99 |
| **41** | 0.65 | -1.47 | -0.9 |
| **42** | 1.21 | -1.03 | -0.81 |
| **43** | -0.43 | -2.4 | -0.73 |
| **44** | 0.76 | -1.38 | -0.64 |
| **45** | 1.54 | -0.77 | -0.56 |
| **46** | 0.53 | -1.57 | -0.47 |
| **47** | 2.51 | -0.04 | -0.38 |
| **48** | 1.43 | -0.86 | -0.29 |
| **49** | 4.38 | 1.37 | -0.2 |
| **50** | 1.88 | -0.51 | -0.11 |
| **51** | 2 | -0.42 | -0.02 |
| **52** | 1.77 | -0.6 | 0.08 |
| **53** | 2.37 | -0.14 | 0.18 |
| **54** | 2.8 | 0.18 | 0.29 |
| **55** | 2.12 | -0.33 | 0.4 |
| **56** | 3.32 | 0.57 | 0.52 |
| **57** | 3.13 | 0.43 | 0.65 |
| **58** | 3.53 | 0.72 | 0.79 |
| **59** | 3.76 | 0.9 | 0.94 |
| **60** | 2.65 | 0.07 | 1.12 |
| **61** | 2.96 | 0.3 | 1.33 |
| **62** | 6.82 | 3.39 | 1.59 |
| **63** | 4.04 | 1.11 | 1.95 |
| **64** | 4.84 | 1.72 | 2.55 |
| **65** | 5.58 | 2.32 | 3.62 |

Convert Table in the Cognitive Area

| SCORE | V2_MEASURE | V3_MEASURE | V3_F_MEASURE |
|---|---|---|---|
| **0** | -8.29 | -7.05 | -8.33 |
| **1** | -7.07 | -5.92 | -7.2 |
| **2** | -6.35 | -5.29 | -6.57 |
| **3** | -5.91 | -4.92 | -6.2 |
| **4** | -5.59 | -4.66 | -5.93 |
| **5** | -5.33 | -4.44 | -5.72 |
| **6** | -5.1 | -4.26 | -5.54 |
| **7** | -4.9 | -4.09 | -5.37 |
| **8** | -4.72 | -3.95 | -5.23 |

| | | | |
|---|---|---|---|
| **9** | -4.56 | -3.81 | -5.09 |
| **10** | -4.4 | -3.68 | -4.96 |
| **11** | -4.25 | -3.56 | -4.84 |
| **12** | -4.1 | -3.44 | -4.72 |
| **13** | -3.97 | -3.33 | -4.61 |
| **14** | -3.83 | -3.22 | -4.5 |
| **15** | -3.7 | -3.11 | -4.39 |
| **16** | -3.58 | -3.01 | -4.29 |
| **17** | -3.46 | -2.91 | -4.19 |
| **18** | -3.34 | -2.81 | -4.09 |
| **19** | -3.22 | -2.71 | -3.99 |
| **20** | -3.11 | -2.62 | -3.9 |
| **21** | -3 | -2.53 | -3.81 |
| **22** | -2.89 | -2.44 | -3.72 |
| **23** | -2.79 | -2.35 | -3.63 |
| **24** | -2.68 | -2.26 | -3.54 |
| **25** | -2.58 | -2.18 | -3.46 |
| **26** | -2.48 | -2.09 | -3.38 |
| **27** | -2.38 | -2.01 | -3.3 |
| **28** | -2.29 | -1.93 | -3.22 |
| **29** | -2.19 | -1.85 | -3.14 |
| **30** | -2.1 | -1.78 | -3.06 |
| **31** | -2.01 | -1.7 | -2.99 |
| **32** | -1.92 | -1.62 | -2.91 |
| **33** | -1.83 | -1.55 | -2.84 |
| **34** | -1.75 | -1.48 | -2.77 |
| **35** | -1.66 | -1.41 | -2.7 |
| **36** | -1.57 | -1.33 | -2.63 |
| **37** | -1.49 | -1.26 | -2.56 |
| **38** | -1.41 | -1.19 | -2.49 |
| **39** | -1.33 | -1.13 | -2.42 |
| **40** | -1.24 | -1.06 | -2.36 |
| **41** | -1.16 | -0.99 | -2.29 |
| **42** | -1.08 | -0.92 | -2.22 |
| **43** | -1.01 | -0.86 | -2.16 |
| **44** | -0.93 | -0.79 | -2.1 |
| **45** | -0.85 | -0.73 | -2.03 |
| **46** | -0.77 | -0.66 | -1.97 |
| **47** | -0.7 | -0.6 | -1.91 |
| **48** | -0.62 | -0.54 | -1.84 |
| **49** | -0.54 | -0.47 | -1.78 |
| **50** | -0.47 | -0.41 | -1.72 |
| **51** | -0.39 | -0.35 | -1.66 |

