



2016

CONSTRUCTION AND VALIDATION OF A HOLISTIC EDUCATION SCHOOL EVALUATION TOOL USING MONTESSORI ERDKINDER PRINCIPLES

Anthony Philip Setari

University of Kentucky, anthony.setari@uky.edu

Digital Object Identifier: <http://dx.doi.org/10.13023/ETD.2016.167>

Recommended Citation

Setari, Anthony Philip, "CONSTRUCTION AND VALIDATION OF A HOLISTIC EDUCATION SCHOOL EVALUATION TOOL USING MONTESSORI ERDKINDER PRINCIPLES" (2016). *Theses and Dissertations--Education Science*. Paper 12.
http://uknowledge.uky.edu/edsc_etds/12

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 Science 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.

Anthony Philip Setari, Student

Dr. Kelly D. Bradley, Major Professor

Dr. Robert Shapiro, Director of Graduate Studies

CONSTRUCTION AND VALIDATION OF A HOLISTIC EDUCATION
SCHOOL EVALUATION TOOL USING MONTESSORI ERDKINDER PRINCIPLES

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

Anthony Philip Setari

Co-Directors: Dr. Kelly D. Bradley, Professor of Educational Policy & Evaluation
and Dr. Michael D. Toland, Associate Professor of Educational, School, &
Counseling Psychology

Lexington, KY

Copyright © Anthony Philip Setari 2016

ABSTRACT OF DISSERTATION

CONSTRUCTION AND VALIDATION OF A HOLISTIC EDUCATION SCHOOL EVALUATION TOOL USING MONTESSORI ERDKINDER PRINCIPLES

The purpose of this study was to construct a holistic education school evaluation tool using Montessori Erdkinder principles, and begin the validation process of examining the proposed tool. This study addresses a vital need in the holistic education community for a school evaluation tool. The tool construction process included using Erdkinder literature to justify the development of each item through the use of an item matrix, ultimately leading to the development of the 23 item formative Montessori Erdkinder School Evaluation Survey. The validation process included a series of three Rasch Rating Scale Model analyses with data from a sample school. The validation process used item anchoring estimates from the earlier analyses in the later analyses and included determining the tool's dimensionality, reliability, item fit, possible differential item functioning, and comparing the order of item difficulty levels to the holistic model of Maslow's hierarchy of needs. Results of the study showed that six items had issues with fit and would need to be revised, and that the items in the cognitive and moral facet will need to be revised to better match Maslow's model. This study provides the foundation for the development of a holistic education evaluation or accreditation system, and constructed a resource that could be directly implemented in schools.

KEYWORDS: Rasch Rating Scale Model, Holistic Education, Montessori Erdkinder, School Evaluation, Education Policy

Anthony P. Setari

April 18, 2016

Date

CONSTRUCTION AND VALIDATION OF A HOLISTIC EDUCATION
SCHOOL EVALUATION TOOL USING MONTESSORI ERDKINDER PRINCIPLES

By

Anthony Philip Setari

Kelly D. Bradley, Ph.D.

Co-Director of Dissertation

Michael D. Toland, Ph.D.

Co-Director of Dissertation

Robert Shapiro, Ph.D.

Director of Graduate Studies

April 18, 2016

For Renee

“Well, there I was, ’way off my ambitions, getting deeper in love every minute,
and all of a sudden I didn’t care.”

- F. Scott Fitzgerald, “The Great Gatsby”

ACKNOWLEDGMENTS

I would like to begin this acknowledgment section by thanking my co-advisors and mentors, Dr. Kelly Bradley and Dr. Michael Toland. Dr. Bradley has been a tireless supporter of my work and was the person who taught me how to be a quantitative researcher. She is a marvel of a person and an academic that I cannot thank enough for all she has done in supporting me to reach this milestone. Dr. Toland has challenged me to be an expert in the field, working to assure I have the precision of skill to be an academic leader in educational measurement. He has pushed me towards a level of excellence I did not realize I could achieve at this point in my career, and for that I am grateful. Ultimately, the combined efforts of Dr. Bradley and Dr. Toland are largely responsible for making me the academic that I am and making this dissertation possible.

I would also like to thank my other committee members, Dr. Kate Akers and Dr. James Ziliak. Dr. Akers is a wonderful person, a brilliant academic, and someone who saw great possibilities for me. Support like that provided by Dr. Akers is immeasurable but truly special, and I thank her tremendously. Dr. Ziliak has been a supporter of me for years and encouraged me to continue my academic pursuits into the field of education. Dr. Ziliak has pushed me to think broader and apply my work to the larger field of education. I thank him for reminding me of the great implications my work can have and for supporting me for all this time. I would also like to thank Dr. Lee Ann Jung for being the outside examiner of my dissertation and for all her useful feedback.

In addition, I would like to thank all of the individuals who have shaped me into the researcher that I have become. I have always enjoyed academics, and that is largely the result of the wonderful people I have encountered throughout my K-12 education in

the Broward County Public School System, in my undergraduate work at Georgia Southern University, and my graduate work at The Johns Hopkins University and here at the University of Kentucky. I would like to thank my past middle school students in the Baltimore City School System that serve as a reminder as to why my work in education matters and for whom my work ultimately serves.

Finally, I would like to thank my family and friends. A special group of friends supported me throughout this dissertation process and deserve recognition: Chris Booth, Jakub Mikuška, Zuzana Gaššová, Brian and Stephanie Armes, and Eric and Courtney Casero. My family has always supported my schooling and certainly deserves to be recognized. My family members Michelle Mazzacca, Antoinette Setari, Sue Kwap, Eve Holliday, and Cindy and Roger Didio, are wonderful people whom I love tremendously and owe a great deal. I would like to particularly recognize my family members Ann Masi, Rose Masi, and Philip Masi, whom I owe the world. I must thank my dachshund Shorticus, my Italian greyhound Baci, and my Burmese tuxedo cat Oliver, for staying up with me many late nights as I completed my coursework and wrote this work. To my wife Renee, there is simply no amount of gratitude I can show for all she has done. Renee is truly an incredible person with an amazing ability to love, and for everything that she is, has been, and will be for me, I am thankful.

I would like to conclude by thanking the Roman Catholic Church for strengthening my faith and God for making this all possible. Finally, I would like to thank you, old sport. I greatly appreciate you taking the time to read my work.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER I: INTRODUCTION.....	1
Purpose.....	5
Research Questions	8
Contribution	9
Terms and Definitions.....	11
Assumptions.....	12
Conclusion	13
CHAPTER II: LITERATURE REVIEW	14
Montessori’s Philosophy on Adolescence	14
Montessori’s Erdkinder.....	16
History.....	18
Curriculum	20
Interpretations.....	22
Holistic Education.....	23
Item Response Theory	25
One-Parameter Logistic Model	26
Rasch Model	29
Differential Item Functioning.....	31
Rating Scale Model	32
Conclusion	34
CHAPTER III: METHODOLOGY	35
Purpose and Significance.....	35
Research Questions.....	36

Construction of the Montessori Erdkinder School Evaluation Survey	36
Considerations	39
Data Collection	40
Sample.....	42
Analysis.....	43
Answering Research Questions	44
Conclusion	47
CHAPTER IV: RESULTS.....	49
Analysis Procedure	49
Samples	50
July 2015	50
October 2015	51
Analysis 1.....	52
Analysis 2.....	59
Analysis 3.....	67
Conclusion	74
CHAPTER V: DISCUSSION.....	76
Research Questions and Findings	76
Limitations	85
Conclusion	86
Appendix A. Montessori Erdkinder School Evaluation Survey (MESES).....	91
Appendix B. Item Development Matrix	93
REFERENCES	106
VITA.....	115

LIST OF TABLES

Table 1, Facets and Items.....	38
Table 2, July 2015 Sample Descriptive Statistics (N=18)	51
Table 3, October 2015 Sample Descriptive Statistics (N=36)	52
Table 4, Analysis 1 Variance Estimates	53
Table 5, Analysis 1 Item Estimates.....	55
Table 6, Analysis 1 Differential Item Functioning Estimates for Students and Parents....	59
Table 7, Analysis 2 Variance Estimates	60
Table 8, Analysis 2 Item Estimates.....	62
Table 9, Analysis 2 Differential Item Functioning Estimates for Student and Parents	66
Table 10, Analysis 3 Variance Estimates	67
Table 11, Analysis 3 Item Estimates.....	70
Table 12, Analysis 3 Differential Item Functioning Estimates for Students and Parents..	74

LIST OF FIGURES

Figure 1, Item Characteristic Curve.....	28
Figure 2, Rasch Rating Scale Model Difficulties and Thresholds.....	34
Figure 3, Analysis 1 Standardized Residual Loadings for First Contrast.....	54
Figure 4, Analysis 1 Wright Map	57
Figure 5, Analysis 2 Standardized Residual Loadings for First Contrast.....	61
Figure 6, Analysis 2 Wright Map	64
Figure 7, Analysis 3 Standardized Residual Loadings for First Contrast.....	69
Figure 8, Analysis 3 Wright Map	72

CHAPTER I

INTRODUCTION

Alternative education schools continue to gain popularity in the United States (Kena et al., 2014). State and local governments have been encouraging the development of alternative education schools as a means to lower their own costs and shift towards a free-market approach to education (Chubb & Moe, 1990; Hursh, 2006). Parents and students have sought out these alternative education schools as a way to avoid what they see as underperforming public schools overly concerned with national standards and testing (Finn, Caldwell, & Raub, 2006; Lipman, 2011). As these alternative education schools grow, evaluation tools and complete evaluation systems are needed to examine school quality in these unique educational settings.

Holistic education outcomes have become the basis for many alternative education schools (Forbes & Martin, 2004). Research on holistic education schools, specifically the long lineage of research on the outcomes of Catholic school students, shows that these non-traditional schools demonstrate positive student outcomes, particularly for minority and low-income students in the areas of higher education enrollment and standardized testing (Hoffer, 2000; R. Miller, 1990). These positive research findings, coupled with families' dissatisfaction with traditional schooling, makes holistic education methods a desirable focus for the founders of new alternative education schools. The premise of holistic education is that schools should focus on the development of a "whole-child," rather than educational outcomes alone (Mayes & Williams, 2013; R. Miller, 1990; J. P. Miller, 2010). This focus leads holistic education schools to develop curricula that encourage students' psychological, emotional, social,

and spiritual development (Mayes & Williams, 2013; R. Miller, 1990; J. P. Miller, 2010). Holistic education often aims to encourage a sense of peace and communal understanding in students, as well as to find methods for students to reach their full potential through personal expression and personality development. Holistic education does not exist as its own cohesive theory, but rather refers to schools and academics whose ideas encompass many of the wider aims of holistic education. Importantly, holistic education is certainly not a new concept, and has served as the basis for many experimental schools in the United States and endured as the basis for many religiously affiliated schools (R. Miller, 1990; J. P. Miller, 2010). Given the history of holistic education, new alternative education schools have the opportunity to select complete philosophies and even materials from the past or other sources.

Montessori high schools are among the holistically focused alternative education school types that began to increase in the 2000s (Kahn, 2011). Although decreasing in number since peaking in the 1960s and 1970s (Barker, 2011), the 2016 North American Montessori Teachers' Association (NAMTA) school directory included over 100 Montessori high schools (NAMTA, 2016). Montessori high schools are known for fostering the educational, social, and psychological development of adolescents, within an engaging yet nurturing environment (Kahn, 2011; Rathunde, 2001; Tornar, 2011). Parents seek out Montessori schools because they believe these schools will provide their children an education that stresses moral and social development (Parker, 2007; Zarybnisky, 2010). Parents view these schools as having caring teachers who can provide individual attention to children and create a positive educational experience. Given the holistic education focus of Montessori high schools and the positive perception of

Montessori education by parents, there is reason to believe that these schools will continue to grow in number and student enrollment.

Although Montessori high schools have grown in popularity, one of the vital materials missing from their available resources is a rigorously examined common school evaluation tool that determines how well schools are implementing adolescent holistic education according to Montessori philosophy. Such a tool would support school administrators attempting to develop a school that aligns to Montessori's adolescent schooling philosophy. For Montessori high schools, there are several probable reasons for why no common school evaluation tool exists. The most likely reason is that Montessori left behind scant details on what a Montessori high school should incorporate (Barker, 2011). In her writings on developing a Montessori high school, Montessori focused on discussing the broad components of setting up such a school, such as types of learning experiences for students and where to establish schools, not particularly delving into detail on any matter (Barker, 2011; Montessori, 1973). To make up for this lack of information from Montessori herself, the Montessori academic community (e.g. the teachers, administrators, and researchers who are interested in studying the effects of the schools and developing better Montessori schools) have worked to interpret these principles and expand upon them (Barker, 2011; Kahn, 2011; Kahn & Pendleton, 2007). This has led to Montessori high schools ranging in their composition and approaches to schooling, yet each maintaining the intent to follow the principles of Montessori in their unique expanded formats (American Montessori Society, 2015; Kahn, 2011). Attempting to make an evaluation tool that covers the range of Montessori high school interpretations is a challenging task. Schools may vary dramatically in educational methods and goals,

with some attempting to create the outdoor experiences Montessori described, and others trying to infuse methods used with younger children.

This lack of a common school evaluation tool can also be attributed to several other issues. First, there is a lack of overall research in the field of Montessori high school education, particularly in the area of measurement for evaluation and assessment purposes (Barker, 2011; Kahn & Pendleton, 2007; NAMTA, 2015b). Although research articles and journals on Montessori high school education do exist, much of this work is committed to examining educational outcomes in schools and on the refinement of Montessori secondary teaching methods. In addition to the lack of research on Montessori high schools, there is resistance toward evaluation and assessment tools in the Montessori community (Montessori Foundation, 2015; Pottish-Lewis, 2013). This resistance can be attributed to the warnings against evaluation and assessment tools by Montessori herself, and her warnings about the ability of assessment tools to accurately capture students' outcomes within the unique school contexts (Montessori, 1973; Montessori Foundation, 2015; NAMTA, 2015b; Pottish-Lewis, 2013). Finally, Montessori high schools seek to develop holistic education aspects of the child beyond the cognitive, such as personality, morality, and character (Kahn, 2011; Rathunde, 2001; Tornar, 2011), which are abstract concepts challenging to operationalize in a school evaluation tool. Each of these described issues has possibly contributed to the absence of a Montessori high school common evaluation tool, and represents possible barriers to the creation of such a tool.

A common school evaluation tool rooted in the holistic education principles of Montessori's philosophy for adolescent education and developed for Montessori high

schools has the potential to be of great value to both individual schools and families considering these schools for their children. Montessori high schools seek to demonstrate that they can develop aspects beyond the cognitive, and an evaluation tool incorporating these aspects would allow them to do so (Kahn, 2011; Rathunde, 2001; Tornar, 2011). A common school evaluation tool would allow individual schools to market themselves to interested families and stakeholders. In addition, a common school evaluation tool would allow for Montessori high schools to make direct comparisons with one another for the purposes of school improvement. The results from these school evaluations would provide families considering a Montessori high school with additional information about how well the school supports Montessori's beliefs on adolescent education, and families could then use this information to make an informed choice about whether or not to send their adolescent to the school. Each of these reasons supports the practical development of a common Montessori high school evaluation tool. The construction and validation of such a tool, however, goes beyond just benefitting Montessori high schools, as it could prove useful for any schools concerned with holistic education outcomes. The Montessori adolescent education philosophy, known as the Erdkinder, is a model for holistic education. Erdkinder outcomes are valued holistic education outcomes that are transferable to any school concerned with holistic education. Therefore, other types of holistic education schools would gain many of the same benefit as Montessori high schools from implementing a common school evaluation tool.

Purpose

The purpose of this study was to construct a holistic education school evaluation tool based on holistic education outcomes that can be implemented in the range of

Montessori high schools that have been established, and to begin the validation process of examining the proposed tool. The tool in this study was designed to evaluate the presence of Montessori's holistic education Erdkinder principles for appropriate adolescent development within Montessori high school settings. This study uses perceived student outcomes and experiences as the measure of determining school quality. Presently, there is no school evaluation tool in place for widespread use by Montessori high schools that examines holistic education school outcomes. NAMTA (2015a) provides a single example of a Montessori high school evaluation tool; however, there are several issues with the example evaluation tool. The example provided by NAMTA includes only a few educational outcomes, many outcomes are specific to farm schools and would be challenging to find in urban Erdkinder schools, and the outcomes are not presented in a survey format that can be easily implemented in schools (Hershey Montessori School Adolescent Community, 2015a). Creating a new evaluation tool for Montessori high schools that uses Erdkinder principles would allow these schools to be more easily evaluated based on their common values.

To create an effective evaluation tool, a strong theoretical and research backing is needed (Bond & Fox, 2007; De Ayala, 2009). The tool development process in this study was guided by the student outcomes proposed by Kahn in his work "Eight pictures at an exhibition: A Montessori retrospective on the discovery of the adolescent" (2011). Kahn is considered a leading expert in the Montessori Erdkinder academic community, and given Montessori's lack of writing on Erdkinder, provides a modern authoritative view on Montessori's Erdkinder (Kahn, 2011; Ludwig, 2011). Support for these outcomes was then sought from Montessori's Erdkinder principles, as presented in her seminal work

From Childhood to Adolescents: Including Erdkinder and the Function of the University (1973), and her subsequent lectures on the topic of adolescent development (Montessori 2011a; 2011c). If specific support could not be identified from Montessori herself or from the additional literature on Montessori high schools, then the item was removed. The final tool constructed included 23 items focused on Erdkinder principles separated into four holistic education facets: cognitive, social, emotional, and moral. Each item included a Likert-type format with four categories asking respondents how strongly they disagree or agree with a statement. In addition, demographic items and a series of open-ended items for usage by individual schools were included. The tool itself was created using the fundamentals of quality-tool design to reduce issues with response clarity (De Ayala, 2009; Nardi, 2006; Willis, 1999).

To determine the quality of an evaluation tool, a pilot validation study should be conducted. To validate the proposed tool, a validation procedure using the Rasch model analysis was developed and implemented. The procedure involved conducting a series of Rasch model analyses. Rasch model analysis is a common means of validating a tool as it allows for the examination of issues at the item level, specifically, the Rasch model examines how items are functioning within a tool based on people's responses (Bond & Fox, 2007; Fisher, 2006; Royal & Elahi, 201). In addition, the Rasch model assigns a difficulty level to each item and an ability level for each person. Although, the Rasch model is typically applied to dichotomous item responses, the model can also be used with polytomous item response data, which may be more appropriate for tools that are not standard academic assessments (Andrich, 1978; Bond & Fox, 2007; Wright &

Masters, 1982). The Rasch model is a proven means of refining and validating an evaluation tool.

Research Questions

The research questions guiding this study were designed to support the validation process of examining a newly created tool. A total of four research questions were created for this study:

1. How well does the evaluation tool measure the latent trait holistic education concept of the Montessori Erdkinder school principles?
2. How well do the individual items fit the Montessori Erdkinder school principles latent trait reflected on the evaluation tool?
3. To what degree do the item difficulty levels from items of specific holistic education facets compare to the theoretical item difficulty levels of another holistic model, specifically, Maslow's hierarchy of needs?
4. Are there differences in how people from different groups, with similar levels of the latent trait, are able to respond to items?