| | | | |
|---|---|---|---|
| **52** | -0.32 | -0.29 | -1.6 |
| **53** | -0.25 | -0.22 | -1.54 |
| **54** | -0.17 | -0.16 | -1.48 |
| **55** | -0.1 | -0.1 | -1.42 |
| **56** | -0.02 | -0.04 | -1.36 |
| **57** | 0.05 | 0.02 | -1.3 |
| **58** | 0.12 | 0.08 | -1.24 |
| **59** | 0.2 | 0.14 | -1.18 |
| **60** | 0.27 | 0.2 | -1.12 |
| **61** | 0.34 | 0.27 | -1.06 |
| **62** | 0.42 | 0.33 | -1 |
| **63** | 0.49 | 0.39 | -0.94 |
| **64** | 0.56 | 0.45 | -0.88 |
| **65** | 0.64 | 0.51 | -0.82 |
| **66** | 0.71 | 0.57 | -0.76 |
| **67** | 0.78 | 0.63 | -0.7 |
| **68** | 0.86 | 0.69 | -0.64 |
| **69** | 0.93 | 0.75 | -0.58 |
| **70** | 1.01 | 0.81 | -0.52 |
| **71** | 1.08 | 0.88 | -0.45 |
| **72** | 1.15 | 0.94 | -0.39 |
| **73** | 1.23 | 1 | -0.33 |
| **74** | 1.3 | 1.06 | -0.27 |
| **75** | 1.38 | 1.12 | -0.21 |
| **76** | 1.45 | 1.19 | -0.15 |
| **77** | 1.53 | 1.25 | -0.08 |
| **78** | 1.61 | 1.31 | -0.02 |
| **79** | 1.68 | 1.37 | 0.04 |
| **80** | 1.76 | 1.44 | 0.11 |
| **81** | 1.84 | 1.5 | 0.17 |
| **82** | 1.91 | 1.57 | 0.23 |
| **83** | 1.99 | 1.63 | 0.3 |
| **84** | 2.07 | 1.69 | 0.36 |
| **85** | 2.15 | 1.76 | 0.43 |
| **86** | 2.23 | 1.82 | 0.49 |
| **87** | 2.31 | 1.89 | 0.56 |
| **88** | 2.38 | 1.95 | 0.62 |
| **89** | 2.46 | 2.02 | 0.69 |
| **90** | 2.55 | 2.09 | 0.75 |
| **91** | 2.63 | 2.15 | 0.82 |
| **92** | 2.71 | 2.22 | 0.89 |
| **93** | 2.79 | 2.29 | 0.96 |
| **94** | 2.88 | 2.35 | 1.02 |

| | | | |
|---|---|---|---|
| **95** | 2.96 | 2.42 | 1.09 |
| **96** | 3.05 | 2.49 | 1.16 |
| **97** | 3.14 | 2.57 | 1.24 |
| **98** | 3.22 | 2.64 | 1.31 |
| **99** | 3.32 | 2.71 | 1.38 |
| **100** | 3.41 | 2.79 | 1.46 |
| **101** | 3.51 | 2.87 | 1.54 |
| **102** | 3.61 | 2.95 | 1.62 |
| **103** | 3.71 | 3.03 | 1.71 |
| **104** | 3.82 | 3.12 | 1.79 |
| **105** | 3.94 | 3.22 | 1.89 |
| **106** | 4.06 | 3.32 | 1.99 |
| **107** | 4.2 | 3.42 | 2.09 |
| **108** | 4.34 | 3.54 | 2.21 |
| **109** | 4.5 | 3.67 | 2.34 |
| **110** | 4.67 | 3.81 | 2.49 |
| **111** | 4.88 | 3.98 | 2.65 |
| **112** | 5.12 | 4.19 | 2.86 |
| **113** | 5.43 | 4.45 | 3.12 |
| **114** | 5.85 | 4.81 | 3.48 |
| **115** | 6.56 | 5.45 | 4.12 |
| **116** | 7.77 | 6.6 | 5.27 |

Convert Table in the Fine Motor Area

| SCORE | V2_MEASURE | V3_MEASURE | V3_F_MEASURE |
|---|---|---|---|
| **0** | -8.68 | -7.63 | -7.61 |
| **1** | -7.37 | -6.54 | -6.52 |
| **2** | -6.52 | -5.86 | -5.86 |
| **3** | -5.96 | -5.42 | -5.42 |
| **4** | -5.52 | -5.08 | -5.06 |
| **5** | -5.14 | -4.78 | -4.76 |
| **6** | -4.82 | -4.53 | -4.49 |
| **7** | -4.52 | -4.3 | -4.24 |
| **8** | -4.24 | -4.08 | -4.01 |
| **9** | -3.99 | -3.88 | -3.79 |
| **10** | -3.74 | -3.69 | -3.59 |
| **11** | -3.51 | -3.51 | -3.39 |
| **12** | -3.29 | -3.33 | -3.21 |
| **13** | -3.07 | -3.16 | -3.03 |
| **14** | -2.87 | -3 | -2.86 |
| **15** | -2.67 | -2.84 | -2.69 |
| **16** | -2.47 | -2.68 | -2.54 |

| | | | |
|---|---|---|---|
| **17** | -2.28 | -2.53 | -2.38 |
| **18** | -2.1 | -2.38 | -2.24 |
| **19** | -1.92 | -2.24 | -2.1 |
| **20** | -1.74 | -2.1 | -1.96 |
| **21** | -1.57 | -1.96 | -1.83 |
| **22** | -1.4 | -1.83 | -1.7 |
| **23** | -1.24 | -1.7 | -1.58 |
| **24** | -1.08 | -1.57 | -1.45 |
| **25** | -0.92 | -1.45 | -1.34 |
| **26** | -0.77 | -1.33 | -1.22 |
| **27** | -0.62 | -1.21 | -1.11 |
| **28** | -0.47 | -1.1 | -0.99 |
| **29** | -0.33 | -0.98 | -0.88 |
| **30** | -0.18 | -0.87 | -0.78 |
| **31** | -0.04 | -0.76 | -0.67 |
| **32** | 0.09 | -0.65 | -0.56 |
| **33** | 0.23 | -0.55 | -0.46 |
| **34** | 0.37 | -0.44 | -0.36 |
| **35** | 0.5 | -0.34 | -0.26 |
| **36** | 0.63 | -0.23 | -0.16 |
| **37** | 0.76 | -0.13 | -0.06 |
| **38** | 0.89 | -0.03 | 0.04 |
| **39** | 1.02 | 0.07 | 0.14 |
| **40** | 1.15 | 0.17 | 0.24 |
| **41** | 1.28 | 0.27 | 0.34 |
| **42** | 1.41 | 0.37 | 0.44 |
| **43** | 1.54 | 0.47 | 0.54 |
| **44** | 1.67 | 0.57 | 0.64 |
| **45** | 1.81 | 0.67 | 0.74 |
| **46** | 1.94 | 0.78 | 0.84 |
| **47** | 2.07 | 0.88 | 0.94 |
| **48** | 2.21 | 0.99 | 1.05 |
| **49** | 2.35 | 1.1 | 1.15 |
| **50** | 2.49 | 1.21 | 1.26 |
| **51** | 2.63 | 1.32 | 1.37 |
| **52** | 2.78 | 1.43 | 1.49 |
| **53** | 2.94 | 1.55 | 1.61 |
| **54** | 3.1 | 1.68 | 1.73 |
| **55** | 3.27 | 1.81 | 1.86 |
| **56** | 3.46 | 1.95 | 2 |
| **57** | 3.65 | 2.1 | 2.16 |
| **58** | 3.86 | 2.27 | 2.32 |
| **59** | 4.1 | 2.45 | 2.5 |

| | | | |
|---|---|---|---|
| **60** | 4.36 | 2.66 | 2.71 |
| **61** | 4.66 | 2.9 | 2.95 |
| **62** | 5.02 | 3.18 | 3.23 |
| **63** | 5.45 | 3.52 | 3.58 |
| **64** | 6.02 | 3.97 | 4.02 |
| **65** | 6.87 | 4.65 | 4.7 |
| **66** | 8.18 | 5.75 | 5.8 |

Convert Table in the Gross Motor Area

| SCORE | V2_MEASURE | V2onV3_MEASURE |
|---|---|---|
| **0** | -10.95 | -10.75 |
| **1** | -9.72 | -9.62 |
| **2** | -8.97 | -8.96 |
| **3** | -8.47 | -8.54 |
| **4** | -8.07 | -8.19 |
| **5** | -7.71 | -7.88 |
| **6** | -7.37 | -7.59 |
| **7** | -7.03 | -7.29 |
| **8** | -6.7 | -7 |
| **9** | -6.38 | -6.71 |
| **10** | -6.08 | -6.45 |
| **11** | -5.81 | -6.2 |
| **12** | -5.55 | -5.98 |
| **13** | -5.32 | -5.78 |
| **14** | -5.1 | -5.59 |
| **15** | -4.9 | -5.41 |
| **16** | -4.7 | -5.24 |
| **17** | -4.52 | -5.07 |
| **18** | -4.34 | -4.92 |
| **19** | -4.17 | -4.77 |
| **20** | -4 | -4.62 |
| **21** | -3.84 | -4.48 |
| **22** | -3.69 | -4.34 |
| **23** | -3.53 | -4.21 |
| **24** | -3.38 | -4.07 |
| **25** | -3.24 | -3.94 |
| **26** | -3.09 | -3.82 |
| **27** | -2.95 | -3.69 |
| **28** | -2.82 | -3.57 |
| **29** | -2.68 | -3.45 |
| **30** | -2.55 | -3.34 |
| **31** | -2.42 | -3.22 |

| | | |
|---|---|---|
| **32** | -2.29 | -3.11 |
| **33** | -2.17 | -3 |
| **34** | -2.04 | -2.89 |
| **35** | -1.92 | -2.78 |
| **36** | -1.8 | -2.68 |
| **37** | -1.67 | -2.57 |
| **38** | -1.55 | -2.47 |
| **39** | -1.43 | -2.36 |
| **40** | -1.31 | -2.26 |
| **41** | -1.19 | -2.15 |
| **42** | -1.07 | -2.05 |
| **43** | -0.95 | -1.95 |
| **44** | -0.83 | -1.84 |
| **45** | -0.71 | -1.74 |
| **46** | -0.59 | -1.64 |
| **47** | -0.47 | -1.54 |
| **48** | -0.35 | -1.44 |
| **49** | -0.23 | -1.33 |
| **50** | -0.11 | -1.23 |
| **51** | 0 | -1.14 |
| **52** | 0.12 | -1.04 |
| **53** | 0.23 | -0.94 |
| **54** | 0.35 | -0.85 |
| **55** | 0.46 | -0.75 |
| **56** | 0.57 | -0.66 |
| **57** | 0.67 | -0.57 |
| **58** | 0.78 | -0.48 |
| **59** | 0.88 | -0.4 |
| **60** | 0.98 | -0.31 |
| **61** | 1.08 | -0.23 |
| **62** | 1.18 | -0.14 |
| **63** | 1.28 | -0.06 |
| **64** | 1.38 | 0.02 |
| **65** | 1.47 | 0.1 |
| **66** | 1.57 | 0.18 |
| **67** | 1.66 | 0.27 |
| **68** | 1.75 | 0.35 |
| **69** | 1.85 | 0.43 |
| **70** | 1.94 | 0.51 |
| **71** | 2.03 | 0.59 |
| **72** | 2.13 | 0.67 |
| **73** | 2.22 | 0.75 |
| **74** | 2.31 | 0.83 |