Given that this study uses a series of Rasch model analyses, only the final analysis conducted was used to answer the research questions. The first research question was answered by examining the results of Rasch principal components analysis of the residual variances reported. The second research question was answered by examining item infit and outfit statistics. The third research question was answered by examining the item difficulty estimates and comparing the order of item difficulty estimates from the specified facets to the theoretical item difficulty order of Maslow's hierarchy of needs.

The fourth research question was answered by conducting a differential item functioning analysis on student and parent responses.

Contribution

The results of this study have great implications for both the Montessori secondary community and the wider field of alternative schools concerned with holistic education. For the Montessori community, this study creates a needed common school evaluation tool that can be used in Montessori high schools with students, teachers, staff, and parents. This tool can be used by Montessori high schools' stakeholders to evaluate the quality of implementation of Erdkinder principles in these unique schools. Schools can use the results of this tool to set school improvement goals and make data-driven decisions for demonstrating Erdkinder principles. This tool has the potential to provide modern Montessori high schools with an evaluation tool that can be used for in-school decision-making.

For the Montessori community, this evaluation tool also creates a means to compare the performance of schools in a quantitative manner. By creating a common evaluation tool, specific schools can be identified for exceptional performance with respect to Erdkinder implementation. Given that Montessori schools seek to exemplify a Montessori approach to education, these exceptional Erdkinder schools may be examined further for best practices that can be implemented in other schools. In addition, implementation of this evaluation tool in Montessori high schools provides a way to demonstrate to prospective students and parents that a school represents the Erdkinder Montessori philosophy, which allows families to make informed decisions about student

enrollment. Allowing students and parents to make comparisons between schools is valuable in the growing free-market education system.

The Montessori community can also use the results of implementing this evaluation tool as evidence for grants and other funding agencies. Montessori schools struggle to report their holistic student growth outcomes in a quantitative manner. By using this tool, Montessori high schools will be able to justify possible funding dollars and demonstrate the presence of valued holistic education principles in their schools. Montessori high schools can also use this same evidence with school governing boards who wish to see that schools are meeting holistic education expectations. Data from the evaluations provide schools with a quantitative means of demonstrating the presence of holistic education outcomes.

For the wider holistic education community of schools, they would receive many of the same benefits as Montessori high schools if they were to implement the described tool. These other schools would have a school evaluation tool capable of examining the presence of holistic education outcomes, which could provide data for school improvement, grants, marketing, and many other purposes. Notably, any of these schools existing as public charters are subject to the same or similar evaluation standards as traditional public schools in several states; thus, this tool could prove useful for the demonstrative evaluations and data-driven decision-making to which charter schools are subjected (Fryer, 2012; Jordan, 2013). The development of the tool in this study would provide holistic education schools of all types with a long-sought, rigorously validated school evaluation tool that could be used for examining holistic education components of their schools.

Terms and Definitions

Erdkinder – The principles of adolescent high school education that Maria Montessori laid out primarily in *From Childhood to Adolescents: Including Erdkinder and the Function of the University* (1973), and is German for “the children of the soil” (Montessori, 1973, p. 97; Barker, 2011). Although Montessori explained some of her principles of high school education, the description of the Erdkinder was not as detailed as her plans for preschool and elementary education, leaving a great deal of room for interpretation (Barker, 2011).

Holistic education – For this study, holistic education refers to the paradigm described by R. Miller (1990), “The holistic paradigm dissolves the traditional dichotomies between mind and body, between spirit and matter. The central tenet ... is its emphasis on the integration of the inner qualities of human life with the outer physical, social world” (p. 59).

Item – For the purpose of this study, an item refers to the individual questions or statements used on a tool to measure the presence of a latent trait.

Item Response Theory – A measurement technique which examines “the probability of a person’s expected response to an item is the joint function of that person’s ability, or location on the latent trait, and one or more parameters characterizing the item” (Bond & Fox, 2007, p. 311).

Montessori high school – Refers to the variety of high schools that claim to follow a Montessori philosophy of education. There is no governing body that authorizes the usage of the Montessori name for high schools and no evaluation from a governing agency that examines if the school follows the Montessori Erdkinder philosophy

(American Montessori Society, 2015; Kahn, 2011). These schools vary greatly in their makeup and the experiences of students.

Rasch model analysis – A form of the one-parameter item response theory model where the constant value is held at 1.0 (Bond & Fox, 2007; De Ayala, 2009). The Rasch model analysis estimates item difficulty levels and person-ability levels on the same logit scale.

Tool – For the purposes of this study, a tool refers to an assessment, survey, or scale being used to measure the presence of a latent trait.

Validity – For the purposes of this study, validity is taken from Kline's (1998) definition, "A test is said to be valid if it measures what it purports to measure" (p. 34).

Assumptions

There are two types of assumptions in this study, those associated with the data collection process and those associated with the Rasch analysis. Regarding data collection, it was assumed that the items were presented in an understandable format for respondents. It was also assumed that given the items are not of a sensitive manner and are anonymously collected, respondents answered honestly (Willis, 1999). The primary assumption of the Rasch model is that there is a unidimensional latent trait being measured by the tool, although this will be determined by the first research question (Bond & Fox, 2007). However, as a form of item response theory (IRT) model, the Rasch also includes three additional assumptions, the assumption of conditional independence, the functional form assumption, and the item level fit assumption (De Ayala, 2007; Sick, 2010; Toland, 2014). These assumptions are discussed in detail in Chapter 2.

Conclusion

This chapter introduced the educational alternative movement and provided an overview of Montessori high school philosophy. This chapter also introduced the purpose of this study to construct an evaluation tool for usage in Montessori high schools, and provided an overview of the research methods and research questions. The following chapter will provide a thorough review of the literature on measurement and the Montessori Erdkinder.

CHAPTER II

LITERATURE REVIEW

This chapter provides an overview of the literature that serves as the foundation of this work. The chapter begins with a discussion of Montessori's approach to adolescent education and the Erdkinder system. A review of the Erdkinder system's implementation in the United States follows with additional details on holistic education philosophy. Finally, this chapter includes a description of the primary analysis method used for this study, the Rasch model of an item-response theory analysis.

Montessori's Philosophy on Adolescence

Maria Montessori was an early researcher of human development and education principles (Guttek, 2004; Standing, 1998). She believed that humans developed through an ordered series of *planes of education* that started at birth and ended with adulthood (Grazzini, 2004). These planes formed the foundation of Montessori's belief about age appropriate education. Montessori's *planes of education* existed in four stages, with each stage lasting six years. The first plane of education involved the child's early childhood education, and the final plane of education ended with higher education. Montessori believed that the first two planes of education mirrored the last two planes of education, an important consideration when examining how Montessori viewed adolescence.

Adolescents in Montessori's third plane of development are undergoing the same level of physical and psychological stress as infants in the first plane (Grazzini, 2004; Kahn, 2011). These stressors required adolescents to receive special educational considerations in schooling to assure they undergo healthy human development (Barker, 2011; Montessori, 1973; Montessori, 2011a; Montessori, 2011b; Montessori, 2011c).

Montessori saw the third plane as where children physically became adults, and moved socially from being the wards of their parents to individuals free of their parents' influences and preparing to live in society (Montessori, 1973; Montessori, 2011a; Montessori, 2011b; Montessori, 2011c). This challenging movement towards independence guided much of the human development and education principles Montessori laid out for the adolescent age group (Barker, 2011; Kahn, 2011; Tornar, 2011).

Montessori believed that adolescents had two primary concerns, "to be protected during the delicate physical transition period, and to be placed in a position to understand the man's role which he will play in society" (Barker, 2011, p. 97). Montessori believed that personality development was the primary educational goal for adolescents, which would ultimately prepare them for life in society (Montessori, 1973; Montessori, 2011a; Montessori, 2011b; Tornar, 2011). She referred to this period as one in which adolescents were social newborns who were to be guided (Montessori, 2011c). As social newborns, adolescents' success in life would be determined by a belief in their abilities, the capacity to adapt their abilities, and the belief that their abilities could be applied to improving the world (Montessori, 2011c; Tornar, 2011).

In Montessori's philosophy, work was the primary means of personality development, and she held a set of fundamental beliefs about how adolescents should view work (Montessori, 1973; Montessori, 2011a; Tornar, 2011). First, Montessori believed work would provide adolescents the ability to become economically independent of their parents and self-sufficient in their lives. Second, work needed to be viewed by adolescents as a noble endeavor; hence, it would be vital that adolescents

never feel work as being forced or being given without reason (Montessori, 1973). If adolescents began to dislike work, then it could lead to undesirable adolescent personality outcomes, such as avoidance of work. Finally, Montessori believed adolescents needed to understand work existed in both a physical and intellectual capacity, both of which were worthwhile and were complementary to each other. She warned against adolescents desiring one type of work over the other, as work needed to be understood as a multifunction tool for self-help and social adaptability. In the Montessori approach, an appreciation and understanding of work provided the foundation for much of an adolescent's personal development (Montessori, 1973; Montessori, 2011a; Tornar, 2011).

Montessori also believed personality development occurred through adolescents' experiences, particularly through engagement with the community (Kahn, 2011; Montessori, 1973). She encouraged adolescents to identify issues in their community and work to address these issues. The larger goal of these social experiences was to understand the ability of individuals to positively impact society and have faith in the ability of humans to improve social issues. Montessori also desired for adolescents to collaborate with one another, as well as experience other cultures through these interactions with others. Montessori believed that these diverse experiences with others would promote an adolescent's understanding of the importance of cooperation.

Montessori's Erdkinder

Montessori laid out her plans for adolescent education in *From Childhood to Adolescence: Including Erdkinder and the Function of the University*, the work that would largely be used as the source for interpreting Montessori's beliefs for a high school education (Barker, 2011). The principles of Montessori's high school education were to

take place in the Erdkinder, a term which does not have its etymology in education, but rather is a somewhat religious archaic German term for “the children of the soil” (Barker, 2011, p. 97; Montessori, 1973). In the exemplar Erdkinder, adolescents would be housed in a school on a rural farm, not far from a city (Kahn, 2011; Montessori, 1973).

Adolescent students would live at the school as a means to gain independence from their parents (Montessori, 1973). Students would interact with one another and develop together so to foster the students’ individual personality development.

Working on a farm was vital to the adolescent in the Montessori Erdkinder (Kahn, 2011; Montessori, 1973). Adolescents at the Erdkinder would be actively working the land, growing food, caring for livestock, and using machinery (Montessori, 1973). The intent of this farm work, however, was not for students to learn how to become farmers. By working on a farm, students would appreciate the principles of life and death in the natural world, gain a foundation in scientific thought, and connect with the scientific thought process. Through harvesting and selling their crops, students would gain access to the principles of production and sale, which would aid them in later achieving economic independence. Montessori saw farm work as vital for the development of the adolescent’s personality; however, she understood it would be a challenging model to implement.

Montessori also saw the Erdkinder as introducing students to the principles of independence and economic self-sustainability (Kahn, 2011; Montessori, 1973).

Montessori envisioned this as occurring through providing adolescents access to a storefront in a nearby city and a hotel-type business on the farm (Montessori, 1973). In the store, the students would work and sell the produce they grew on the farm, as well as

any other creative materials. In the hotel, students would further the principles of hospitality they learned from earlier Montessori experiences. Through both experiences, students would learn to interact with their parents in a new manner that would aid in achieving independence by altering the perception of adult parents as only having a parental role. Montessori saw these experiences as encouraging independence and economic self-sustainability, which were the two ultimate goals of adulthood.

History

The desire for a Montessori plan of education for high school-aged students began to grow following the publication of her work on the second plane of education in 1916 (Barker, 2011). Montessori, however, had not laid out plans for high school students at this point and was focused still on the development of the earlier childhood stages.

Attempts at the formation of a Montessori high school can be traced back to as early as 1923, however, the first high school bearing the Montessori name did not open until 1930 in Amsterdam. By the time Montessori did begin to discuss her principles on adolescent development, there were several high schools in Europe bearing the Montessori name. Although Montessori consulted with some of these high schools, she never approved the actions of these schools to the same extent as with schools focused on pre-school and elementary aged children.

Montessori began discussing her principles of adolescent education in a series of widely attended lectures in 1936 (Barker, 2011). A few years later, “The Erdkinder” essay was published. In 1948, the first version of *From Childhood to Adolescence: Including Erdkinder and the Function of the University* was published. The work included Montessori’s Erdkinder essay and notes from her earlier lectures. *From*

Childhood to Adolescence: Including Erdkinder and the Function of the University would not make it to the United States until 1973. Shortly after, Montessori high schools would start being developed in the United States.

The interest in Montessori education in the 1970s is attributed to how middle class Americans perceived the possible positive impact on children's social and emotional growth of sending their children to these schools (R. Miller, 1990). Montessori did not stress democratic principles or social upheaval through education, but rather encouraged a strong focused work ethic as a means of gaining economic success. Montessori also saw the present school system as the flaw. R. Miller (1990) argues that these factors led middle class Americans to strongly support these schools, as they provided a means to encourage social change in a peaceful manner that worked well with American ideas of culture. R. Miller (1990) warns that it is easy for individuals to misinterpret the intent of Montessori schools, with some believing they are rigid institutions creating peaceful conformist students, while others view these schools as promoting disorder.

Montessori advocates were aware that the premise of an Erdkinder experience would be challenging for both students and parents to accept, and for school administrators to implement fully (Kahn, 2011). This awareness led Montessori advocates to label Montessori high schools as experiments that would eventually lead to the full creation of an Erdkinder. An estimated 550 schools have been created since the 1970s, each ideally leading towards the progression of a true Erdkinder school. Presently, there are about 12 Erdkinder schools actually located on farms, one of the most

prominent of these schools is The Farm School at the Hershey Montessori School in Pennsylvania.

Montessori high schools have had mixed success in the United States. However, two U.S. schools are worth noting for their influence on the U.S. Montessori high school community: The Farm School at the Hershey Montessori School and the Montessori High School at University Circle (Ludwig, 2011; Kahn, 2011). The Farm School has become a center for individuals attempting to understand the Montessori Erdkinder philosophy or implement a farm-like setting into their school practices. Because the school is located on a farm, it is considered to be a close representation of what Montessori had intended for adolescent education (Kahn, 2011). The Montessori High School at University Circle is considered a model for the implementation of a Montessori high school in an urban setting, specifically regarding urban education practices. Each of these schools is highly influential in the Montessori high school community. These two schools provide the basis for much of the thinking about the appropriate way to implement an Erdkinder school and provide many of the widely used resources actually implemented in Montessori high schools. Neither of these schools has published a school evaluation tool that has been rigorously examined for reliability and validity.

Curriculum

Montessori laid out the general principles for an Erdkinder educational curriculum that she believed would lead to the desired developmental outcomes, particularly the development of the adolescent personality (Montessori, 1973). Montessori (1973) divided the principles of her curriculum into three categories:

1. To open the way to the possibilities of the adolescent for personal expression. That is, to facilitate, by exercises and exterior means, the development of the interior personality.
2. To supply that which we consider to be the creative elements necessary for the physical being of man in general.
3. To put the adolescent into relation with present civilization by bringing him general culture and by experience. (p. 116)

Each of these three principles would be achieved through the school curriculum (Montessori, 1973). The first goal of personal expression would be gained through artistic tasks such as music, poetry, drama, and art. Personal expression was to be understood as an endeavor connected to work. The second goal of creative elements education should be composed of moral education, mathematics, foreign languages, and linguistics. One of the primary points of creative elements education is to teach students that abstract concepts can be placed into physical forms that can be manipulated. The final goal was to put adolescents in connection with civilization through the natural sciences, engineering, history, and elective learning. Montessori had a particular desire for adolescents to work with machines and to understand that machines should be used to improve humanity. Elective learning would provide an opportunity for learning in the specific fields of students' interest. Montessori laid out these curricular principles, but never developed materials similar to what she developed at the elementary level (Barker, 2011).

In addition to those three curricular principles, Montessori had specific beliefs about the nature of student work in the Erdkinder (Kahn, 2011; Montessori, 1973). Montessori desired for the Erdkinder curriculum to be based largely on the principle of

choice and expression. Montessori wanted the adolescent curriculum to encourage work, but teach that specialization was only a means of entering into the workforce and that individuality should not be lost due to specialization (Montessori, 1973). She did not believe that adolescents should be forced to complete school work, but also that they should not be allowed to waste their possible work potential.

Importantly, Montessori understood some of the practicalities of a high school education. Montessori believed that the Erdkinder should dedicate time in the final two years of high school to prepare for university entrance examinations (Montessori, 1973). Although Montessori was referring to preparations for academic success, she suggested a possible examination of character as a way of understanding if the adolescent is prepared to enter society. Examinations did not fit with the general principles of Montessori, but she acknowledged their value in the larger culture.

Interpretations

In comparison to the early-education and elementary plans that Montessori laid out, little was left behind on high school (Barker, 2011). Given that little exists beyond the seminal text and a few recently released lectures, Montessori scholars have had to conjecture heavily about how to interpret the few Montessori texts available and the ideas put forth in other works. For example, scholars have used Montessori's beliefs about peace through education to develop curricula with peace as an end-goal (Kahn, 2011; Kahn & Pendleton, 2007).

The group largely responsible for the advancement of high school Montessori methods is the North American Montessori Teachers of Association (NAMTA) in conjunction with the works of its leader David Kahn (Ludwig, 2011). Many of the

Erdkinder materials and methods of teaching have come from NAMTA. Kahn was instrumental in the creation of one of the closest examples of a true Erdkinder in the United States, the Hershey Montessori Farm School, and is considered an expert in the field of Montessori high school education (Kahn, 2011; Ludwig, 2011). Kahn and NAMTA were largely influential in the creation of an introductory program to Montessori high school education and the creation of an annual colloquium for interest on the topic.

In his work, “Eight Pictures at an Exhibition: A Montessori Retrospective on the Discovery of the Adolescent,” Kahn broke down Montessori high school outcomes into four holistic education facets; moral, emotional, cognitive, and social (Kahn, 2011). Each of these holistic education facets included a series of student outcomes based on his review of the Montessori literature. Kahn intended for these outcomes to serve as a foundation of formal research on Montessori high school education. Kahn (2011) intended for his outcomes to go beyond the typical education outcomes and examine the “social goals’ aimed at understanding the whole personality” (p. 25). Through these goals, Kahn provides a format for examining Montessori high school outcomes beyond the standard educational outcomes.

Holistic Education

The Montessori education is considered to be a holistic approach to education (R. Miller, 1990; J. P. Miller, 2010; Taggart, 2001). Holistic education, also known as whole child education, is committed to creating within students a sense of oneness with the universe and a rejection of overly material forms of education (Mayes & Williams, 2013; R. Miller, 1990; J. P. Miller, 2010). Miller (1990) explains that the movement encourages

“inner human qualities, such as mind, emotion, creativity, imagination, compassion, a sense of wonder and reverence, and the urge for self-realization” (p. 58). The holistic education approach rejects education objectives that are considered to promote materialism, instead promoting educational expression and freedom in learning (R. Miller, 1990; J. P. Miller, 2010).