| | | |
|---|---|---|
| **75** | 2.41 | 0.91 |
| **76** | 2.5 | 0.99 |
| **77** | 2.6 | 1.07 |
| **78** | 2.69 | 1.16 |
| **79** | 2.79 | 1.24 |
| **80** | 2.88 | 1.32 |
| **81** | 2.98 | 1.41 |
| **82** | 3.08 | 1.49 |
| **83** | 3.18 | 1.58 |
| **84** | 3.28 | 1.67 |
| **85** | 3.38 | 1.76 |
| **86** | 3.48 | 1.85 |
| **87** | 3.58 | 1.94 |
| **88** | 3.69 | 2.03 |
| **89** | 3.8 | 2.13 |
| **90** | 3.9 | 2.22 |
| **91** | 4.01 | 2.32 |
| **92** | 4.13 | 2.42 |
| **93** | 4.24 | 2.52 |
| **94** | 4.36 | 2.63 |
| **95** | 4.49 | 2.74 |
| **96** | 4.61 | 2.85 |
| **97** | 4.75 | 2.97 |
| **98** | 4.89 | 3.09 |
| **99** | 5.04 | 3.22 |
| **100** | 5.19 | 3.36 |
| **101** | 5.36 | 3.51 |
| **102** | 5.55 | 3.67 |
| **103** | 5.75 | 3.84 |
| **104** | 5.97 | 4.04 |
| **105** | 6.23 | 4.26 |
| **106** | 6.52 | 4.53 |
| **107** | 6.89 | 4.84 |
| **108** | 7.36 | 5.27 |
| **109** | 8.11 | 5.94 |
| **110** | 9.35 | 7.08 |

Convert Table in the Social Communication Area

| SCORE | V2_MEASURE | V3_MEASURE | V3_F_MEASURE |
|---|---|---|---|
| **46** | -8.52 | -8.05 | -8.35 |
| **47** | -7.28 | -6.8 | -7.11 |
| **48** | -6.53 | -6.05 | -6.36 |

| | | | |
|---|---|---|---|
| **49** | -6.07 | -5.58 | -5.89 |
| **50** | -5.72 | -5.23 | -5.54 |
| **51** | -5.43 | -4.94 | -5.25 |
| **52** | -5.19 | -4.69 | -5 |
| **53** | -4.97 | -4.47 | -4.78 |
| **54** | -4.77 | -4.27 | -4.58 |
| **55** | -4.58 | -4.09 | -4.39 |
| **56** | -4.4 | -3.91 | -4.21 |
| **57** | -4.23 | -3.75 | -4.04 |
| **58** | -4.07 | -3.59 | -3.87 |
| **59** | -3.91 | -3.44 | -3.71 |
| **60** | -3.76 | -3.3 | -3.56 |
| **61** | -3.61 | -3.16 | -3.4 |
| **62** | -3.46 | -3.03 | -3.26 |
| **63** | -3.32 | -2.89 | -3.11 |
| **64** | -3.18 | -2.76 | -2.97 |
| **65** | -3.04 | -2.64 | -2.82 |
| **66** | -2.9 | -2.51 | -2.68 |
| **67** | -2.76 | -2.39 | -2.54 |
| **68** | -2.63 | -2.27 | -2.41 |
| **69** | -2.5 | -2.15 | -2.27 |
| **70** | -2.36 | -2.04 | -2.14 |
| **71** | -2.23 | -1.92 | -2 |
| **72** | -2.11 | -1.81 | -1.87 |
| **73** | -1.98 | -1.69 | -1.74 |
| **74** | -1.85 | -1.58 | -1.62 |
| **75** | -1.73 | -1.47 | -1.49 |
| **76** | -1.61 | -1.37 | -1.37 |
| **77** | -1.49 | -1.26 | -1.25 |
| **78** | -1.37 | -1.15 | -1.13 |
| **79** | -1.26 | -1.05 | -1.01 |
| **80** | -1.14 | -0.94 | -0.89 |
| **81** | -1.03 | -0.84 | -0.78 |
| **82** | -0.92 | -0.74 | -0.66 |
| **83** | -0.81 | -0.64 | -0.55 |
| **84** | -0.7 | -0.54 | -0.44 |
| **85** | -0.59 | -0.44 | -0.33 |
| **86** | -0.48 | -0.34 | -0.22 |
| **87** | -0.37 | -0.24 | -0.11 |
| **88** | -0.26 | -0.15 | 0 |
| **89** | -0.16 | -0.05 | 0.11 |
| **90** | -0.05 | 0.05 | 0.22 |
| **91** | 0.06 | 0.14 | 0.32 |

| | | | |
|---|---|---|---|
| **92** | 0.16 | 0.24 | 0.43 |
| **93** | 0.26 | 0.34 | 0.53 |
| **94** | 0.37 | 0.43 | 0.64 |
| **95** | 0.47 | 0.53 | 0.74 |
| **96** | 0.58 | 0.63 | 0.85 |
| **97** | 0.68 | 0.72 | 0.95 |
| **98** | 0.78 | 0.82 | 1.06 |
| **99** | 0.89 | 0.92 | 1.16 |
| **100** | 0.99 | 1.01 | 1.27 |
| **101** | 1.09 | 1.11 | 1.37 |
| **102** | 1.2 | 1.21 | 1.47 |
| **103** | 1.3 | 1.31 | 1.58 |
| **104** | 1.4 | 1.4 | 1.68 |
| **105** | 1.5 | 1.5 | 1.78 |
| **106** | 1.6 | 1.6 | 1.88 |
| **107** | 1.71 | 1.7 | 1.99 |
| **108** | 1.81 | 1.8 | 2.09 |
| **109** | 1.91 | 1.9 | 2.19 |
| **110** | 2.01 | 2.01 | 2.3 |
| **111** | 2.12 | 2.11 | 2.4 |
| **112** | 2.22 | 2.21 | 2.5 |
| **113** | 2.32 | 2.31 | 2.61 |
| **114** | 2.43 | 2.42 | 2.71 |
| **115** | 2.53 | 2.53 | 2.82 |
| **116** | 2.64 | 2.63 | 2.93 |
| **117** | 2.75 | 2.74 | 3.03 |
| **118** | 2.86 | 2.85 | 3.14 |
| **119** | 2.97 | 2.97 | 3.26 |
| **120** | 3.08 | 3.08 | 3.37 |
| **121** | 3.19 | 3.2 | 3.49 |
| **122** | 3.31 | 3.32 | 3.61 |
| **123** | 3.43 | 3.44 | 3.73 |
| **124** | 3.56 | 3.57 | 3.85 |
| **125** | 3.69 | 3.7 | 3.98 |
| **126** | 3.82 | 3.84 | 4.12 |
| **127** | 3.97 | 3.99 | 4.26 |
| **128** | 4.12 | 4.14 | 4.41 |
| **129** | 4.27 | 4.3 | 4.57 |
| **130** | 4.44 | 4.47 | 4.74 |
| **131** | 4.63 | 4.66 | 4.93 |
| **132** | 4.83 | 4.87 | 5.14 |
| **133** | 5.07 | 5.1 | 5.37 |
| **134** | 5.34 | 5.38 | 5.64 |

| | | | |
|---|---|---|---|
| **135** | 5.67 | 5.71 | 5.98 |
| **136** | 6.12 | 6.17 | 6.43 |
| **137** | 6.86 | 6.91 | 7.17 |
| **138** | 8.1 | 8.15 | 8.41 |

Convert Table in the Social Emotion Area

| SCORE | V2_MEASURE | V3_MEASURE | V3_F_MEASURE |
|---|---|---|---|
| **0** | -7.26 | -6.74 | -6.95 |
| **1** | -5.9 | -5.66 | -5.86 |
| **2** | -5.02 | -5.02 | -5.22 |
| **3** | -4.44 | -4.62 | -4.82 |
| **4** | -4 | -4.31 | -4.51 |
| **5** | -3.64 | -4.05 | -4.26 |
| **6** | -3.32 | -3.84 | -4.04 |
| **7** | -3.04 | -3.64 | -3.85 |
| **8** | -2.78 | -3.46 | -3.68 |
| **9** | -2.55 | -3.3 | -3.52 |
| **10** | -2.33 | -3.15 | -3.37 |
| **11** | -2.13 | -3 | -3.23 |
| **12** | -1.93 | -2.87 | -3.1 |
| **13** | -1.75 | -2.74 | -2.97 |
| **14** | -1.57 | -2.62 | -2.85 |
| **15** | -1.4 | -2.5 | -2.74 |
| **16** | -1.23 | -2.39 | -2.63 |
| **17** | -1.07 | -2.28 | -2.52 |
| **18** | -0.92 | -2.17 | -2.42 |
| **19** | -0.77 | -2.07 | -2.32 |
| **20** | -0.62 | -1.97 | -2.22 |
| **21** | -0.47 | -1.88 | -2.12 |
| **22** | -0.32 | -1.79 | -2.03 |
| **23** | -0.18 | -1.69 | -1.93 |
| **24** | -0.04 | -1.6 | -1.84 |
| **25** | 0.1 | -1.51 | -1.75 |
| **26** | 0.24 | -1.42 | -1.66 |
| **27** | 0.38 | -1.34 | -1.57 |
| **28** | 0.52 | -1.25 | -1.48 |
| **29** | 0.66 | -1.16 | -1.39 |
| **30** | 0.8 | -1.08 | -1.29 |
| **31** | 0.94 | -0.99 | -1.2 |
| **32** | 1.08 | -0.9 | -1.11 |
| **33** | 1.23 | -0.81 | -1.02 |

| | | | |
|---|---|---|---|
| **34** | 1.37 | -0.72 | -0.92 |
| **35** | 1.52 | -0.62 | -0.82 |
| **36** | 1.68 | -0.53 | -0.73 |
| **37** | 1.84 | -0.43 | -0.62 |
| **38** | 2 | -0.33 | -0.52 |
| **39** | 2.17 | -0.22 | -0.41 |
| **40** | 2.35 | -0.11 | -0.29 |
| **41** | 2.54 | 0.01 | -0.17 |
| **42** | 2.74 | 0.14 | -0.04 |
| **43** | 2.95 | 0.28 | 0.1 |
| **44** | 3.19 | 0.44 | 0.26 |
| **45** | 3.46 | 0.62 | 0.44 |
| **46** | 3.77 | 0.83 | 0.65 |
| **47** | 4.14 | 1.09 | 0.91 |
| **48** | 4.64 | 1.45 | 1.27 |
| **49** | 5.42 | 2.06 | 1.88 |
| **50** | 6.69 | 3.16 | 2.98 |

REFERENCES


Andrich, D. (1988). Rasch models for measurement (Vol. 68): Sage.