Holistic education is guided by a series of three principles, each with the intent of challenging the standard educational school practices (R. Miller, 1990). First, holistic education seeks the balancing of holistic goals with the materialistic goals, for example, imaginative work with reasoning. Second, students need to feel a sense of inclusion with others and within their classroom. Finally, students must begin to examine the relationships that exist in the world, and should not see the universe as parceled out but rather as a cohesive entity. Overall, these three principles guide much of the work in the holistic classroom.

Holistic education does not exist as a singular theory of education, but rather refers to the many educational theorists and researchers whose ideas embody what are considered holistic education principles (Martin, 2002; R. Miller, 1990). These researchers include Thomas Maslow, whose hierarchy of needs is credited as providing the foundation for many holistic education principles (Martin, 2002; R. Miller, 1990; Taggart, 2001). In Maslow’s hierarchy of needs, the principles of what a human needs are laid out in the order of physiological needs, safety, belongingness and love, esteem, and self-actualization (Maslow, 1943; Maslow, 1987). Physiological needs refers to what is required to keep a human body physically functioning; safety refers to a feeling of security and sense of stability; belongingness and love refers to feelings of affection, and

sense of being part of a community; esteem refers to feelings of self-worth and confidence; and self-actualization is a point at which an individual is reaching a proposed state which is truly reflective of their nature and performing the work which fits this nature. These needs described by Maslow would guide the development of many holistic education models (Martin, 2002; R. Miller, 1990; Taggart, 2001).

Both Maslow and Montessori are considered to be hallmarks of the holistic education movement (Martin, 2002; R. Miller, 1990; J. P. Miller, 2010; Taggart, 2001). In a 2011 article, Weinberg discussed the connection between Maslow's theory and Montessori principles of education. Weinberg argues that each stage of Maslow's hierarchy of needs is reflective of Montessori education principles, and these principles ultimately lead to the self-actualization of a child. The relationship between Maslow and Montessori as hallmarks of the holistic education movement, as well as the connection between the two philosophies described by Weinberg (2011), provides support for the interpretation of the Montessori educational approach through the usage of Maslow's hierarchy of needs.

Item Response Theory

IRT models are a common method of examining tool quality, as they allow for tool developers to estimate the latent trait characteristics of both a tool's items and its respondents (De Ayala, 2009; Goldstein & Wood, 1989; Toland, 2014). Specifically, IRT allows for the estimation of how much of a latent trait is within a person (person-ability level) and how much of a latent trait is required to endorse (item difficulty level) an item. The item level analysis is what makes the IRT model distinctive, as it allows for researchers to identify issues with specific items instead of only on the whole-tool level.

One-Parameter Logistic Model

The one-parameter logistic (1PL) model is IRT's most basic form (De Ayala, 2009). The parameter of interest in the 1PL IRT model is the item difficulty parameter (Brown, 2006; De Ayala, 2009). The one-parameter IRT model estimates an item's difficulty by identifying the point at which a person with the same estimated level of a latent trait has a 50/50 probability of either positively endorsing the question (or answering the question correctly in education assessments). The item difficulty levels are reported on a logit scale, usually ranging from -3.0 to 3.0 (De Ayala, 2009; Toland 2014). Items below the 0.0 point of the logit scale are considered easier to endorse than items above the 0.0 point. Items at the extreme ends of the logit scale (e.g. -3.0 and 3.0) are considered to be the easiest and most difficult to endorse, respectively. Theoretically, these logit values can go beyond the -3.0 and 3.0, however, they would only be reported if items fell at these levels.

In addition to the item difficulty parameter, the one-parameter IRT model provides a person-ability level for each person (Brown, 2006; De Ayala, 2009; Toland, 2014). A person-ability level is determined through a series of steps predicated on the probability of a person endorsing an item. Similar to item difficulty levels, person-ability level is reported using a logit scale with the same reporting premises. The estimates developed by the IRT model are considered invariant across a given population, and thus, person ability estimates can be developed for any individual considered to be a part of the population. It is important to note that IRT as a model is most effective at estimating person-ability levels when they are close to item difficulty levels (De Ayala, 2009; Toland, 2014). To estimate a range of abilities it is important to have item difficulties that

range throughout the logit scale which can determine the person-ability levels of a range of people.

The mathematical formula for the IRT unidimensional 1PL model is presented as follows: (Rasch 1960/1980; Brown, 2006, p. 397)

$$P(y_{is} = 1 | \theta_s, b_i) = \frac{\exp(\theta_s - b_i)}{[1 + \exp(\theta_s - b_i)]}$$

As the 1PL model equation shows, the left side of the formula is the logistic function of the model estimating if the item is endorsed “1.” The logistic function of the formula can be directly interpreted as “the probability (P) that y equals 1 for a specific item (i) and participant (s), given (|) the participant’s (ability) level (θ_i) and the item’s difficulty (b_i)” (Brown, 2006, p. 397). The *exp* (*exponential*) on the right side of the equation is approximately 2.718 and can be used to calculate specific person’s probabilities (De Ayala, 2009). The actual θ and b used in the 1PL model are commonly calculated through a process called joint maximum likelihood estimation (JLME; De Ayala, 2009).

JLME simultaneously develops person-ability levels and item difficulty levels. JLME first estimates item difficulty location through persons’ responses and then uses these estimates to determine person-ability levels (De Ayala, 2009). The person-ability levels are then used to determine new item difficulty levels. The new item difficulty-levels are then used to determine new person-ability levels. This process is repeated until the estimates remain generally consistent. JLME is commonly used in Rasch model programs such as Winsteps, however, it is not the only means of developing estimates for the 1PL model.

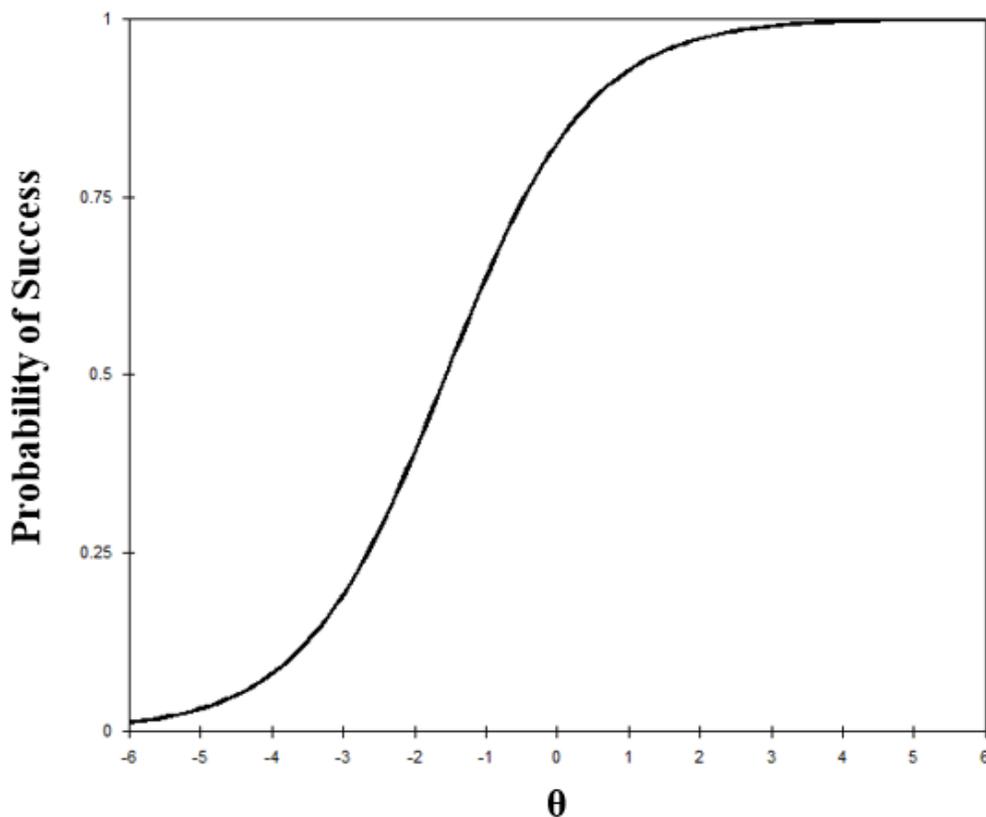
In addition to the basic model, the 1PL model can be written as having a multiplier α , as shown below: (Birnbaum, 1968; Brown, 2006, p. 398)

$$P(y_{is} = 1 | \theta_s, b_i, \alpha_i) = \frac{\exp[\alpha_i(\theta - b_i)]}{1 + \exp[\alpha_i(\theta - b_i)]}$$

The multiplier α is relevant when interpreting item characteristic functions and the information they provide (Brown, 2006; De Ayala, 2009). Figure 1 is an item characteristic curve that demonstrates the probability of a person endorsing an item at a given difficulty level based on a person's ability level (De Ayala, 2009).

Figure 1

Item Characteristic Curve



The multiplier α of the exponent impacts the slope of the item characteristic curve (De Ayala, 2009). In the 1PL model, the item characteristic curves have a constant slope,

better known as a discrimination parameter. The constant slope means that the information provided by each item does not alter. When slopes are steeper (through the increase of α) they are able to reduce uncertainty about the ability levels of persons and when slopes are wider (through the decrease of α) they increase uncertainty. The 1PL model commonly has a constant α of 1, which is also known as the Rasch model for dichotomous item response data.

Rasch Model

Similar to other IRT models, the Rasch model is an accepted means of examining assessment and evaluation tools (Bond & Fox, 2007). The primary assumption of the Rasch model is that a single unidimensional latent concept is being examined, although multidimensional Rasch models do exist and can be examined using alternative Rasch modeling methods (Bond & Fox, 2007; Linacre, 2015a). When examining Rasch model outputs, in addition to item difficulty and person ability scores, it is important to examine item fit statistics. Item fit allows for researchers to interpret if an item is appropriate for a tool. Fit statistics are reported for each item and person on a tool, specifically the infit and outfit statistics reported as mean squares and standardized t -values. Infit means squares can be determined using the formulas: (Engelhardt, 2013, p.178)

$$\text{Person Mean Squares Infit} = U_i = \frac{\sum_n^N z_{ni}^2}{N}$$

$$\text{Item Mean Squares Infit} = V_i = \frac{\sum_n^N y_{ni}^2}{\sum_n^N Q_{ni}}$$

Outfit mean squares can be determined using the formulas: (Engelhardt, 2013, p. 178)

$$\text{Person Mean Squares Outfit} = U_n = \frac{\sum_i^I z_{ni}^2}{L}$$

$$\text{Item Mean Squares Outfit} = U_i = \frac{\sum_n^N z_{ni}^2}{N}$$

These mean square formulas require the use of the score residuals (Y_{ni}) and standardized score residuals (Z_{ni} ; Engelhardt, 2013). Mean square formulas report the size of a misfit, however, do not indicate if this misfit is statistically significant (Engelhardt, 2013; Linacre, 2012a). Identifying if the misfit is statistically significant requires standardizing the mean squares (Linacre, 2012a). Mean squares are standardized to z-scores (ZSTD) using the Wilson-Hilferty transformation: (Linacre, 2012a, p. 27)

$$q^2 = \frac{2}{d.f.}$$

$$ZSTD = \left(MnSq^{\frac{1}{3}} - \right) \left(\frac{3}{q} \right) + \left(\frac{q}{3} \right).$$

The Wilson-Hilferty transformation requires the usage of the degrees of freedom from the selected mean squares ($MnSq$) statistic (Linacre, 2012a). Although these standardized fit statistics are calculated as ZSTD scores, they are considered to be t-values and reported as such (Bond & Fox, 2007; Wright & Masters, 1990). They are considered t-statistics because in the Rasch model, the ZSTD scores “approximate a unit normal distribution corresponding to a t-statistic with infinite degrees of freedom” (Wright & Masters, 1990). A t-statistic above 1.96 is considered to be statistically significant (Linacre, 2015a), and demonstrates misfit, as it rejects the null hypothesis “these data fit the Rasch model” (Wright & Masters, 1990, p. 84). The 1.96 expectation is rounded to 2.0 in programs such as Winsteps, and thus items with fit below -2.0 and above 2.0 are considered to have an issue with misfit (Bond & Fox, 2007; Linacre, 2015a). Misfit is likely to occur when the responses to an item are erratic in comparison to what the model expects the responses to be (Bond & Fox, 2007).

In addition to the fit statistics, the Rasch model also provides reliability estimates (Bond & Fox, 2007; Linacre, 2012b; Linacre, 2015a). The Rasch model reports two types

of reliability, person level and item level. Person level reliability refers to the expectation that a different group of people with similar ability levels would respond to items in a manner similar to the sample used in the analysis (Bond & Fox, 2007). Item-reliability refers to if the sample used had enough participants to allow for acceptable statistics to be generated in the report. The formula for identifying these reliability estimates is as follows: (Linacre, 2012b, p. 26)

$$Reliability = \frac{True\ Variance}{(True\ Variance + Error\ Variance)} = \frac{True\ S.D.^2}{(True\ S.D.^2 + S.E.^2)}$$

True variance can be identified with the square of the true standard deviation (*True S.D.*) reported for person-ability and item difficulty estimates, respectively (Linacre, 2012b; Linacre, 2015a). The error variance can be found by identifying the mean of the standard errors (*S.E.*) for these estimates. Linacre (2015a) suggests a minimum of .80 for both person and item reliability estimates. When there is an issue identified with reliability, Linacre (2012b; 2015b) suggests adding additional items to address person reliability issues, and adding additional persons to the sample to address item reliability issues. Person and item reliability estimates allow for greater confidence in interpreting the results generated by a Rasch model analysis (Bond & Fox, 2007).

Differential Item Functioning

Within the Rasch model, variations in item difficulties between groups can be determined through examining differential item functioning (DIF; Bond & Fox, 2007). DIF analyses are valuable in examining measurement tools, as variations in how groups respond to items within a tool may suggest there are issues with the tool. DIF is determined by estimating item difficulty levels for two groups and comparing the difference between these groups. Identifying DIF requires identifying a Welch's t-

statistic (Welch & Miller, 1995) for each item on the tool using the following formula:
(Linacre, 2015a, p. 429)

$$t = \frac{DIF\ Contrast}{Joint\ S.E.} = \frac{(Group_1\ Difficulty - Group_2\ Difficulty)}{\sqrt{(Group_1\ S.E.^2 + Group_2\ S.E.^2)}}$$

The formula requires identifying the difficulty of the item (*Difficulty*) for each group and the respective standard error of the measure (*S.E.*; Linacre, 2015a). The corresponding *p*-value for the t-statistic is then found using the specified degrees of freedom. For an item to be considered as having an issue with DIF there must be a statistically significant difference between the ability of individuals of different groups to respond to the item as demonstrated by a t-statistic above 1.96 (Linacre, 2015a; Tristan, 2006). In interpreting DIF, it is important to consider if the difference is a result of item bias or an issue beyond the test developer’s control (Bond & Fox, 2007; Linacre, 2015a). For example, DIF may be a result of one group not having the sufficient experience or background knowledge related to an item’s content.

Rating Scale Model

The standard Rasch model uses dichotomous item response data. However, the model can be extended to polytomous item response data by means of the Rasch Rating Scale Model (RSM; Wright & Masters, 1982). The Rasch RSM can be written as follows: (Ostini & Nering, 2005, p. 39)

$$P_{ig} = \frac{e^{\sum_{g=0}^l [\theta - (b_i + \tau_g)]}}{\sum_{h=0}^m e^{\sum_{g=0}^h [\theta - (b_i + \tau_g)]}}$$

The Rasch RSM equation identifies the probability of endorsing a particular item’s category (*g*; Ostini & Nering, 2005). In the equation, *m* is the number of categories and “*h*=0, 1, ..., *g*, ..., *m*...with *g* representing the specific category being modeled from *m*+1

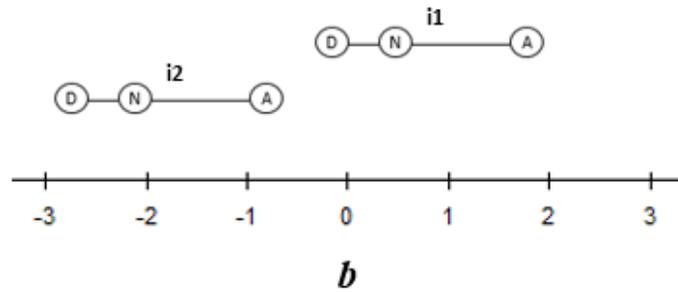
categories” (Ostini & Nering, 2005, p. 39). The Rasch RSM also estimates a threshold parameter (τ) that is unique for the analysis of polytomous data.

The threshold estimate is predicated on the assumption that moving between response categories on a scale is not equal (Bond & Fox, 2007). For example, choosing to endorse “Disagree” over “Strongly Disagree” on an item may be more challenging for a person than choosing to endorse “Agree” over “Disagree.” Thus, a threshold parameter is developed to indicate this difference in the ability of a person to endorse between item categories or specifically, “the level at which the likelihood of being observed in a given response category...is exceeded by the likelihood of being observed in the next higher category” (Bond & Fox, 2007, p. 105). In the Rasch RSM, a single set of thresholds is estimated, the Rasch-Andrich thresholds. The Rasch-Andrich thresholds are developed with the assumption that each item has a similar rating scale structure, and thus, only a single set of thresholds needs to be generated that can be used to demonstrate the thresholds for each items categories (Bond & Fox, 2007).

The thresholds are important to how item difficulty estimates are interpreted when data are being analyzed with the Rasch RSM. Under the Rasch RSM, item difficulties represent the point where responding to categories above the point is as likely as responding to categories below (Bond & Fox, 2007). The item difficulty estimate does not necessarily correspond to a specific category on an item’s scale and may fall between two possible responses. Furthermore, the estimated thresholds are important when comparing the probability of a person endorsing a category on an item. As Figure 2 demonstrates, with Rasch RSM difficulties and thresholds, it may be easier for a person to endorse “Disagree” (D) on one item ($i1$) than “Agree” (A) on another item ($i2$).

Figure 2

Rasch Rating Scale Model Difficulties and Thresholds



Thresholds in the Rasch RSM model alter aspects of interpreting the analysis results, particularly when considering item difficulty and person-ability.

Conclusion

This chapter included a discussion of the relevant contextual and methodological information necessary for this study. The following chapter details the complete methodology and the analyses conducted.

CHAPTER III

METHODOLOGY

This chapter begins by recapping the purpose of the study and the research questions. This is followed by a discussion of the survey construction and data collection methods. The chapter then describes the analysis procedure and study sample. Finally, there is an explanation of how each research question will be answered using the study results.

Purpose and Significance

The purpose of this study was to construct and validate a holistic education school evaluation tool for usage in Montessori high schools. The tool was designed to examine the presence of Montessori's holistic education Erdkinder principles in a school as determined by perceptions of student outcomes. Enrollment in Montessori high schools continues to grow as parents and students seek out alternative education schools to replace public schooling, which they view as overly focused on standardized testing and traditional concepts of learning (Finn et al., 2006; Kahn, 2011; Lipman, 2011). However, without a common school evaluation tool, Montessori high schools are unable to identify areas of school improvement and market their current successes on the vital Erdkinder principles. Presently, the Montessori academic community has not presented a common school evaluation tool that has been examined for issues of validity and reliability.