Arrindell, W. A., & Van der Ende, J. (1985). An empirical test of the utility of the observations-to-variables ratio in factor and components analysis. Applied Psychological Measurement, 9(2), 165-178.

Bagnato, S. J., Neisworth, J. T., & Capone, A. (1986). Curriculum-based assessment for the young exceptional child: Rationale and review. Topics in early childhood special education, 6(2), 97-110.

Bagnato, S. J., & Neisworth, J. T. (1999). Collaboration and teamwork in assessment for early intervention. Child and Adolescent Psychiatric Clinics, 8(2), 347-363.

Bailey, E., & Bricker, D. (1986). A psychometric study of a criterion-referenced assessment instrument designed for infants and young children. Journal of the Division for Early Childhood, 10(2), 124-134.

Bertenthal, B. I., & Clifton, R. K. (1998). Perception and action.

Black, C. (1974). First Chance Network: Directory and Abstracts.

Bond, T., Yan, Z., & Heene, M. (2020). Applying the Rasch model: Fundamental measurement in the human sciences: Routledge.

Bond, T. G., Fox, C. M., & Lacey, H. (2007). Applying the Rasch model: Fundamental measurement. Paper presented at the in the social sciences (2nd).

Bond, T.G., & Fox, C.M. (2013). Applying the Rasch model: Fundamental measurement in the human sciences. Hove, UK: Psychology Press.

Boone, W. J., Staver, J. R., & Yale, M. S. (2013). Rasch analysis in the human sciences: Springer.

Boone, W. J. (2016). Rasch analysis for instrument development: why, when, and how? CBE—Life Sciences Education, 15(4), rm4.

Bricker, D., Clifford, J., Yovanoff, P., Pretti-Frontczak, K., Waddell, M., Allen, D., & Hoselton, R. (2008). Eligibility determination using a curriculum-based assessment: A further examination. Journal of Early Intervention, 31(1), 3-21.

Bricker, D., & Pretti-Frontczak, K. (1996). AEPS measurement for three to six years (Vol. 3). Baltimore MD: Brookes.

Bricker, D., & Pretti-Frontczak, K. (1998). Treatment validity of the Assessment, Evaluation, and Programming System Test for three to six years. International Division of Early Childhood, New Orleans, LA.

Budescu, D. (1985). Efficiency of linear equating as a function of the length of the anchor test. Journal of educational measurement, 22(1), 13-20.

Ricker, K. L., & von Davier, A. A. (2007). The impact of anchor test length on equating results in a nonequivalent groups design. ETS Research Report Series, 2007(2), i-19.

Bricker, D., & Waddell, M. (2002). Curriculum for Birth to Three Years. Assessment, Evaluation, and Programming System for Infants and Children (AEPS): ERIC.

Bricker, D., & Waddell, M. (1996). AEPS Curriculum for Three to Six Years. Volume 4. Baltimore: Paul H. In: Brooks Publishing Co.

Bricker, D., Yovanoff, P., Capt, B., & Allen, D. (2003). Use of a curriculum-based measure to corroborate eligibility decisions. Journal of Early Intervention, 26(1), 20-30.

Cascio, E. U. (2021). Early childhood education in the United States: What, when, where, who, how, and why (No. w28722). National Bureau of Economic Research.

Castaneda-Villa, N., & James, C. (2007). Objective source selection in blind source separation of AEPs in children with cochlear implants. Paper presented at the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

CDC. (2021). Child Development - Milestone. Retrieved from https://www.cdc.gov/ncbddd/actearly/milestones/index.html

Choi, Y. J., & Asilkalkan, A. (2019). R packages for item response theory analysis: Descriptions and features. Measurement: Interdisciplinary research and perspectives, 17(3), 168-175.

Dennis, D. V. (2017). Learning from the past: What ESSA has the chance to get right. The Reading Teacher, 70(4), 395-400.

Donoghue, J. R., & Isham, S. P. (1998). A comparison of procedures to detect item parameter drift. Applied Psychological Measurement, 22(1), 33-51.

Early Childhood Technical Assistance Center & Center for IDEA Early Childhood Data Systems. (2017). State child outcomes measurement system framework. Retrieved from http://ectacenter.org/eco/pages/childoutcomes.asp#frameworks

Education, V. D. o. (2019). K-12 Inclusive Practices Guide. Retrieved from https://www.doe.virginia.gov/.../inclusive/k-12-inclusive-practices-guide.pdf

Fischer, L., Rohm, T., Carstensen, C. H., & Gnambs, T. (2021). Linking of Rasch-scaled tests: Consequences of limited item pools and model misfit. Frontiers in psychology, 12.

Gao, X., & Grisham-Brown, J. (2011). The Use of Authentic Assessment to Report Accountability Data on Young Children's Language, Literacy and Pre-Math Competency. International Education Studies, 4(2), 41-53.

Grisham-Brown, J., Hallam, R. A., & Pretti-Frontczak, K. (2008). Preparing Head Start personnel to use a curriculum-based assessment: An innovative practice in the "age of accountability". Journal of Early Intervention, 30(4), 271-281.

Grisham-Brown, J., & Pretti-Frontczak, K. (2011). Assessing Young Children in Inclusive Settings: The Blended Practices Approach: ERIC.

Grisham, J., Waddell, M., Crawford, R., & Toland, M. (2021). Psychometric Properties of the Assessment, Evaluation, and Programming System for Infants and Children–Third Edition (AEPS-3). Journal of Early Intervention, 43(1), 24-37.

Hallam, R. A., Lyons, A. N., Pretti-Frontczak, K., & Grisham-Brown, J. (2014). Comparing apples and oranges: The mismeasurement of young children through the mismatch of assessment purpose and the interpretation of results. Topics in early childhood special education, 34(2), 106-115.

Hamilton, D. A. (1995). The utility of the assessment evaluation programming system in the development of quality IEP goals and objectives for young children, birth to three, with visual impairments. University of Oregon.

He, Y., Cui, Z., Fang, Y., & Chen, H. (2013). Using a linear regression method to detect outliers in IRT common item equating. Applied Psychological Measurement, 37(7), 522-540.

Hebbeler, K. M., Smith, B. J., & Black, T. L. (1991). Federal early childhood special education policy: A model for the improvement of services for children with disabilities. Exceptional Children, 58(2), 104-112.

Heo, K. H., & Squires, J. (2012). Cultural adaptation of a parent completed social emotional screening instrument for young children: Ages and stages

Hu, H., Rogers, W. T., & Vukmirovic, Z. (2008). Investigation of IRT-based equating methods in the presence of outlier common items. Applied Psychological Measurement, 32(4), 311-333.

Huang, C. Y., & Shyu, C. Y. (2003). The impact of item parameter drift on equating. In Annual Meeting of the National Council on Measurement in Education.

Kane, M. (2010). Validity and fairness. Language testing, 27(2), 177-182.

Kolen, M. J., & Brennan, R. L. (2014). Linking. In Test Equating, Scaling, and Linking (pp. 487-536). Springer, New York, NY.

Kusner, M., Sun, Y., Kolkin, N., & Weinberger, K. (2015). From word embeddings to document distances. Paper presented at the International conference on machine learning.

Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. Paper presented at the International conference on machine learning.

Lee, D. D., Bagnato, S. J., & Frontczak, K. P. (2015). Utility and validity of authentic assessments and conventional tests for international early childhood intervention

purposes: Evidence from US national social validity research. Journal of Intellectual Disability-Diagnosis and Treatment, 3(4), 164-176.

Li, D., Jiang, Y., & von Davier, A. A. (2012). The accuracy and consistency of a series of IRT true score equatings. Journal of educational measurement, 49(2), 167-189. Hu, H., Rogers, W. T., & Vukmirovic, Z. (2008). Investigation of IRT-based equating methods in the presence of outlier common items. Applied Psychological Measurement, 32(4), 311-333.

Linacre, J. (1998). Structure in Rasch residuals: Why principal components analysis (PCA)? Rasch Measurement Transactions, 12 (2), 636.

Linacre, J. (2003). Size vs. significance: Standardized chi-square fit statistic. Rasch Measurement Transactions, 17(1), 918.

Linacre, J. (2021). A user ‘s Guide to WINSTEPS Rasch-Model Computer Programs. Winsteps. com. Retrieved from https://www.winsteps.com/winman/copyright.htm

Linacre, J. M. (2000). Comparing and choosing between "Partial Credit Models" (PCM) and "Rating Scale Models" (RSM). Rasch Measurement Transactions, 14(3), 768. Retrieved from https://www.rasch.org/rmt/rmt143k.htm

Linacre, J. M. (2009). Investigating dimensionality. Linacre, JM (2009): Practical Rasch Measurement. Further Topics. Online Course Statistics. com. Arlington: Virginia: The Institute for Statistics Education.

Linacre, J. M. (2023). Winsteps help. Retrieved from www.winsteps.com/winman/webpage.htm.

Liu, J., Sinharay, S., Holland, P. W., Curley, E., & Feigenbaum, M. (2011). Test score equating using a Mini-Version anchor and a midi anchor: A case study using SAT® data. Journal of educational measurement, 48(4), 361-379.

Lynch, B. K., & McNamara, T. F. (1998). Using G-theory and many-facet Rasch measurement in the development of performance assessments of the ESL speaking skills of immigrants. Language testing, 15(2), 158-180.

Macy, M. G., Bricker, D. D., & Squires, J. K. (2005). Validity and reliability of a curriculum-based assessment approach to determine eligibility for Part C services. Journal of Early Intervention, 28(1), 1-16.

McLean, M. (2005). Using curriculum-based assessment to determine eligibility: Time for a paradigm shift? Journal of Early Intervention, 28(1), 23-27.

Mihalcea, R., Corley, C., & Strapparava, C. (2006). Corpus-based and knowledge-based measures of text semantic similarity. Paper presented at the Aaai.

Michaelides, M. P., & Haertel, E. H. (2014). Selection of common items as an unrecognized source of variability in test equating: A bootstrap approximation assuming random sampling of common items. Applied Measurement in Education, 27(1), 46-57.

Newborg, J., & Company, R. P. (2005). Battelle developmental inventory: Riverside Pub.

Neisworth, J. T., & Bagnato, S. J. (2000). Recommended practices in assessment. DEC recommended practices in early intervention/early childhood special education, 17-27.