The challenge with examining the presence of Montessori's Erdkinder principles is that Montessori herself left little behind on the topic. The few texts left by Montessori have been widely interpreted by the Montessori community, sometimes in competing capacities. Furthermore, Montessori warned against tracking, which has led some in the

Montessori high school community to view data-based evaluation as inappropriate. To develop a tool examining Montessori high schools, it is appropriate to not only examine Montessori's works, but also overall holistic education research and the works of academics using the Montessori adolescent philosophy.

Research Questions

The research questions guiding this study were designed to determine the quality of the tool being tested. The research questions are:

1. How well does the evaluation tool measure the latent trait holistic education concept of the Montessori Erdkinder school principles?
2. How well do the individual items fit the Montessori Erdkinder school principles latent trait reflected on the evaluation tool?
3. To what degree do the item difficulty levels from items of specific holistic education facets compare to the theoretical item difficulty levels of another holistic model, specifically, Maslow's hierarchy of needs?
4. Are there differences in how people from different groups, with similar levels of the latent trait, are able to respond to items?

These research questions were used to determine if the holistic education school evaluation tool developed in this study needed revisions and the specifics of how this tool could be revised.

Construction of the Montessori Erdkinder School Evaluation Survey

The survey constructed for this study is called the Montessori Erdkinder School Evaluation Survey (MESES), and measures the presence of Montessori's holistic education Erdkinder principles for appropriate adolescent development. The MESES

includes a series of 23 items separated into four facets: social, cognitive, emotional, and moral. These four facets reflect the holistic education facets outlined by Kahn (2011). Survey items consisted of statements that asked people to respond how strongly they agree or disagree with the statement. Each item had five possible responses: *strongly disagree*, *disagree*, *agree*, and *strongly agree*, as well as the option, *do not know*. The statement “Students at the Montessori high school...” was placed above the series of items, to be a stem for each item. This stem system had respondents focus their answers towards the school with which they are currently affiliated. The complete MESES can be found in Appendix A.

Item creation for the MESES occurred following an extensive process demonstrated through the item matrix found in Appendix B. The item creation process began with the identification of the student learning goals and their respective facets specified in Kahn’s “Eight Pictures at an Exhibition: A Montessori Retrospective on the Discovery of the Adolescent Work” (2011). These goals were then directly connected to the texts from Montessori herself and if this could not be done, additional support was sought from other Montessori Erdkinder scholars. When a goal could not be connected to additional support from the additional sources, an item was not developed to represent that goal. The item creation process eventually led to the development of 23 items. The 23 items with their facet groupings and subsequent labels are in Table 1, and they provide the basis for the evaluation of the Erdkinder practices within a school.

Table 1

Facets and Items

<i>Facet</i>	<i>Items</i>
Cognitive	<ul style="list-style-type: none"> - Are taught how to find structure in nature. (c1) - Are taught how math applies to their life. (c2) - Are taught about other cultures through foreign language instruction. (c3) - Learn how society has changed over time. (c4) - Are encouraged to philosophically reflect on the universe. (c5) - Express themselves through creative arts. (c6)
Social	<ul style="list-style-type: none"> - Are taught to challenge the rules of society. (s1) - Feel they should help others in society. (s2) - Learn that working is important to their future success. (s3) - Volunteer in the community. (s4) - Are encouraged to work with each other. (s5)
Emotional	<ul style="list-style-type: none"> - Are supported in building self-confidence. (e1) - Are taught problems in society are fixable. (e2) - Find enjoyment in working. (e3) - Feel their future has already been decided. (e4) - Are aware of their own talents. (e5) - Learn responsibility. (e6)
Moral	<ul style="list-style-type: none"> - Get to choose their schoolwork. (m1) - Respect one another. (m2) - Set goals to improve society. (m3) - Are taught to address the problems in society. (m4) - Feel that working is shameful. (m5) - Discuss their ideas. (m6)

Note. Item label is in parentheses.

In addition to the 23 Erdkinder items, three open-ended response questions were included on the MESES. These open-ended response questions were included to provide additional information for the formative evaluation purposes of the school and allow schools to gather additional information that may aid in the development of individual school evaluation surveys. The open-ended response items include questions asking respondents to identify strengths and weaknesses of the school, and a question about other questions to possibly include on the survey in the future. Information from these open-ended response questions may be particularly fruitful for individual schools who plan to use the MESES over several iterations.

In addition to the open-ended response questions, basic demographic items were included. Respondents were asked to identify their sex and what their relationship to the school was: student, parent/guardian, teacher, or staff. An additional category of *recent graduate* was included for the July 2015 collection phase. These demographic questions did not include additional items to aid in protecting respondent anonymity.

Considerations

There are a wide array of means to develop a survey and write survey items. Certain choices were made for the development of the MESES to aid in increasing the quality of the measurement. For example, although the items were placed in an on-line format, the item order was not randomized. Items were not randomized because schools implementing a paper-and-pen format will not have the ability to randomize items, and the reliability and validity results should relate to individuals using both an on-line and paper-and-pencil format. In addition, three of the items were negatively written and would be reverse-coded in the analysis later conducted. By having negatively phrased

items that require reverse-coding, respondents are prevented from simply answering items without regard for what the item is saying (Nardi, 2006). The survey was also kept to a short form, thus increasing the likelihood that the survey would be completed by respondents, as it requires less time for completion. Furthermore, the short-form reduces the likelihood of the respondent becoming disinterested or fatigued with the survey and not answering as they normally would. Finally, items were written to measure a single concept, and thus, items did not include the usage of conjunctions or clauses that may lead to the measurement of multiple concepts. These survey development considerations aid in the creation of high-quality measurement tools and were implemented in this study.

Data Collection

A Montessori high school in the southeastern United States was selected as the data collection site. This site was selected because the MESES was developed for this specific type of educational setting. Students, parents/guardians, teachers, and staff from the school were selected to be the study sample. These groups were selected for inclusion in the sample because they provide a range of perspectives on the school.

Approval for the distribution of the survey and collection of data was first received from the school's governing board and Director. This approval included the requirement that the survey data collected be provided to the school for their own formative evaluation purposes. The study was then approved by the University of Kentucky's Institutional Review Board, and a waiver of consent was obtained that allowed for data to be collected from students without the need for parental or guardian documented consent. This consent waiver required that the Director of the school contact parents and guardians in advance of the survey's distribution, notifying them that them

about the study, and that the survey was going to be distributed to students. Parents and guardians were notified they could request their student not participate in the study by contacting the Director or primary investigator.

The proposed survey was then created in an on-line format, specifically through SurveyMonkey. The survey was designed so that item ordering would remain consistent and each respondent would receive the same version of the survey. The on-line format was selected following consultation with the school Director, who expressed that school evaluation data were standardly collected through the usage of on-line tools. The on-line format was also selected to capture responses during summer months when accessing students may be challenging.

The survey data was collected through two phases, in July 2015 and October 2015. Two collection phases were required in order to conduct the specified analysis; specifically, results of the Rasch Rating Scale Model (RSM) analysis from the July 2015 collection period would be needed for anchoring when conducting the Rasch RSM analysis with the October 2015 data. At the beginning of each collection phase, the school's Director e-mailed possible respondents that they would receive the survey in three days. This preliminary e-mail specified that the data collected would be used for research purposes and by the participating school for evaluation purposes, and respondents had two weeks from receiving the initial survey to respond. During each phase, the Director sent out the actual survey three days after sending the preliminary e-mail. For the October 2015 collection, possible respondents were sent a follow up e-mail asking them to complete the survey if they had not done so previously. The October 2015 follow-up email was sent a week after the possible respondents initially received the

survey, and again two days before the close of the data collection window. Follow up e-mails were intended to be sent during the July 2015 collection phase as well. However, issues unrelated to the study arose that kept the school's Director from sending the follow-up e-mails in this phase.

Sample

The sample selected for this study was students, teachers, parents/guardians, and staff members associated with a Montessori high school located in the southeastern United States. The sampling procedure included two waves of data collection, the first in July 2015 and the second in October 2015. The July 2015 collection wave garnered 18 respondents and the October 2015 sample garnered 36 respondents. For a Rasch RSM analysis, 50 respondents are suggested when using polytomous data (Linacre, 1994). The July 2015 data does not meet the minimum suggested number of respondents to conduct a Rasch RSM. However, the first sample is only being used for advanced calibration purposes through the anchoring process described. Thus, the small sample was not a concern. Similarly, the October 2015 sample also did not meet the minimum suggested number of respondents to conduct a Rasch RSM. Although neither sample met the suggested sample size requirements for conducting a Rasch RSM, there are several factors that suggest the analysis is still appropriate. First, item reliability estimates provided as part of the Rasch RSM demonstrate if sufficient data is available to provide interpretable results. As such, item reliability estimates provide a means of interpreting if a sufficient sample size has been collected. In addition, although the suggested number of participants for a Rasch RSM analysis is 50 respondents, the Rasch is capable of generating usable statistics with smaller than recommended samples (Bond & Fox, 2007;

Linacre, 1994). For example, Wright and Stone's (1979) "Best Test Design" included only 35 respondents (Linacre, 1994). Finally, the item-anchoring procedure used in the analysis allowed for the development of better estimates for the final Rasch RSM analysis, partially making up for the loss in estimate precision that comes from having small samples.

Analysis

The Rasch RSM was selected to be the primary means of analysis for this study. The Rasch RSM analysis technique was selected because it is assumed that respondents interpreted the items' response categories in a similar manner (Bond & Fox, 2007; Linacre, 2000). The Rasch RSM results would be examined for unidimensionality, reliability, item fit, item difficulty spread, item difficulty hierarchy, and differential item functioning (DIF), in a process that built upon the tool validation process discussed by Royal and Elahi (2011). To aid in the production of more accurate results to answer this study's research questions, a series of Rasch RSM analyses was conducted using an item-anchoring strategy. An item-anchoring strategy provides better estimates of item difficulty levels and person-ability levels, as anchoring uses estimates from a prior analysis in the development of estimates for the present analysis (Bond & Fox, 2007; Linacre, 2015a). The analysis procedure for this study included conducting three Rasch RSM analyses, with each analysis being run in Ministep version 3.90.1 (Linacre, 2015b).

The first Rasch RSM analysis was conducted using the data collected from the July 2015 sample. The items that demonstrated appropriate item fit in the first analysis were then marked for inclusion as anchor item difficulty estimates in the second Rasch RSM analysis.

The second Rasch RSM analysis was conducted using data from the October 2015 sample and the previously designated item difficulty estimate anchors from the first Rasch RSM analysis. Item difficulty anchoring involves items that demonstrated appropriate fit in the initial analysis having their item difficulty estimates directly used as the item difficulty estimates, for their respective corresponding items, in the second analysis (Bond & Fox, 2007; Linacre 2015b). This anchoring strategy means that item difficulty estimates for the non-anchored items and person ability estimates would be developed using the specific item difficulty estimates from the anchored items. The results of this second analysis will then be examined for item displacement, also known as drift, which is the difference in the anchored item difficulty with what these items difficulties would have been had they not been anchored. This analysis used the recommendation given by O'Neill, Peabody, Tan, and Du (2013), that items with displacement at or above 0.6 not be anchored, and thus these items demonstrating displacement were not anchored in the third analysis.

The third Rasch RSM analysis used the data from the October 2015 sample and included the anchored items that were not marked as having an issue with displacement or misfit during the second analysis. The third Rasch RSM analysis was then used to answer the research questions for this study.

Answering Research Questions

The methodology was selected to support the validation purpose of the study and to answer the specified research questions. Only the results of the third, final Rasch RSM analysis was used to answer the study's research questions.

To answer the first research question, *How well does the evaluation tool measure the latent trait holistic education concept of the Montessori Erdkinder school principles?*, the results of the principal components analysis (PCA) of the Rasch residuals, which were provided as part of the output of the Rasch RSM, were examined. The PCA of the Rasch residuals is a factor extraction procedure that attempts to determine if an additional factor can be identified from the data residuals or if the residuals represents additional statistical randomness that is not accounted for in the model (Linacre, 1998). The PCA of the Rasch residuals was used to determine if a single unidimensional concept was being measured in the survey. A single concept being measured would support that the survey was measuring the proposed Montessori Erdkinder concept although additional validation studies will be needed to confirm the measurement. Multiple concepts being measured, however, would suggest that the MESES is measuring several factors and may need to undergo an item-reduction strategy that would reduce the survey to a single unidimensional concept (DeVellis, 2012). If the PCA of the Rasch residuals yields a first contrast with an eigenvalue above 2.0, then the grouping of the specific positive and negative items would be examined in the first contrast loading table (Linacre, 2015a). If the positive and negative items of the respective groupings were determined to be related, such as being associated with a specific facet yet located on opposite ends, it would be assumed that multiple dimensions are making up the tool. Although, this clustering of items within a specific facet may not be of concern because it is reasonable that these facets make up the latent trait being examined. However, should examining the first contrast loadings demonstrate a clustering of items that appear unrelated to the latent

trait, such as items of various facets containing a similar theme, the tool would be considered as having multiple dimensions that needs to be examined in future analyses.

To answer the second research question, *How well do the individual items fit the Montessori Erdkinder school principles latent trait reflected on the evaluation tool?*, item fit and item difficulty range were examined. Item fit was determined by looking at item infit statistics and item outfit statistics. Items demonstrating appropriate fit were considered to be measuring Montessori Erdkinder principles well, and items demonstrating misfit were considered to not be measuring Montessori Erdkinder principles well. Items that had infit statistics or outfit t-statistics that fell outside the -2.0 to 2.0 range were considered to be misfit items. Any misfit items were tagged for either removal or revision in future iterations of the MESES. Item difficulty range was then examined to determine if the survey included a range of easy-to-difficult items. Having a range of item difficulties would suggest that the survey items are capable of capturing respondents who are likely to endorse items at varying ranges (Bond & Fox, 2007). A clustering of items around a difficulty level, or lack of items at a certain difficulty, would suggest that the tool could not capture the ability levels of all persons.

To answer the third research question, *To what degree do the item difficulty levels from items of specific holistic education facets compare to the theoretical item difficulty levels of another holistic model, specifically, Maslow's hierarchy of needs?*, the order of item difficulty levels with their respective facet label was compared to the holistic model proposed, specifically, Maslow's hierarchy of needs. To conduct this comparison, the order of the item difficulties was compared to Maslow's hierarchy of needs difficulty order. In Maslow's model, the easiest to most difficult items would fall in the order of

cognitive, social, emotional, and then moral items. For the item hierarchy, cognitive items were considered the easiest to endorse, as they are not relevant to Maslow's hierarchy of needs and reflect direct instructional outcomes (Maslow, 1987). Items that match the proposed order were considered as meeting the holistic model, and items that did not match the proposed order were considered to fall outside the proposed model. Items that fell outside the proposed model were examined for possible mismatch reasons, and tagged for revision or removal in future iterations of the MESES.

To answer the fourth research question, *Are there differences in how people from different groups, with similar levels of the latent trait, are able to respond to items?*, a DIF analysis of student and parent responses was conducted. These two groups were selected because it was anticipated that these two groups would have the largest samples. The DIF analysis determined if there were variations in perceptions of respondents for specific items. This variation would demonstrate that either respondents were viewing items differently, or an issue with the item possibly led to a difference in views (Bond & Fox, 2007). Items that showed no DIF would be considered to demonstrate that persons in the different respondent groups were able to endorse the items at a similar level. DIF would be identified by determining what items following the DIF analysis had a t-value above 1.96 and thus, a statistically significant *p*-value at the .05 level. Items with DIF were identified for either removal or revision in future versions of the MESES.

Conclusion

This chapter presented an overview of the methods implemented in this study. Specifically, this chapter described the survey creation process and the data collection process in a sample Montessori high school. This chapter also described the Rasch RSM

analyses procedure, which served as the primary means of examining the tool's validity and reliability, building upon the validity process of Royal and Elahi (2011). Finally, this chapter discussed how each research question would be answered by the methodology. The following chapter includes the results of the Rasch RSM analyses.

CHAPTER IV

RESULTS

This chapter details the results of the study. The chapter begins with a review of the purpose of the study and the analysis procedure used. The chapter then continues on to provide details about the two waves of sample data. Finally, the chapter provides the results for Analysis 1, Analysis 2, and Analysis 3, and interprets each analysis.

Analysis Procedure

The purpose of this study was to construct a holistic education school evaluation tool for Montessori high schools and begin the validation process of examining the created tool. For this study, the Montessori Erdkinder School Evaluation Survey (MESES) was created following a process that developed items based on Erdkinder literature. To complete the validation component of this study, a validation procedure was designed using the Rasch Rating Scale Model (RSM) as the primary method of analysis. The Rasch RSM was selected because this analysis type allows the tool to be examined at the item level and acknowledges the usage of rating scale categories for the subject responses. The validation process includes examining the results of the Rasch RSM for item reliability, person reliability, the Rasch principal components analysis (PCA) of residual variances, item infit, item outfit, item difficulty spread, the comparison of item difficulty spread to a theoretical model of item difficulties, and the differential item functioning (DIF) between two groups of respondents.

The analysis procedure for this study included a series of three Rasch RSM analyses. This procedure required the usage of two waves of sample data as part of an item-anchoring strategy that allows for item difficulty estimates from earlier analyses to

be used in later analyses as a means of creating better overall model estimates. The two waves of sample data for this study were collected from subjects in a Montessori high school, the first in July 2015 and the second in October 2015. The first Rasch RSM analysis was conducted using the July 2015 sample data. Items from the first Rasch RSM that did not demonstrate issues of misfit, having infit or outfit statistics falling outside the range between -2.0 and 2.0, were marked for item anchoring in the second Rasch RSM analysis. The second Rasch RSM analysis results were then examined for item drift, a process that estimates how different an items difficulty estimate would have been had anchored items not been included in the analysis. Items from the second analysis that had item drift estimates at 0.6 or above and anchored items that demonstrated misfit were marked for removal from item anchoring in the third Rasch RSM analysis. The third Rasch RSM analysis was run using the item anchors that did not demonstrate item misfit or item drift in either of the prior analyses. The final Rasch RSM analysis was used to answer the research questions for this study.

Samples

July 2015

The July 2015 sample included 18 respondents. The descriptive statistics for this sample are reported in Table 2. The sex of the respondents included six males (33.3%) and 11 females (61.1%). Only one respondent (5.6%) chose not to identify their sex. Respondents' relationship to the school was most commonly reported as being that of parent/guardian, with seven respondents (38.9%) identifying as such. The remaining sample included a total of five students (27.8%), four teachers (22.2%), and one recent

graduate (5.6%). None of the respondents identified as being a staff member, and one respondent (5.6%) did not identify their relationship to the school.