Notari, A. R., & Drinkwater, S. G. (1991). Best practices for writing child outcomes: An evaluation of two methods. Topics in early childhood special education, 11(3), 92-106.

Pretti-Frontczak, K., & Bricker, D. (2000). Enhancing the quality of individualized education plan (IEP) goals and objectives. Journal of Early Intervention, 23(2), 92-105.

Pretti-Frontczak, K., & Bricker, D. (2000). Enhancing the quality of individualized education plan (IEP) goals and objectives. Journal of Early Intervention, 23(2), 92-105.

Rasch, G. (1993). Probabilistic models for some intelligence and attainment tests: ERIC.

Reichow, B., Boyd, B. A., Barton, E. E., & Odom, S. L. (2016). Handbook of early childhood special education: Springer.

Ricker, K. L., & von Davier, A. A. (2007). The impact of anchor test length on equating results in a nonequivalent groups design. ETS Research Report Series, 2007(2), i-19.

Sandall, S., McLean, M. E., & Smith, B. J. (2000). DEC recommended practices in early intervention/early childhood special education: ERIC.

Schachter, R. E., Piasta, S., & Justice, L. (2020). An Investigation into the Curricula (and Quality) Used by Early Childhood Educators. NHSA Dialog, 23(2).

Schütze, H., Manning, C. D., & Raghavan, P. (2008). Introduction to information retrieval (Vol. 39, pp. 234-265). Cambridge: Cambridge University Press.

Shin, S.-H. (2009). How to treat omitted responses in Rasch model-based equating. Practical Assessment, Research, and Evaluation, 14(1), 1.

Sinharay, S., & Holland, P. W. (2007). Is it necessary to make anchor tests mini-versions of the tests being equated or can some restrictions be relaxed? Journal of educational measurement, 44(3), 249-275.

Slentz, K. L. (1987). Evaluating The Instructional Needs of Young Children With Handicaps: Psychometric Adequacy of The Evaluation And Programming System--Assessment Level II (EPS-II).

Smith, R. M. (1996). A comparison of the Rasch separate calibration and between-fit methods of detecting item bias. Educational and Psychological Measurement, 56(3), 403-418.

Smith, R. M. (1995). Using item mean squares to evaluate fit to the Rasch model.

Straka, E. A. (1996). Assessment of young children for communication delays.

Sundberg, R. (2019). Statistical modelling by exponential families (Vol. 12): Cambridge University Press.

Toland, M. D., Grisham, J., Waddell, M., Crawford, R., & Dueber, D. M. (2021). Scale Evaluation and Eligibility Determination of a Field-Test Version of the Assessment, Evaluation, and Programming System Third Edition. Topics in early childhood special education, 0271121420981712.

Vanderheyden, A. M. (2005). Intervention-driven assessment practices in early childhood/early intervention: Measuring what is possible rather than what is present. Journal of Early Intervention, 28(1), 28-33.

Vinovskis, M. A. (2008). The birth of Head Start: Preschool education policies in the Kennedy and Johnson administrations: University of Chicago Press.

Wang, H.-T., Sandall, S. R., Davis, C. A., & Thomas, C. J. (2011). Social skills assessment in young children with autism: A comparison evaluation of the SSRS and PKBS. Journal of Autism and Developmental Disorders, 41(11), 1487-1495.

Waterbury, G. T. (2019). Missing data and the Rasch model: The effects of missing data mechanisms on item parameter estimation. Journal of applied measurement, 20(2), 154-166.

Winchell, B. N. (2011). A critical examination of the technical adequacy of a curriculum-based assessment using Rasch analyses. Kent State University,

Wright, B. D. (1999). Fundamental measurement for psychology. The new rules of measurement: What every psychologist and educator should know, 65-104.

Wright, B. D., & Masters, G. N. (1982). Rating scale analysis: MESA press.

Wright, B. D., & Linacre, J. M. (1989). Observations are always ordinal; measurements, however, must be interval. Archives of physical medicine and rehabilitation, 70(12), 857-860.

Wright, R. J. (2007). Educational assessment: Tests and measurements in the age of accountability. Sage Publications.

Yang, W.-L., & Houang, R. T. (1996). The Effect of Anchor Length and Equating Method on the Accuracy of Test Equating: Comparisons of Linear and IRT-Based Equating Using an Anchor-Item Design.

Ye, M., & Xin, T. (2014). Effects of item parameter drift on vertical scaling with the nonequivalent groups with anchor test (NEAT) design. Educational and Psychological Measurement, 74(2), 227-235.

Zhao, Y., & Hambleton, R. K. (2017). Practical consequences of item response theory model misfit in the context of test equating with mixed-format test data. Frontiers in Psychology, 8, 484.

VITA

Name: Yuyan (Summer) Xia

Education

- M. Ed, Interdisciplinary Early Childhood Education, University of Kentucky 2016 -2018

Work Experience

- Research Assistant (Evaluation center, University of Kentucky) 2020 - present
- Teaching Assistant (Early childhood lab, University of Kentucky) 2016 -2019

Publications

Sampson, S. O., Xia, Y., Parsons, J. M., & Cardarelli, R. (2022). Reduce, reuse, and recycle: Saving resources by repurposing data to address evaluation questions.

Ke, S., Xia, Y., & Zhang, J. (2020). What really matters in early bilingual and biliteracy acquisition. Home language and literacy input in Chinese heritage language learners.