Table 2

July 2015 Sample Descriptive Statistics (N=18)

<i>Descriptive</i>	<i>n</i>	<i>%</i>
Sex		
Male	6	33.3
Female	11	61.1
Relationship to School		
Student	5	27.8
Parent/Guardian	7	38.9
Teacher	4	22.2
Staff Member	0	0.0
Recent Graduate	1	5.6

October 2015

The October 2015 sample included 36 respondents. The descriptive statistics for this sample are reported in Table 3. The sex of the respondents included 15 males (41.7%) and 20 females (55.6%). The sample included one (2.8%) respondent who did not identify their sex. The respondents' relationship to the school was primarily as parent/guardian, with 16 (44.4%) respondents (44.4%) identifying as such. The sample also included 13 students (36.1%), five teachers (13.9%), and two staff members (5.6%).

Table 3

October 2015 Sample Descriptive Statistics (N=36)

<i>Descriptive</i>	<i>n</i>	<i>%</i>
Sex		
Male	15	41.7
Female	20	55.6
Relationship to School		
Student	13	36.1
Parent/Guardian	16	44.4
Teacher	5	13.9
Staff Member	2	5.6

Analysis 1

Analysis 1 was a Rasch RSM analysis using the July 2015 sample data. For Analysis 1, the person reliability estimate was .89 and the item-reliability estimate was .77. Although the person reliability level was above the .80 threshold suggested by Linacre (2015a), the item reliability estimate was not. These results suggest that for Analysis 1, there was not a large enough sample to estimate item difficulty levels effectively, although the analysis was able to estimate person-ability levels effectively.

After examining the reliability estimates, dimensionality of the MESES in Analysis 1 was determined. Table 4 includes the results of the principal components analysis of the Rasch residuals for Analysis 1.

Table 4

Analysis I Variance Estimates

<i>Identifier</i>	<i>Eigenvalue</i>	<i>%</i>
Total raw variance in observations	42.0	100.0
Raw variance explained by measures	19.0	45.3
Persons	8.6	20.4
Items	10.5	24.9
Raw unexplained variance (total)	23.0	54.7
1st contrast	4.7	11.2
2nd contrast	4.0	9.6
3rd contrast	3.1	7.4
4th contrast	2.5	6.0
5th contrast	2.3	5.4

The eigenvalue of the first contrast was 4.7. This contrast eigenvalue being above 2.0 suggests that the tool is multidimensional. Table 5 includes the standardized residual loadings for the first contrast and Figure 3 is the mapping of these loadings.

Figure 3

Analysis 1 Standardized Residual Loadings for First Contrast

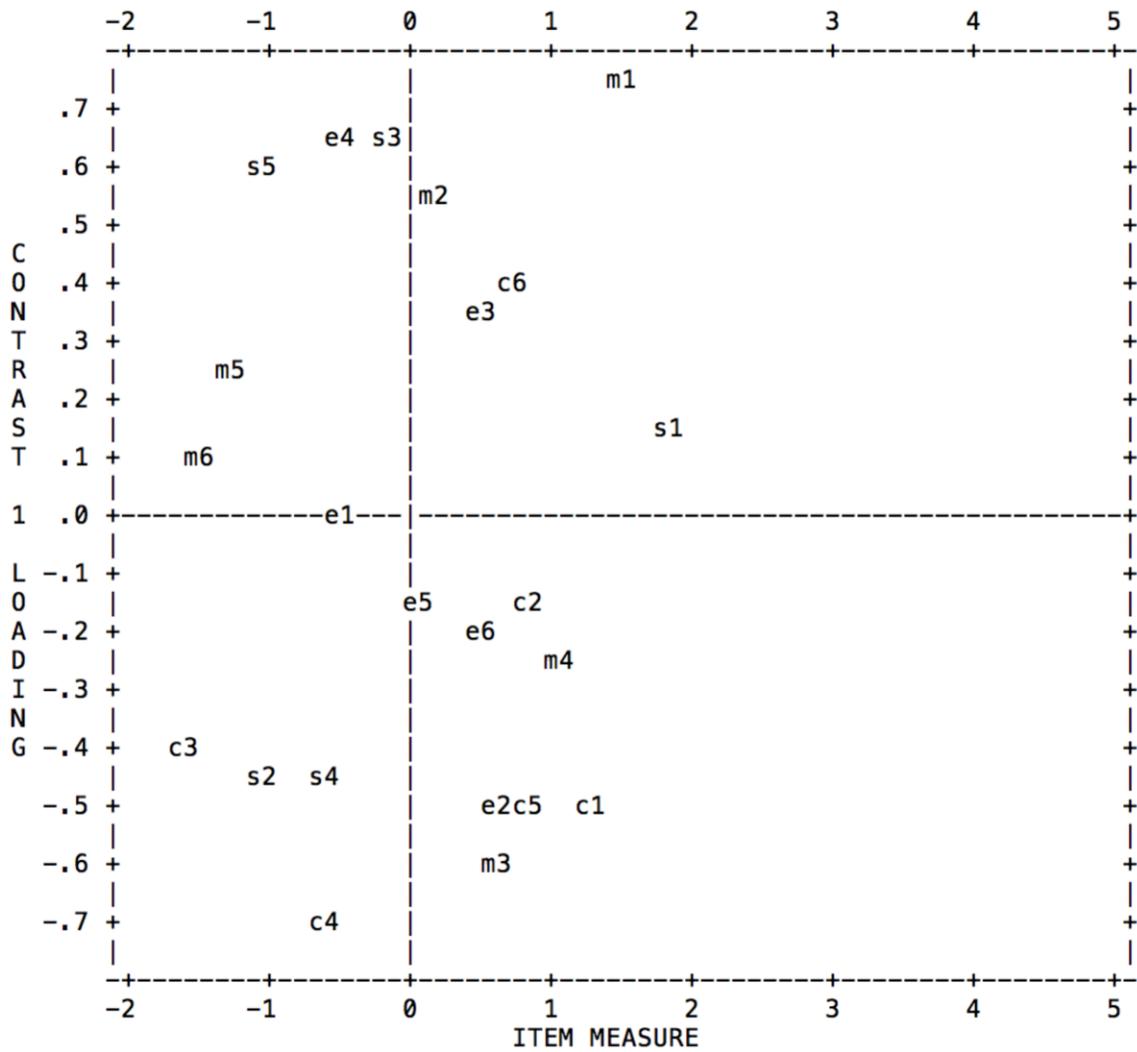


Table 5
Analysis 1 Item Estimates

<i>Item</i>	<i>First Contrast Loading</i>	<i>Difficulty</i>	<i>Model S.E.</i>	<i>Infit</i>		<i>Outfit</i>	
				<i>MNSQ</i>	<i>t</i>	<i>MNSQ</i>	<i>t</i>
c1	-.48	0.99	.41	0.63	-1.0	0.55	-1.3
c2	-.13	0.77	.39	1.29	0.9	1.12	0.4
c3	-.41	-1.71	.51	1.33	1.0	1.24	0.6
c4	-.71	-0.69	.47	1.20	0.6	1.17	0.5
c5	-.49	0.63	.39	1.49	1.3	1.33	0.9
c6	.38	0.65	.41	1.06	0.3	1.01	0.2
s1	.15	1.80	.38	2.40	3.2	3.03	4.0
s2	-.44	-1.13	.46	1.14	0.5	0.91	-0.1
s3	.63	-0.35	.42	0.69	-0.8	0.69	-0.8
s4	-.44	-0.73	.44	0.73	-0.7	0.70	-0.7
s5	.62	-1.13	.46	1.14	0.5	1.05	0.3
e1	.02	-0.52	.45	1.15	0.5	1.06	0.3
e2	-.52	0.58	.43	0.50	-1.4	0.47	-1.5
e3	.33	0.47	.46	0.87	-0.2	1.05	0.3
e4	.67	-0.51	.46	0.74	-0.7	0.68	-0.7
e5	-.15	0.05	.43	0.34	-2.2	0.36	-2.0
e6	-.21	0.48	.39	0.73	-0.7	0.73	-0.7
m1	.74	1.46	.36	1.07	0.3	1.03	0.2
m2	.55	0.16	.42	1.04	0.2	0.88	-0.2
m3	-.58	0.59	.42	0.55	-1.3	0.47	-1.5
m4	-.27	1.02	.41	0.49	-1.6	0.50	-1.5
m5	.25	-1.34	.47	1.52	1.4	1.18	0.5
m6	.10	-1.57	.48	0.93	-0.1	0.92	0.0

Note. MNSQ = mean-square.

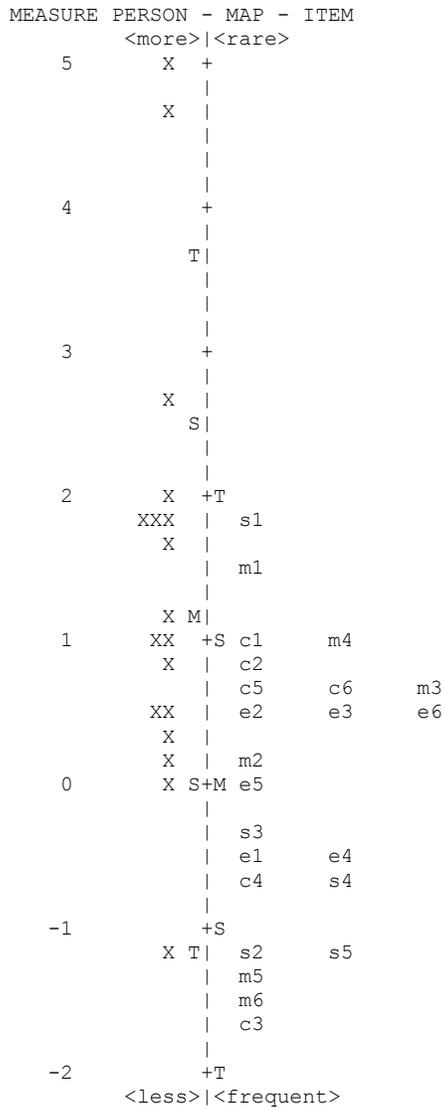
The contrast loadings show that neither the positive loading items nor the negative loading items are composed primarily of a single facet, nor can all the items of a single facet be identified as entirely positive or negative. The positive and negative loading items are divided into nearly equal groups, with 11 items having positive loadings and 12 items having negative loadings. Examining the content of the positive loading items and the negative loading items do not provide any indication of a clear dimension that is being measured in either grouping. This lack of a clear dimension being measured would suggest that the possibility of multidimensionality is not of concern, and occurs from the inclusion of items from several facets that make up a larger latent trait.

After determining the reliability and dimensionality, the item infit and outfit t-statistics in Analysis 1 were examined for issues of misfit. The infit and outfit t-statistics for Analysis 1 are reported in Table 5. The fit statistics show that all but two of the items, *s1* and *e5*, fell within the acceptable infit and outfit t-statistics. Item *s1* had an infit t-statistic of 3.2 and an outfit t-statistic of 4.0, and item *e5* had an infit t-statistic of -2.2 and an outfit t-statistic of -2.0. These results suggest that items *s1* and *e5* are not effectively measuring respondents' ability levels.

Item difficulty estimates, also reported in Table 5, were then examined to determine the range of item difficulty levels. For Analysis 1 the item difficulty levels were between -1.71 and 1.80. The spread of items shows that item difficulty levels are between the moderately-easy to moderately-challenging to endorse range, meaning that the tool does not include very-easy and very-challenging items. Figure 4 includes the Wright map for Analysis 1.

Figure 4

Analysis 1 Wright Map



The Wright map shows there is a redundancy of items, especially between the 0.5 and 1.0 difficulty level.

To further the validation process, the item difficulty results are compared to the proposed theoretical model of Maslow's hierarchy of needs. Figure 4 shows that the moral facet items for Analysis 1 are some of the most difficult items to endorse, excluding items *m5* and *m6*. The emotional facet items fell within the mid-range of difficulty level, with the emotional items having difficulty levels below the item difficulties of the challenging moral items and above the item difficulties of many of the social items. The social facet items are among the easier items to endorse, excluding item *s1*. The cognitive facet items according to Maslow's theory should have had the lowest item difficulty levels, however, in the analysis they did not. The cognitive facet items had item difficulty levels that ranged from moderately-easy to moderately-difficult to endorse. Although the cognitive facet items did not follow the theoretical model, the majority of the items in the other facets followed Maslow's theoretical model, which suggests that many of the item difficulty levels are appropriate for the tool.

Finally, to determine if there are differences in how the ability of respondents to endorse an item, the results of the DIF analysis between students and parents was examined. For the Analysis 1 DIF analysis, the data from the recent graduate was included as part of the student data. The DIF results for Analysis 1 are reported in Table 6.

Table 6

Analysis 1 Differential Item Functioning Estimates for Students and Parents

<i>Item</i>	<i>t</i>	<i>df</i>	<i>p</i>
c1	-1.31	6	.24
c2	-1.62	8	.14
c3	-0.96	8	.37
c4	-1.56	8	.16
c5	1.04	9	.33
c6	0.36	7	.73
s1	0.57	7	.59
s2	-0.32	9	.76
s3	0.07	9	.95
s4	-1.13	8	.29
s5	-0.12	9	.91
e1	1.79	9	.11
e2	0.81	8	.44
e3	0.72	5	.51
e4	-0.10	7	.92
e5	-0.33	9	.75
e6	0.22	9	.83
m1	0.57	9	.59
m2	1.15	9	.28
m3	-1.62	6	.16
m4	0.35	5	.74
m5	-0.32	9	.76
m6	1.71	9	.12

The results of the DIF analysis shows that there is no evidence of DIF for students and parents.

Analysis 2

Analysis 2 was a Rasch RSM analysis using the October 2015 sample data with item difficulty estimates anchored from the results of Analysis 1. Only the item difficulty estimates for items that did not demonstrate misfit in Analysis 1 were anchored. The

Table 7

Analysis 2 Variance Estimates

<i>Identifier</i>	<i>Eigenvalue</i>	<i>%</i>
Total raw variance in observations	46.7	100.0
Raw variance explained by measures	23.7	50.7
Persons	13.7	29.4
Items	10.0	21.3
Raw unexplained variance (total)	23.0	49.3
1st contrast	3.0	6.4
2nd contrast	2.7	5.7
3rd contrast	2.4	5.2
4th contrast	2.2	4.8
5th contrast	1.9	4.0

reliability estimates were examined as the first part of the validation process. The person reliability estimate for Analysis 2 was .89 and the item-reliability estimate was .88, both above the .80 level recommended by Linacre (2015b). These reliability estimates suggest that the analysis had a sufficient number of items and respondents to develop person-ability and item difficulty estimates that can be interpreted confidently.

After examining the reliability estimates, Analysis 2 was examined for dimensionality. The results of the PCA of the Rasch residuals in Analysis 2 are located in Table 7, and suggest that the structure may be multidimensional. The first contrast had a 3.0 eigenvalue, suggesting that an additional dimension may be present. The standardized residual loadings of the first contrast for Analysis 2 are reported in Table 8 and the mapping of these loadings is demonstrated in Figure 5.

Figure 5

Analysis 2 Standardized Residual Loadings for First Contrast

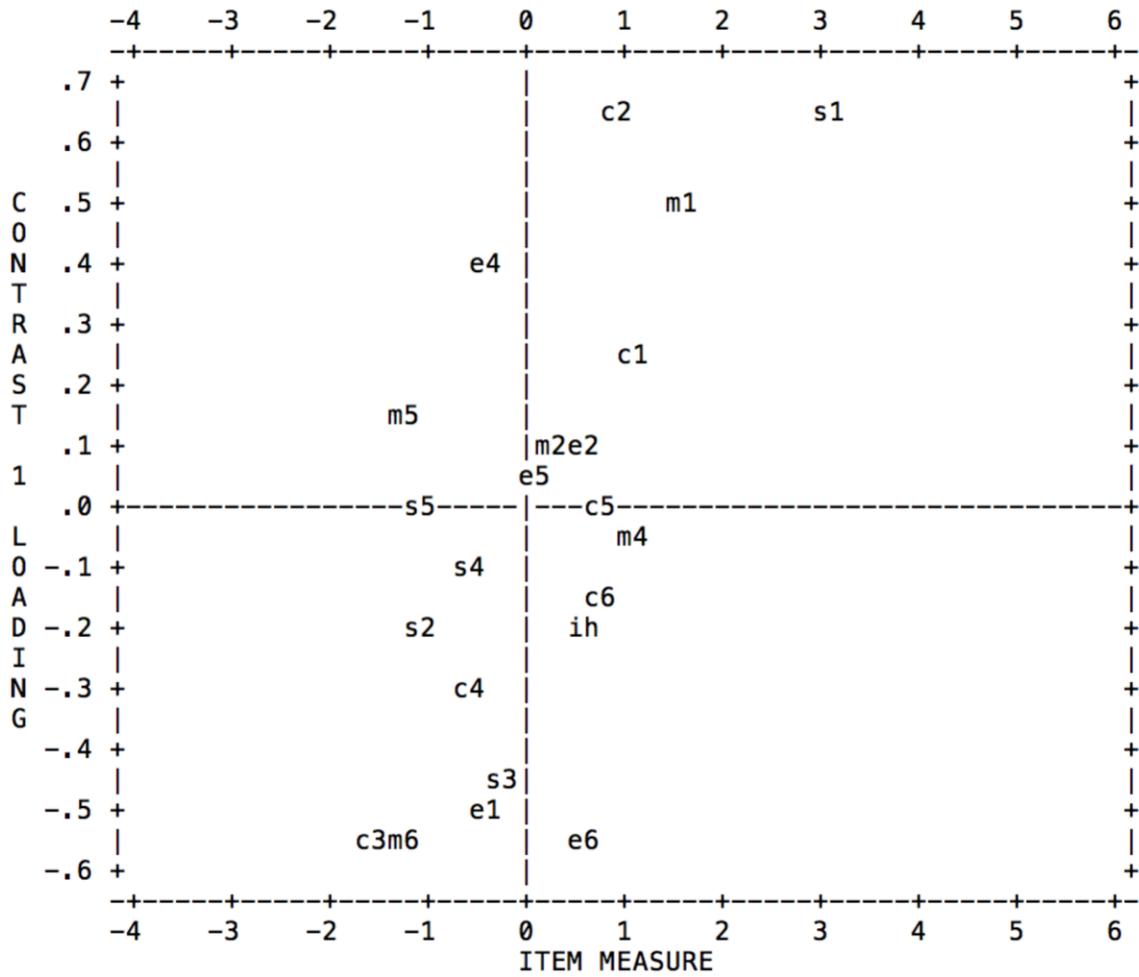


Table 8
Analysis 2 Item Estimates

<i>Item</i>	<i>First Contrast Loading</i>	<i>Difficulty</i>	<i>Model S.E.</i>	<i>Infit</i>		<i>Outfit</i>		<i>Displacement</i>
				<i>MNSQ</i>	<i>t</i>	<i>MNSQ</i>	<i>t</i>	
c1*	.23	0.99	.31	1.03	0.2	0.97	0.0	0.18
c2*	.65	0.77	.32	0.92	-0.2	1.06	0.3	-0.34
c3*	-.53	-1.71	.41	1.37	1.3	1.12	0.4	0.79
c4*	-.30	-0.69	.35	0.80	-0.8	0.84	-0.3	0.61
c5*	.00	0.63	.33	1.05	0.3	1.02	0.2	-0.68
c6*	-.17	0.65	.32	1.08	0.4	1.34	1.1	-0.37
s1	.67	3.00	.28	2.40	4.4	9.90	9.9	0.01
s2*	-.21	-1.13	.36	0.67	-1.5	0.58	-0.9	-0.39
s3*	-.47	-0.35	.33	0.76	-1.0	0.72	-0.8	-0.63
s4*	-.08	-0.73	.35	1.13	0.6	1.26	0.8	0.10
s5*	-.02	-1.13	.37	0.68	-1.4	0.68	-0.6	0.13
e1*	-.48	-0.52	.33	0.81	-0.8	0.70	-0.9	-0.20
e2*	.10	0.58	.32	0.32	-3.4	0.33	-3.0	-0.05
e3*	-.18	-0.47	.32	0.96	-0.1	0.88	-0.3	0.19
e4*	.38	-0.51	.34	2.30	4.0	3.30	4.2	0.80
e5	.06	-0.06	.32	0.78	-0.9	0.82	-0.5	0.00
e6*	-.55	0.48	.32	1.03	0.2	-0.95	-0.1	-1.29
m1*	.52	1.46	.28	1.10	0.5	1.58	1.9	0.29
m2*	.11	0.16	.31	0.43	-2.9	0.43	-2.5	0.08
m3*	-.19	0.59	.30	0.75	-1.0	0.72	-1.0	0.35
m4*	-.07	1.02	.29	0.63	-1.6	0.60	-1.6	-0.67
m5*	.16	-1.34	.37	3.14	5.7	4.00	3.7	0.53
m6*	-.56	-1.57	.40	1.46	1.6	1.67	1.2	0.95

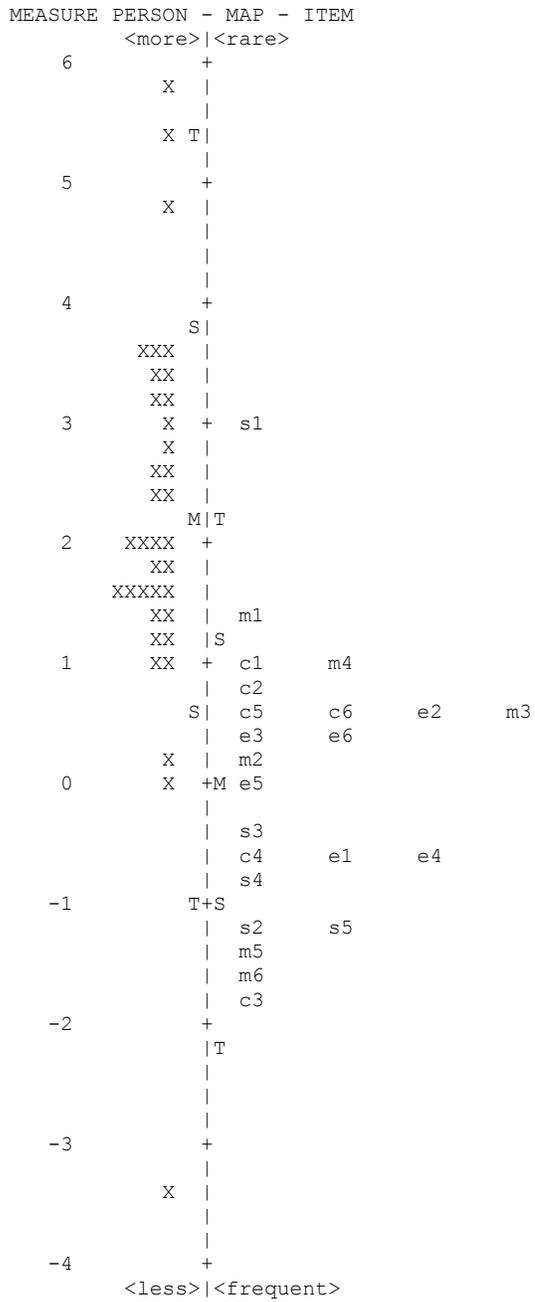
Note. MNSQ = mean-square; *Item uses difficulty estimates anchored from previous analysis.

There are 10 positive loading items and 13 negative loading items. Examining the items in the positive loading did not suggest a particular theme was evident between the associating items, as each item's focus was unconnected and from varying facets. The items in the negative loading also lacked an evident theme and were from several facets. The negative loading, however, did contain the majority of the social items. These results suggest that there may be multidimensionality; however, the lack of clear themes between the associating positive or negative items implies that the tool may be unidimensional.

The item level results in Analysis 2 were examined next; specifically, the item fit and item difficulty levels. The item infit and outfit statistics for Analysis 2 are reported in Table The fit statistics suggest that five items had issues with misfit, specifically items *s1*, *e2*, *e4*, *m2*, and *m5*. Item difficulty levels ranged from -1.71 to 3.00, although the majority of these items were anchored to the item difficulty estimates from Analysis 1. Figure 6 is the Wright map for Analysis 2.

Figure 6

Analysis 2 Wright Map



As the map shows, the majority of items were moderately-easy to moderately-challenging to endorse, with the exception of item *s1* which was very difficult to endorse. Item *s1*, however, was a misfit item that cannot be interpreted as measuring person-ability accurately. Table 8 also reports the item displacement values for the anchored items. There are eight items with displacement values of 0.6 or more, suggesting they should be unanchored in additional analyses. The displaced items were *c3*, *c4*, *c5*, *s3*, *e4*, *e6*, *m4*, and *m6*. Analysis 2 item difficulty results suggested that additional items at the very-easy to endorse and very-challenging to endorse levels should be added, and that several items needed to be unanchored for Analysis 3.

When comparing the item difficulty estimates to the theoretical model of Maslow's hierarchy of needs, the results show that the item difficulty estimates for Analysis 2 follow a pattern that partially replicates Maslow's theoretical model. The moral facet items are among some of the most difficult to endorse, specifically, *m1*, *m3*, and *m4*. The emotional facet items follow in difficulty level, overlapping partially with the moral items in the harder to endorse difficulty level and overlapping partially with the social items in the easier-to-endorse difficulty level. The social facet items are among some of the easiest to endorse items. Although the moral, emotional, and social facet items partially ascribe to the theoretical model, the cognitive facet items do not. Instead of being the easiest items to endorse, many of the cognitive facet items are among the most difficult to endorse. These results demonstrate many of the items have appropriate difficulty level; however, the cognitive facet items do not match expectations.

The next stage of examining the Analysis 2 results was to examine if there was DIF between the students and parents. Table 9 includes the results of the Analysis 2 DIF analysis.

Table 9

Analysis 2 Differential Item Functioning Estimates for Students and Parents

<i>Item</i>	<i>t</i>	<i>df</i>	<i>p</i>
c1	0.79	22	.44
c2	0.47	23	.65
c3	-0.10	25	.92
c4	-0.32	24	.76
c5	-1.15	23	.26
c6	-0.53	22	.60
s1	0.60	20	.56
s2	-0.71	26	.48
s3	-0.52	25	.61
s4	-0.58	25	.56
s5	0.99	25	.33
e1	0.25	26	.80
e2	-0.41	22	.68
e3	0.66	23	.52
e4	0.77	24	.45
e5	-0.21	26	.84
e6	0.16	22	.88
m1	-1.22	25	.24
m2	0.77	26	.45
m3	1.36	25	.19
m4	0.04	25	.97
m5	0.07	26	.95
m6	0.08	25	.94

These results suggest that there was no issue with DIF between the student and parent groups.

Analysis 3

Analysis 3 was the final stage of the validation procedure. Analysis 3 included conducting a Rasch RSM on the October 2015 dataset after unanchoring the misfit items from the previous analyses and the displaced items from Analysis 2. For analysis 3, the reliability estimates were on the high level, with both being above Linacre's (2015b) reliability recommendations. Person reliability was .89 and item reliability was .86. These reliability estimates mean that there was a sufficient number of items to effectively estimate person-ability levels, as well as a sufficiently large sample to determine item difficulty levels. These reliability results suggest that the Rasch RSM analysis estimates in Analysis 3 can confidently be interpreted.

After determining reliability, the dimensionality estimates were examined. The PCA of the Rasch residuals results reported in Table 10 shows that the first contrast had an eigenvalue of 3.0.

Table 10

Analysis 3 Variance Estimates

<i>Identifier</i>	<i>Eigenvalue</i>	<i>%</i>
Total raw variance in observations	49.3	100.0
Raw variance explained by measures	26.3	53.3
Persons	18.3	37.1
Items	8.0	16.3
Raw unexplained variance (total)	23.0	46.7
1st contrast	3.0	6.2
2nd contrast	2.7	5.5
3rd contrast	2.4	5.0
4th contrast	2.2	4.5
5th contrast	1.9	3.8

The first contrast having an eigenvalue above 2.0, suggests that the tool is multidimensional. Table 11 includes the standardized residual loadings for the items in Analysis 3 and the loadings are mapped out in Figure 7.

Figure 7

Analysis 3 Standardized Residual Loadings for First Contrast

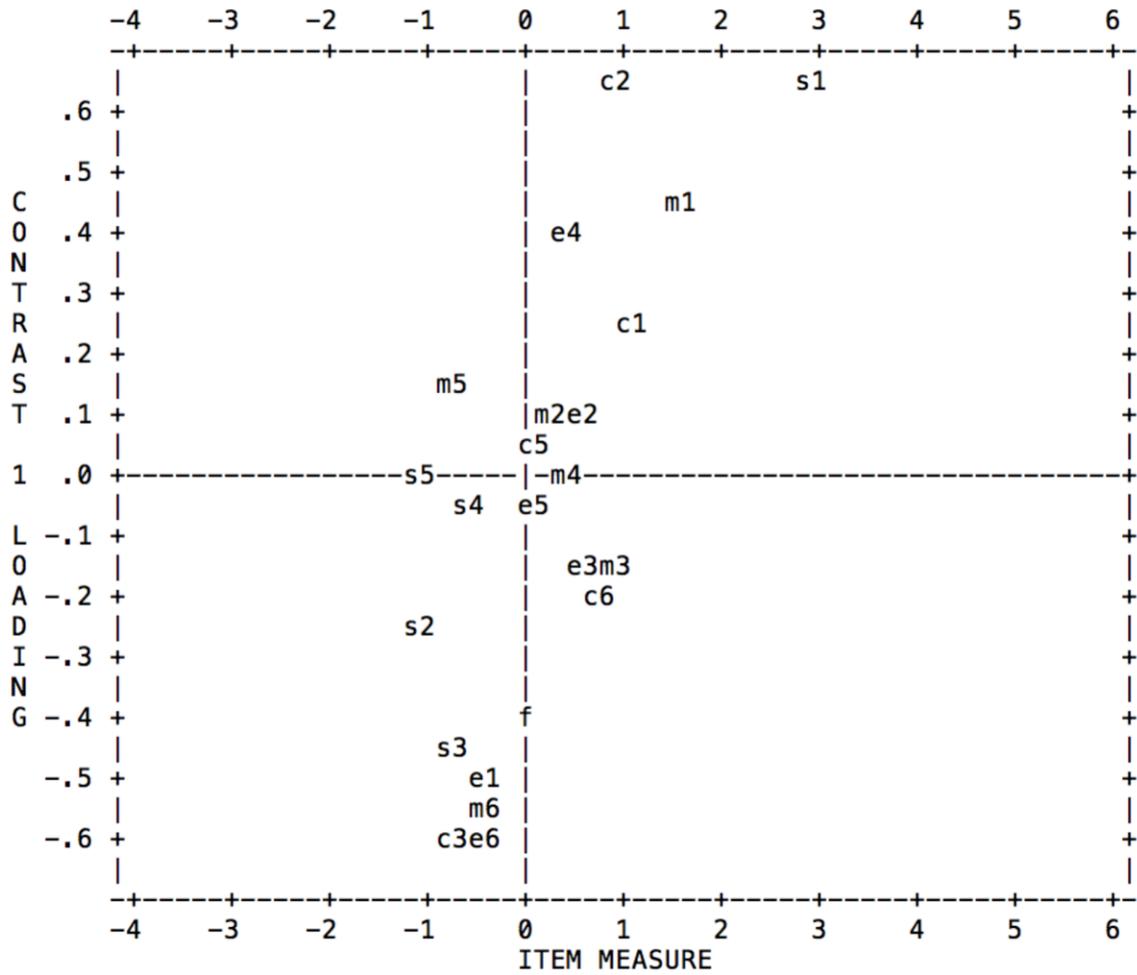


Table 11
Analysis 3 Item Estimates

<i>Item</i>	<i>First Contrast Loading</i>	<i>Difficulty</i>	<i>Model S.E.</i>	<i>Infit</i>		<i>Outfit</i>		<i>Displacement</i>
				<i>MNSQ</i>	<i>t</i>	<i>MNSQ</i>	<i>t</i>	
c1*	.24	0.99	.30	0.98	0.0	0.92	-0.2	0.2
c2*	.63	0.77	.31	0.88	-0.4	0.99	0.1	-0.3
c3	-.60	-0.87	.35	0.84	-0.6	0.70	-0.7	0.0
c4	-.38	-0.08	.32	0.59	-1.9	0.58	-1.5	0.0
c5	.06	-0.02	.34	0.96	0.0	1.05	0.3	0.0
c6*	-.19	0.65	.31	1.02	0.2	1.26	0.9	-0.4
s1	.63	2.88	.28	2.29	4.1	9.90	9.9	0.0
s2*	-.23	-1.13	.36	0.65	-1.5	0.58	-1.0	-0.3
s3	-.43	-0.88	.35	0.77	-0.9	0.77	-0.5	0.0
s4*	-.06	-0.73	.34	1.10	0.5	1.24	0.7	0.2
s5*	-.02	-1.13	.37	0.69	-1.3	0.68	-0.6	0.2
e1*	-.52	-0.52	.33	0.79	-0.9	0.67	-1.0	-0.2
e2	.08	0.52	.32	0.30	-3.6	0.32	-3.1	0.0
e3*	-.16	0.47	.31	0.91	-0.3	0.84	-0.5	0.2
e4	.38	0.29	.31	1.77	2.5	2.05	2.9	0.0
e5	-.03	-0.04	.31	0.74	-1.1	0.77	-0.7	0.0
e6	-.60	-0.68	.35	0.83	-0.7	0.74	-0.6	0.0
m1*	.47	1.46	.27	1.03	0.2	1.44	1.5	0.2
m2	.09	0.24	.30	0.40	-3.1	0.41	-2.7	0.0
m3*	-.17	0.59	.30	0.71	-1.2	0.68	-1.2	0.3
m4	.00	0.34	.30	0.53	-2.2	0.51	-2.0	0.0
m5	.17	-0.78	.34	2.56	4.7	2.81	3.4	0.0
m6	-.57	-0.58	.34	0.87	-0.5	0.86	-0.3	0.0

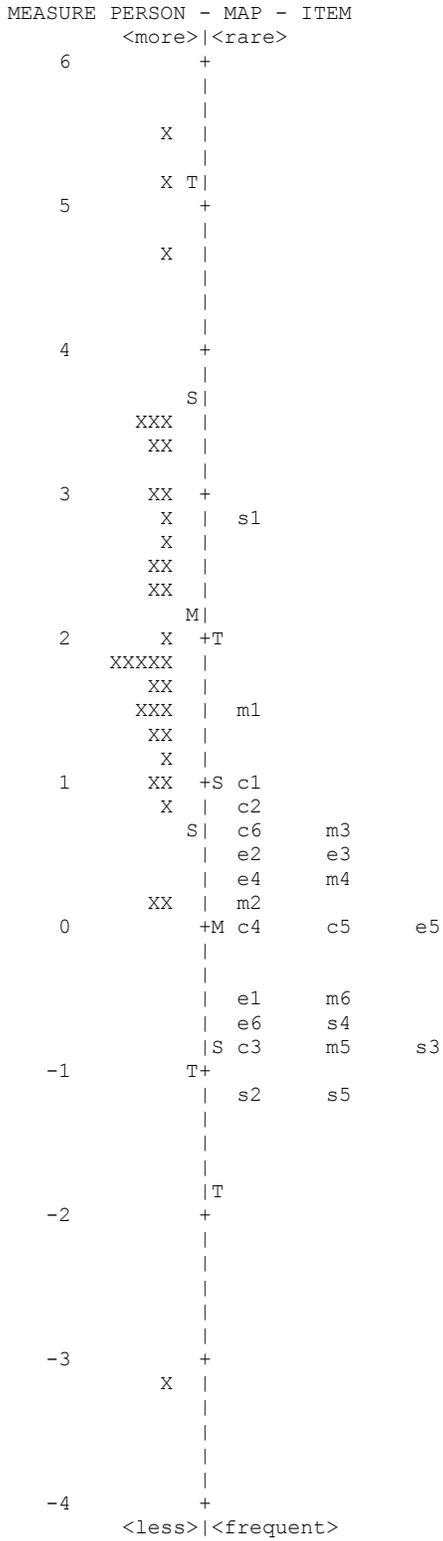
Note. MNSQ = mean-square; *Item uses difficulty estimates anchored from previous analyses.

There are 10 positive loading items with no clear indicator of a theme between the items within the positive loading. Noteworthy among the positive loading items is that the social facet items are lacking from the group, with only *s1* having a positive loading. Furthermore, there are 13 negative loadings items. The negative loading items are from all four facets, with no clear indicator of a theme that connects the items. This lack of a cohesive theme in either the positive or negative loading items suggests that the tool may not be multidimensional, despite the eigenvalue of the first contrast.

Following reliability and dimensionality, the item level results of Analysis 3 were examined. The item fit results for Analysis 3 reported in Table 11 show that six items had misfit. The misfit items in Analysis 3 were *s1*, *e2*, *e4*, *m2*, *m4*, and *m5*. As these results show, half of the moral facets items had an issue with misfit. The item difficulty results, also reported in Table 11, show that the difficulty estimates ranged from -1.13 to 2.88. Although the most challenging to endorse item was *s1* at 2.88, the majority of the challenging to endorse items were at the moderate level. The majority of items ranged in the moderately-difficult to moderately-easy to endorse levels, as demonstrated in the Wright map in Figure 8.

Figure 8

Analysis 3 Wright Map



As the Wright map also shows, there is a great deal of overlap in item difficulty levels within the same facet. For example, the cognitive facet items *c1*, *c2*, and *c6* had similar item difficulty levels, the social facet items *s2*, *s3*, and *s5* had similar item difficulty levels, and the emotional facet items *e1*, *e2*, and *e3* had similar item difficulty levels. Examining the item difficulty estimates shows there is a need for additional items at the most and least challenging endorsement levels, as well as an opportunity for item reduction to remove redundant items within the same facet that have similar difficulty levels.

Examining the item level results provide the basis to continue the validation process by examining how the order of item difficulties in Analysis 3 compares to the theoretical model of Maslow's hierarchy of needs. The results for Analysis 3 suggest that the item difficulty order is somewhat similar to the theoretical model, when considering only the items that fit. When considering only the items that are not misfit, the moral facet items *m1* and *m3* are among the most difficult to endorse, and the social facet items *s2*, *s3*, *s4*, and *s5* are among the easiest to endorse. The emotional facet items *e1*, *e3*, and *e5* fall below the moral fitting items and above the social fitting items. The cognitive facet items range in their difficulty levels. These results suggest that the Analysis 3 order of item difficulties mostly followed Maslow's hierarchy of needs, although the cognitive facet items do not follow the theoretical model.

The final component of the validation process is examining the DIF of Analysis 3. The DIF results for students and parents of Analysis 3 are reported in Table 12.

Table 12

Analysis 3 Differential Item Functioning Estimates for Students and Parents

<i>Item</i>	<i>t</i>	<i>df</i>	<i>p</i>
c1	0.78	22	.45
c2	0.46	23	.65
c3	-0.08	25	.94
c4	-0.28	24	.78
c5	-1.10	23	.28
c6	-0.52	22	.61
s1	0.61	20	.55
s2	-0.68	26	.50
s3	-0.49	25	.63
s4	-0.56	25	.58
s5	0.98	25	.34
e1	0.27	26	.79
e2	-0.40	22	.70
e3	0.66	23	.52
e4	-0.73	24	.47
e5	-0.18	26	.86
e6	0.16	22	.87
m1	-1.18	25	.25
m2	0.77	26	.45
m3	1.34	25	.19
m4	0.06	25	.95
m5	0.08	26	.93
m6	0.11	25	.91

As the results show, there is no indication of DIF between students and parents for any of the items.

Conclusion

This chapter provided a detailed description of the results for the study, beginning with descriptive details about the two samples collected. The chapter then described the reliability, dimensionality, item fit, item difficulty estimates, and DIF estimates for the

three analyses conducted. The next chapter answers the research questions of this study using these results, and discusses the possible implications of this study.

CHAPTER V

DISCUSSION

The purpose of this study was to construct a holistic education school evaluation tool that could be implemented in Montessori high schools, as well as begin the validation process of examining the constructed tool. This study began by discussing the need for the development of such a tool and the research questions that would be guiding the study. Then, a detailed review of the literature on Montessori adolescent philosophy, holistic education, and the item-response theory one-parameter Rasch model was given. Following the review of literature, the process used to construct the Montessori Erdkinder School Evaluation Survey (MESES) and the validation procedure used in this study were described. The validation procedure included using a series of Rasch Rating Scale Model (RSM) analyses, which included item anchoring to develop better Rasch estimates. For the validation procedure, the Rasch RSM estimates were examined for reliability, dimensionality, item fit, item difficulty, and differential item functioning (DIF) between students and parents. Furthermore, the order of the item difficulty estimates was compared to the theoretical model of these estimates according to Maslow's hierarchy of needs. This chapter includes a listing of each research question and a discussion of the answer for each research question based on the findings from the final Rasch RSM analysis. The chapter concludes with a discussion of the implications of this study for schools, the school evaluation process, and public policy.

Research Questions and Findings

1. How well does the evaluation tool measure the latent trait holistic education concept of the Montessori Erdkinder school principles?

To answer the first research question, the dimensionality results for Analysis 3 were examined. The Rasch principal components analysis of the residual variances allowed for the interpretation that the tool may be multidimensional, thus measuring more than a single latent trait. This initial interpretation could be made because the first contrast had an eigenvalue of 3.0, and Linacre (2015a) suggests that a value above 2.0 on the first contrast is a possible indicator of a tool being multidimensional in its measurement. However, Linacre suggests that when there is an indicator of possible multidimensionality, to examine the positive and negative loadings in the first contrast and determine if the items make up a unique measurable concept distinct from the latent trait being measured. If a trait is found, then the tool can be considered multidimensional. However, if no trait is found, then the suspicion of multidimensionality from the earlier examination of the eigenvalues can be considered a byproduct of measuring a latent trait that is comprised of several complex concepts. Upon examining the items that make up the groupings for the positive and negative loadings of the first contrast in Analysis 3, it can be interpreted that the MESES is measuring a single latent trait.

The argument for the MESES measuring a single latent trait is based on examining the items that make up the positive and negative groupings of the first contrast, and finding that there is no conceptually distinct concept discernable in the grouping of the items. The items in the positive loading group of the first contrast were *c1, c2, c5, s1, e2, e4, m1, m2, m4, and m5*. Examining the positive items by their facets supports that there is no distinct concept present. The negative items similarly do not have a unique theme that becomes present when examining the grouping of items of the first contrast. The negative loading items on the first contrast were *c3, c4, c6, s2, s3, s4,*

s5, e1, e5, e6, m3, and m6. Again, when examining how these individual negative loading items are associated with one another, there is no clear indication of an underlying concept that is being measured.

After examining the items in the positive and negative loading groupings respectively, there is no clear theme that is present in either. Given the high reliability results at the person and item levels, these item loading groupings can be interpreted confidently. This evidence supports the conclusion that the MESES is measuring a single latent trait, and that the high eigenvalue of the first contrast can be attributed to the MESES measuring a complex latent trait comprised of several concepts.

2. How well do the individual items fit the Montessori Erdkinder school principles latent trait reflected on the evaluation tool?

To answer the second research question, the item fit statistics for Analysis 3 required examination. For items to be considered as fitting and measuring the latent trait, they needed to have infit and outfit t-statistics at 2.0 or below. The results of Analysis 3 suggest that the majority of the items on the MESES, 17 out of the 23, fit the model and thus measured the latent trait well. A total of six items were considered to be misfit with infit or outfit t-statistics above 2.0, and thus did not measure the latent trait well.

The misfit items on the MESES for Analysis 3 were *s1, e2, e4, m2, m4, and m5*. The item *s1, ... are taught to challenge the rules of society*, had the most extreme misfit of any of the items, with an infit t-statistic of 4.1 and an outfit t-statistic of 9.9. These results suggest that item *s1* had extreme issues with being able to measure the latent trait. The reasoning for the misfit of item *s1* may be attributed to the item being negatively

phrased. As such, the negative nature of the item may have been confusing to respondents. Respondents may have misunderstood the item's relation to Montessori Erdkinder, and may have been unsure of how to respond, especially if they wanted to avoid negatively characterizing the school. The term "challenge" may have been interpreted differently, with some respondents believing that it referred to students being taught to protest, write letters, or commit illegal actions against social establishments. There is no means of knowing how respondents may have differed in their interpretation of this item; however, the extreme misfit of item *s1* suggests it was likely interpreted in a range of capacities by respondents. Future iterations of the MESES may benefit from using a non-negatively phrased item and examining how the newly worded item fits.

A total of two items in the emotional facet had misfit as well, items *e2* and *e4*. Item *e2*, ... *are taught problems in society are fixable*, had an infit t-statistic of -3.6 and an outfit t-statistic of -3.1. There is no clear indicator as to why item *e2* may have misfit, although there are a few possibilities. Respondents may have had an issue with the societal focus of item *e2*, and not connected how students' learning may address societal issues. Respondents may have also had an issue with item *e2* suggesting that there are problems in society that need "fixing" or with the concept that problems can be "fixed." Respondents may have struggled to separate the school's teachings from students' innate beliefs. In the future, item *e2* could be altered to include specific examples of how students may be taught to fix a societal issue, such as through concepts like work, research, or volunteering. Item *e4*, ... *feel their future has already been decided*, was a negatively phrased item and had an infit *t* of 2.5 and an outfit *t* of 2.9. The initial issue with *e4* may be attributed to the reverse-coded nature of the item. The negative phrasing

of item *e4* may have led to confusion about its relevance to Montessori Erdkinder. Respondents may have been unsure about how to respond, particularly if they were attempting to give the school a positive review. An additional possibility is that respondents agreed with item *e4* because they feel that the students are on set academic paths leading towards particular goals and outcomes. A future iteration of the MESES should include a non-negatively phrased version of item *e4*. The MESES would benefit from altering item *e2* and *e4* in the future, as this analysis shows they are both misfit items in their present state.

The moral facet had the most issues with misfit in Analysis 3. A total of three items in the moral facet had misfit, items *m2*, *m4*, and *m5*. Item *m2*, ... *respect one another*, had an infit t-statistic of -3.1 and an outfit t-statistic of -2.7. Item *m2* has no clear semantic indicator as to why it may be a misfit item. It is possible that ‘respect’ was interpreted in varying ways by respondents, with some respondents viewing respect as referring to students’ language and actions, and others viewing respect as how students’ perceived one another. Although there is no clear indicator of the specific reason that item *m2* misfit, altering the item in future iterations to be specific about how students demonstrate respect may improve the item’s fit. Item *m4*, ... *are taught to address the problems in society*, had an infit t-statistic of -2.2 and an outfit t-statistic of -2.0. The misfit in item *m4* may be attributed to a few factors. For item *m4*, respondents may have had issues with the idea that society has problems or respondents may have struggled to understand how students would be taught to address societal problems. There may have also been variation in how respondents interpreted the term “address” in item *m4*, such as respondents believing this referred to physical actions like volunteering and others

believing this referred to political actions like promoting policy changes. Being specific about actions may be a way that item *m4* can be altered to provide a more accurate measurement in future iterations of the MESES. The final misfit moral item was *m5*, ... *feel that working is shameful*. Item *m5* had an infit t-statistic of 4.7 and an outfit t-statistic of 3.4, and was a negatively phrased item. The misfit of item *m5* can possibly be attributed to the item being a negatively phrased item. The negative nature of item *m5* may have confused respondents who either understood the item was not an Erdkinder value and questioned why it was on the MESES, or who believed it was an Erdkinder value but were uncomfortable endorsing the item. The negative nature of item *m5* may have been particularly challenging for respondents who wanted to give the school a positive review, and were confused about how to endorse the item so to reflect their positive view. Including item *m5* in future iterations of the MESES without the negative phrasing may alter the fit and provide a better measurement. Results of Analysis 3 demonstrate that the moral facet items are in particular need of revision.

Overall, the majority of the items on the MESES fit the model and measured the latent trait effectively. The six misfit items, *s1*, *e2*, *e4*, *m2*, *m4*, and *m5*, will need revision, or possibly need to be removed, in future iterations of the MESES.

3. To what degree do the item difficulty levels from items of specific holistic education facets compare to the theoretical item difficulty levels of another holistic model, specifically, Maslow's hierarchy of needs?

To answer research question 3, the items and their respective item difficulty levels needed to be placed in order, and this order then compared to the theoretical order of

these items according to Maslow's hierarchy of needs. In Maslow's hierarchy, moral facet items should be the most challenging to endorse, with emotional facet items being easier to endorse than the moral facet items, followed by social facet items being easier to endorse than moral and emotional facet items, and finally, cognitive facet items being the easiest to endorse. Cognitive facet items should be the easiest to endorse, as the cognitive facet falls outside Maslow's theory, and is considered the general purpose of schooling. If the items of the MESES with their respective facets have an item difficulty order that is similar to Maslow's hierarchy of needs, it provides evidence to support the validation of the MESES as an accurate measure of the presence of the holistic education latent trait. The results of Analysis 3 suggest that the MESES is effective at measuring the social and emotional facets. However, it will need revision of the cognitive and moral facet items in future iterations.

The results of Analysis 3 show that the item difficulty estimates for the MESES are between -1.13 and 2.88. As the Wright map in Figure 8 shows, the most challenging item to endorse was item *s1* at 2.88, followed by item *m1* at 1.46. Although the most challenging item to endorse is in the social facet, it is also an item with a high level of misfit, and thus, is not considered to measure the latent trait well. The moral facet items, that should be the hardest to endorse, are spread throughout the model, and three items have an issue with misfit. Two of the moral facet items that do fit the model are among the more difficult items to endorse, *m1* and *m3*. However, item *m6*, ... *discuss their ideas*, has an item difficulty level of -0.58, meaning the item is among the easier items to endorse and falls outside the theoretical item difficulty level. Item *m6* may have been easy to endorse if the school encourages students to discuss ideas in courses openly, or

possibly requires it as part of class participation. Item *m6* could be made more challenging by asking if students discuss their beliefs and values, as it may be a closer connection to the moral facet.

The emotional facet items and the social facet items that fit the model do appear to fall in the theoretical order according to Maslow's hierarchy of needs. Emotional facet items *e1*, *e3*, *e5*, and *e6*, fall slightly above and below the 0.0 logit mark, with difficulty levels that range between -0.52 and 0.47. Social facet items *s2*, *s3*, *s4*, and *s5* have difficulty levels that range between -1.13 and -0.73. These results suggest that the emotional and social facet items that demonstrated appropriate fit do match the theoretical model of Maslow's hierarchy of needs.

The cognitive items should have theoretically been the easiest items to endorse. However, these items did not match the model well. The item difficulty levels for the cognitive items, all of which fit the model, ranged between -0.87 and 0.99. The easiest cognitive item to endorse was item *c3* at -0.87. Besides item *c1*, the cognitive facet items will need a great deal of revision to better match Maslow's hierarchy of needs. The cognitive facet items may not fit the theoretical model due to the sample coming from a single school, where students may find the teaching of certain subjects more challenging. Piloting a revised MESES with a larger sample from several schools may result in the cognitive items falling into the theoretically appropriate order.

The results of Analysis 3 demonstrate that many of the MESES items do match the theoretical model of Maslow's hierarchy of needs. The emotional facet and social facet items that fit the model followed the theoretical model particularly well. The moral facet items partially follow the model, however, several are in need of revision. The

cognitive facet items, however, do not match the model effectively. Overall, the MESES has many items that meet the proposed theoretical model, but additional revision is needed with continued pilot testing of the tool.

4. Are there differences in how people from different groups, with similar levels of the latent trait, are able to respond to items?

To determine if there are differences in the ability of students and parents to respond to MESES items for research question 4, the differential item functioning (DIF) results for Analysis 3 were examined. The DIF results for Analysis 3 show that there was no significant difference between the responses of students and parents on any of the MESES items, as there was no t -statistic above 1.96, and thus, no statistically significant p -value. These results suggest that the MESES and its items do not have an issue with bias between students and parents, and the ability of members of these groups to endorse the items.

There are several possible reasons as to why no DIF was detected. It may be that the MESES items were written well enough that they could be interpreted in a similar manner by both groups of respondents. For the items that demonstrated misfit, respondents from both groups may have had similar issues with their ability to respond. It is also possible that the respondents share an affinity for the school, and may have chosen to respond to items in a similar positive manner. This possibility of a shared affinity is worth considering, given that there were just as many parent respondents as student respondents and this level of parent response rate is a somewhat uncommon occurrence in school-based research (Schilpzand, Sciberras, Efron, Anderson, & Nicholson, 2015).

The frequency of parental responses may suggest that parents have a particular interest in the school. Finally, it is possible that the sample size may have impacted the ability to detect DIF, and that DIF may become identifiable once the MESES is piloted with a larger sample. Although these are possibilities for why DIF was not detected with the MESES items in this analysis, additional research should be conducted with a larger sample and with respondents from other groups to further examine the possibility of the presence of DIF.

Limitations

Although this study was designed to include a thoughtful survey development process and validation technique that addressed many potential issues, there are limitations present in the study. The primary limitation pertains to the sample sizes of both the July 2015 and October 2015 samples. Both samples included respondent totals below recommended sample sizes (Linacre, 1994). These smaller sample sizes possibly led to inflated standard errors with the item estimates (Linacre, 2015a). Consequently, the item anchoring process using these item estimates with inflated standard errors arguably altered the interpretability of the estimates developed from the item anchoring process, as they are based on estimates with inflated standard errors. Although, the examination of item drift aids in the interpretability of estimates using item anchoring, basing anchoring on items with inflated standard errors may be conceived as an issue. Additional studies would benefit from including larger samples and aiming to have smaller standard errors.

Additional limitations of this study include not having Montessori Erdkinder stakeholders as a more intricate part of the validation procedure, as well as using sources that were not from Montessori herself as part of the survey development process.

Montessori Erdkinder stakeholders could provide a great deal of support for the inclusion or exclusion of items. Conducting qualitative interviews with Montessori Erdkinder stakeholders would provide great insight into quality of the tool and possible issues for school implementation, and thus, is a worthwhile pursuit for future studies. The inclusion of non-Montessori sources as part of the survey development process was necessary given the lack of detail from Montessori herself, and supports the tool being relevant to the values of the current Montessori community. However, the inclusion of these non-Montessori sources could be perceived as a deviation from true Erdkinder. Until additional Montessori Erdkinder original materials can be located or translated from the native Italian, little can be done to address this concern.

Conclusion

Alternative education is growing in the United States, as families continue to seek out schools with holistic education approaches over traditional public schooling. In response to this interest from families, a wide range of schools ascribing to unique and varied holistic education philosophies has emerged. Despite this development, there has been a lack in the construction and validation of evaluation tools for these unique school settings. Given that many of these new schools desire to have holistic education outcomes and have philosophies designed to be very different from traditional public schooling, it is important that evaluation tools are developed to examine the presence of their intended holistic school outcomes. This is particularly true for the new holistically-focused charter schools, which are required in some states to report evaluation data and undergo evaluation procedures similar to traditional public schools (Fryer, 2012; Jordan, 2013). This study demonstrates how a holistic education school evaluation tool can be created

and placed through a validation procedure for usage in Montessori high schools and the range of schools concerned with holistic education.

This study is the first step in a larger validation process. These preliminary study results indicate that the MESES shows great promise to be an effective holistic education school evaluation tool. However, it has potential for improvement through the revision of several items. Once the tool is revised, it should undergo additional validation with a larger Montessori high school sample from several schools. The validation process should also be continued by pilot testing the MESES with non-Montessori high schools that ascribe to holistic education virtues. Validating the tool in non-Montessori holistic education schools gives additional credence to using the MESES in many types of holistic education environments. This validation process will lead to the development of a holistic education school evaluation tool that can be confidently implemented by administrators and policy makers in school settings, and the subsequent data collected used for school reporting, accreditation, evaluation, and many other purposes.

This study was driven by the desire to construct an evaluation tool for holistic education schools that not only value working on the cognitive aspects of students' development, but also value the social, emotional, and moral aspects. This study serves as a great reminder that schools have the ability to influence students beyond just the cognitive capacity, and they can actively work to influence students in multiple domains. This ability for schools to influence students beyond the cognitive area is not a new idea. Students in American Catholic schools have been undergoing holistic education for hundreds of years (Bryk, Lee, & Holland, 1993; Greeley, 1982; Hoffer, 2000; Hunt, 2005). Catholic schools have always been concerned with outcomes beyond just the

cognitive, as students are routinely required to participate in social, athletic, and spiritual practices that are not designed to serve as cognitive development-specific experiences. With these varied experiences, students in Catholic schools have consistently showed high levels of high school achievement and college enrollment. As these positive achievement outcomes suggest, influencing students in areas beyond the cognitive may be a worthwhile endeavor for schools. This study is a reminder that schools can influence students in areas beyond the cognitive.

Perhaps the greatest strength of this study and the MESES is its implications for non-Montessori holistic education school settings. A great deal of research has gone into examining the outcomes of students attending schools with a holistic educational focus, often times by examining Catholic schools, with research suggesting there are benefits to students who attend these types of institutions (Hoffer, 2000; R. Miller, 1990). For example, students in Catholic school settings have consistently high achievement and college enrollment rates, with a particular benefit for minority students (Greeley, 1982; Hoffer, 2000). The challenge is that there has been a noticeable lack of evaluation tools that identify and examine the presence of holistic education, without which these schools cannot evaluate themselves for strengths and weaknesses. Consequently, without such tools, the ability to develop complete evaluation or accountability systems has suffered. Although the MESES provides a format for schools to develop their own unique holistic education tool that addresses their specific holistic education concerns, the MESES can be implemented in non-Montessori holistic education school settings. Holistic education schools, such as Catholic schools and other religious schools mentioned before, could implement the MESES to begin examining the presence of holistic education student

outcomes in their schools, and as the framework for developing a complete holistic education evaluation or accountability system.

Furthermore, this study has great implications for the traditional public education sector. In addition to cognitive outcomes, holistic outcomes are sometimes examined in the traditional public education sector for school reporting and evaluation purposes. For example, the Class Assessment Scoring System used as part of the Head Start school evaluation system includes an emotional support evaluation domain (United States Department of Health and Human Services, Office of Head Start, 2013; United States Department of Health and Human Services, Office of Head Start, 2014; Vitiello & Hadden, 2014). Administrators and policymakers in the traditional public education sector who are interested in holistic education outcomes will be able to confidently use the MESES, as will schools required to participate in evaluation systems that include holistic education outcomes, knowing that it has been examined for validity issues as well as constructed in a transparent process that can be explained to stakeholders. The MESES could be extremely valuable for the traditional public education sector as after additional validation work it can be used as a turnkey resource, ultimately saving an organization the cost of constructing and validating such a resource.

This study and the MESES have wider implications for educational policy given the non-renewal of No Child Left Beyond (NCLB) and the passing of the Every Student Succeeds Act (ESSA) in December 2015 (United States Department of Education, 2016). ESSA transfers a great deal of educational accountability power back to individual states. States are currently in the process of determining ways to create alternative accountability models that require schools have high standards, but also considers additional factors

important for determining school success. Holistic education outcomes could easily become a part of a state's accountability model, particularly given the concern that, under NCLB, schools were unconcerned with these types of outcomes. If any state were to consider holistic education as part of their accountability model, a revised MESES could be implemented as part of a school evaluation, or the process used to create the MESES could be implemented to develop a tool specific to the holistic education concerns of a state. Again, the revised MESES could be used as a turnkey resource that could confidently be implemented in such an accountability system, saving states the cost of having to construct and validate such a tool.

This study has the potential to serve as a great catalyst for the development of a wave of holistic education schooling evaluation and accountability tools. Holistic education concerns were often overlooked as educational outcomes because the challenging nature of quantifying the principles. This study, however, shows that the development of such holistic education tools is possible, and demonstrates how it can be done in a way that respects the diversity of these school settings. The MESES was developed in a way that acknowledges the uniqueness of Montessori Erdkinder philosophy; however, it retains its focus on holistic education. By developing the MESES in this way, it is accessible and usable in any school concerned with holistic education student outcomes. This study and the creation of the MESES has the potential to greatly alter educational evaluation by providing a practical means to examine holistic education student outcomes, and laying the framework for larger evaluation and accountability systems.

Appendix A

Montessori Erdkinder School Evaluation Survey (MESES)

The purpose of the MESES is to evaluate the performance of a Montessori high school. This survey was created to collect feedback from Montessori high school students, parents/guardians, and staff. When filling in this survey, please consider the Montessori high school with which you are affiliated. Your feedback is important! Thank you in advance for your participation.

Section I – School Perceptions

In section I, read the following statements and rate how strongly you agree or disagree with each statement. For each statement, please select one response using the following scale:

1	2	3	4	n/a
<i>Strongly Disagree</i>	<i>Disagree</i>	<i>Agree</i>	<i>Strongly Agree</i>	<i>Not Applicable</i>

1. Students at the Montessori high school ...

... are taught to challenge the rules of society.	1	2	3	4	n/a
... feel they should help others in society.	1	2	3	4	n/a
... learn that working is important to their future success.	1	2	3	4	n/a
... volunteer in the community.	1	2	3	4	n/a
... are encouraged to work with each other.	1	2	3	4	n/a
... get to choose their schoolwork.	1	2	3	4	n/a
... respect one another.	1	2	3	4	n/a
... set goals to improve society.	1	2	3	4	n/a
... are taught to address the problems in society.	1	2	3	4	n/a
... feel that working is shameful.	1	2	3	4	n/a
... discuss their ideas.	1	2	3	4	n/a
... are supported in building self-confidence.	1	2	3	4	n/a
... are taught problems in society are fixable.	1	2	3	4	n/a
... find enjoyment in working.	1	2	3	4	n/a
... feel their future has already been decided.	1	2	3	4	n/a
... are aware of their own talents.	1	2	3	4	n/a
... learn responsibility.	1	2	3	4	n/a
... are taught how to find structure in nature.	1	2	3	4	n/a
... are taught how math applies to their life.	1	2	3	4	n/a
... are taught about other cultures through foreign language instruction.	1	2	3	4	n/a
... learn how society has changed over time.	1	2	3	4	n/a
... are encouraged to philosophically reflect on the universe.	1	2	3	4	n/a
... express themselves through creative arts.	1	2	3	4	n/a

Section II – Strengths and Weaknesses

In section II, please write a brief response to the following questions.

2. What are the strengths of this school?

3. What are the weaknesses of this school?

4. What suggestions do you have for other questions that can be included on this survey?

Section III - Demographics

In section II, please answer the following questions by checking one response category.

5. I am ...

Male

Female

6. I am a ...

Student

Parent/guardian

Teacher

Staff member

Appendix B

Item Development Matrix

Kahn (2011) Outcome Category	Kahn (2011) Objectives	Purpose	Item and Label (In parentheses)	Reverse-Coded Item
	(pp. 24-25)			
Cognitive	Ability to analyze scientific causality in the natural world and the cosmos.	Kahn's (2011) objective on scientific causality seems to be rooted in Montessori's strong focus on scientific learning. Montessori in her 1973 work includes chapters on scientific experiments and educational methods, much of which is centered on teaching the causality of biological and chemical processes. The purpose of this item is to examine if there is a perception that students are learning the causality of natural processes with which Montessori was concerned.	Are taught how to find structure in nature. (c1)	No

Cognitive	Increased understanding of mathematics directly connected to the practical needs of the farm environment and the symbolizing of scientifically observed data.	Montessori (1973) left somewhat clearer expectations regarding actual expectations of how adolescents should experience learning certain subjects. For example, Montessori (1973) writes, "Mathematics: Human intelligence today is no longer a natural intelligence but a mathematical intelligence. Without a mathematical education it is impossible to understand the progress of our time or to participate in it... It is necessary then, because of the vital importance of mathematics, that the school employ 'special methods' in its teaching of the subject and that it render the individual concepts clear and understandable by the help of concrete examples" (p. 119). In her quote, Montessori is arguing for the usage of practical examples. Kahn's (2011) objective also applies to practical examples, however, he places it in the context of the Montessori farm community. This item removes the aspect of the farm community, and examines if there is a perception that students are learning how to apply math in a concrete manner, specifically through the application to practical life.	Are taught how math applies to their life. (c2)	No
-----------	---	--	---	----

Cognitive	Increased facility in a variety of languages and the ability to use language to penetrate different cultures and improve human understanding.	Montessori (1973) writes, "Languages: The development of language is part of the personality itself. Words are, in effect, the natural means to express an idea and consequently to establish understanding standing between men" (p. 120). Montessori argues that foreign languages should serve as a means of connecting adolescents to other cultures, rather than just for the purposes of communication. Kahn's (2011) directly replicates this idea. The purpose of this item is to examine if there is a perception that students are learning foreign languages with the intent of connecting to other cultures.	Are taught about other cultures through foreign language instruction. (c3)	No
-----------	---	--	--	----

Cognitive	Ability to connect the history of life and its civilizations with principles of the evolving self as well as the social evolution of a human community.	Montessori (1973) writes, "Another facet of history particularly suited to the following period is that which treats human development in relation to geographic events: contacts and cross-breedings among the different peoples, the assimilation of different races and cultures, the wars and the conquests of empires - all accompanied by an examination of feelings and customs, of the influence of religion and patriotic sentiment, and the behavior of man" (p. 123). Montessori's believed that history should be largely about how people have changed over time and civilization has formed over time through community interactions. Kahn's (2011) objective is supported by Montessori's quote, as his objective requires students to be able to understand how civilization has shaped over time to make the present society. The item examines if there is a perception that students are being taught to examine that society has been morphed over time within the specific context of a history course.	Learn how society has changed over time. (c4)	No
-----------	---	--	---	----

Cognitive	Philosophical contemplation of nature and cosmos.	Montessori (1973) writes, "The observation of nature is not only an enrichment of the mind from the philosophical and scientific points of view" (p. 106), which suggests that Montessori connected philosophical reflection to nature. This is the core of Kahn's (2011) objective. The purpose of this item is to examine if there is a perception that students are reflecting on the universe. The universe was selected as a means of representing nature and the cosmos, as it keeps the item from being focused singularly on the environment or a particular interpretation of cosmos.	Are encouraged to philosophically reflect on the universe. (c5)	No
Cognitive	Integration of personal expression within a variety of artistic, speaking, musical, and media modalities in direct relation to occupations and role development within the community.	Montessori (1973) argues that one of the primary objectives of an Erdkinder school is "To open the way to the possibilities of the adolescent for personal expression:...art tasks...are only intended to facilitate the expression of personal artistic feeling" (p. 118). Montessori's quote relates to her philosophy on the importance of personal expression in child/adolescent development. The quote supports Kahn's (2011) objective, as the objective is centered on art as a means of demonstrating the adolescent's personal expression. The purpose of this item is to examine if there is a perception that students are given the opportunity to personally express themselves through an artistic format.	Express themselves through creative arts. (c6)	No

Social	Knowledge of the meaning of rules and their importance to harmonious living.	Tornar (2011) indicates that personality development is the goal of Montessori secondary schooling and the secondary school environment must provide "the opportunity of ... being independent in the process of decision making; choosing independently but respecting the socially shared rules" (p. 117). Kahn's (2011) objective is directly connected to the idea of understanding and appreciating the purposes of rules, making it a valid objective. The item establishes if there is a perception that the school teaches the importance of rules in society.	Are taught to challenge the rules of society. (s1)	Yes
Social	Adaptation to a variety of work demands for the sake of others; the beginning of social consciousness.	Montessori (1973) argues that adolescents' ability to adapt is vital for their futures because of the uncertainty of the world. She specifically states, "We must develop the possibility of supple and live adaptation without rigid specialization. In the fierce battle which social life has become, man needs a strong character and a fast mind in addition to his courage. He needs to reinforce these principles by moral training and be endowed with practical capabilities to face up to the difficulties of life" (pp. 97-98). Kahn's (2011) objective interprets this concept for the purposes of addressing social needs rather than only employment. This item is concerned with the concept of student adaptability and establishes if there is a perception that the school is having students adapt their skills to various challenging situations.	Feel they should help others in society. (s2)	No

Social	Understanding of work as a producer of commerce necessary to community, life, leading to a beginning view of economic independence and interdependence.	Economic independence is a principle of an Erdkinder education which Montessori spent a great deal of time discussing in the 1973 work. Montessori (1973) states "The essential (school) reform therefore consist in putting the adolescent in condition to be able to acquire economic independence" (p. 102). Kahn's (2011) objective focuses on work as the means of leading to this economic independence. This item establishes if there is a perception that students are learning the importance of work and connecting it to their own lives and future endeavors.	Learn that working is important to their future success. (s3)	No
Social	Balancing of individual initiatives in relation to community goal.	Montessori (1973) states, "After twelve years, we must develop in the child a feeling of society, which ought to contribute to more understanding among men and, as a result, more love. Let us develop admiration and understanding for work and for life of man to this end. ...We have the child participate in social work of some kind" (p. 96). Montessori in this quote is connecting the idea of individual work to participation in society. Kahn's (2011) objective is directly supported by this idea, as it establishes the principle of individuals working towards a common outcome to better the community. This item establishes if there is a perception that the school is having students participate in work which addresses community needs.	Volunteer in the community. (s4)	No

Social	Understanding of interdependency and the need to cooperate with adults and peers in relation to the rest of the world.	Montessori (2011a) writes that, "At puberty another form of intercourse arises, a new kind of social instinct: society is conceived of as work. New real work, work of absorbing interests, work that is worth the effort is not in the main creative, hence not individual; it requires the collaboration of others, and this cooperation implies association and discipline" (pp. 59-60. This quote demonstrates that Montessori believed in encouraging adolescents to work with one another. Kahn's (2011) objective directly relates to encouraging the need for students to learn the importance of working together. This item establishes if there is a perception that a school has students working with others, which is the practical application of understanding cooperation with others.	Are encouraged to work with each other. (s5)	No
Emotional	Feelings of usefulness and an understanding of one's 'many sided powers of adaptation.	Tornar (2011) states that in the Montessori philosophy, adolescents need "to enhance self confidence" (p. 117). Kahn's objective of "feelings of usefulness" can be interpreted as a self-confidence objective, which as an Erdkinder principle, would be supported by Tornar's statement. This item examines if there is a perception that a school is building students' self-confidence.	Are supported in building self-confidence. (e1)	No

Emotional	Belief in human capacity to solve problems and in the spiritual source of strength to overcome adversity.	Montessori (2011a) states, "Two 'faiths' can raise up the man: faith in God and faith in himself. The two faiths must coexist: The former in the interior life of the man; the latter concerns the social man" (p. 59). This quote supports Kahn's (2011) objective as it supports his argument for belief in self and in a spiritual source. This item focuses on the belief in self and asks if students are perceived as being encouraged to have that confidence to make an impact.	Are taught problems in society are fixable. (e2)	No
Emotional	Inner harmony and happiness due to personal contribution, love of work, love of study and achievement, and a personal participation in the work of society.	Montessori (2011c) discusses how adolescents develop an "abnormal life because they lead their lives in sharp contrast to their real desires and ambitions" (p. 77) and argues this a result of adolescents not being able to participate in fields of their own interest. Kahn's (2011) objective is supported by the statement as it shows Montessori believed in the need for students to complete work they were enjoying. This item examines if there is a perception that students are enjoying their work which would demonstrate inner harmony and happiness.	Find enjoyment in working. (e3)	No

Emotional	Feeling that one can be in control of change, internally and externally, in one's personal and social evolution.	Montessori (1973) states that adolescence "implies the acquiring of a sense of the power to act alone; the possibility of carrying out some useful and important action without help from others; the being able to solve one's problems for oneself, to reach a difficult goal of one's own efforts" (p. 58). This statement demonstrates that Montessori believed in adolescents having a sense of being able to control their lives, which is the central tenant of this Kahn (2011) objective. This item examines if there is a perception that students are being taught the idea that they can shape their own futures.	Feel their future has already been decided. (e4)	Yes
Emotional	Revelation of the innermost 'vocation' or deep calling of the soul; a sense of mission or commitment to one's work and life.	Montessori (1973) argues that for adolescents, "success depends on self-confidence, on the awareness of one's own talents and of the many possibilities in their adaptation" (p. 102). This quote supports Kahn's (2011) objective by demonstrating that Montessori believed in students having an awareness of their own talents. The item focuses on the work aspect, not the cosmic education component. The item examines if there is a perception that students are aware of their talents, which is connected to both Kahn's vocation and Montessori's awareness.	Are aware of their own talents. (e5)	No

Emotional	Feeling of self-sufficiency, of taking care of self and others; the feeling of self-confidence.	Montessori (2011a) states that in adolescence, "this man is now ready to begin to live mysteriously, yet positively for other people, to dedicate his life to others, to make sacrifices, to give his life to protect" (p. 74). This quote supports the idea of responsibility and living for others that is found in Kahn's (2011) objective. This item examines specifically the responsibility component, by examining if there is a perception that students are learning responsibility.	Learn responsibility. (<i>e6</i>)	No
Moral	Individual initiative; commitment to freely chosen work.	Montessori (2011a) writes, "Thus, at all the various and successive stages of development, independence is a valid guide for education" (p. 58). This quote supports Kahn's (2011) objective of schools the need for individually selected work. The item examines if there is a perception that students are given the opportunity to select their own work.	Get to choose their schoolwork. (<i>m1</i>)	No
Moral	Respect for others and their roles.	Montessori (1973) states, "Let us develop admiration and understanding for work, and for the life of man to this end" (p. 96). This phrase is supportive of Kahn's (2011) objective of learning to respect others. Montessori in her 1973 does not specifically state respect for others, but the theme is present in statements like the previous quote. This item establishes the perception that the school encourages students to respect others.	Respect one another. (<i>m2</i>)	No

Moral	Development of a mission orientation and service to the universal needs of a larger humanity.	Montessori (2011c) states, "It is his destiny to take part, bit by bit, in the organization of society. We can predict that he will live for others in this world, ready to make sacrifices for others" (p. 74). This statement supports Kahn's (2011) objective, as both statements suggest the adolescent will commit aspects of their lives to the larger society/humanity/world, instead of focusing only on the individual. This item establishes the perception that adolescents are being encouraged to reflect on how they will impact society in the future.	Set goals to improve society. (m3)	No
Moral	Ability to grapple with social and moral problems, such as the right use of the natural environment or the ethics of science.	Montessori (1973) states, "Laborers today need education. They need to understand the complex problems of our time. Otherwise the role that their work plays in the plan of society will be nothing else than that of a pair of ignorant hands" (p. 98). This statement supports Kahn's (2011) objective, as both are concerned with the understanding of societal issues. Montessori's comment takes Kahn's objective further by addressing this need in the context of the worker (whom, in her example, is associated with the developing adolescent). The purpose of this item focuses on the connected theme of addressing societal problems. Specifically, this item addresses if there is a perception that the school is having students examine societal issues.	Are taught to address the problems in society. (m4)	No

Moral	A sense that work is noble and assumption of adult-like responsibility.	Montessori (1973) states, "All work is noble...it is essential to understand the value of work in all its forms, be they manual or intellectual" (p. 103). This quote directly supports Kahn's (2011) objective, as they are both concerned with the perception of work as noble. Kahn pushes this objective to include the act of working. The item is focused on the perception students are gaining an appreciation for work.	Feel that working is shameful. (m5)	Yes
Moral	Conscience exercised by community values and responsible dialogue.	Tornar (2011) indicates that personality development is the goal of Montessori secondary schooling and the secondary school environment must provide "the opportunity of ... exploring the world, thinking, expressing and debating their own and others' ideas" (p. 117). This quote supports the responsible dialogue aspect of the Kahn (2011) objective. The item examines if there is a perception that schools are providing students opportunities to discuss their ideas.	Discuss their ideas. (m6)	No

REFERENCES

- American Montessori Society. (2015). *Montessori schools*. Retrieved from <http://amshq.org/Montessori-Education/Introduction-to-Montessori/Montessori-Schools>
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*, 561-573.
- Barker, D. (2011). A historical look at Montessori's Erdkinder. *Communications*, *1-2*, 96-112.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397-479). Reading, MA: Addison-Wesley.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. New York, NY: Routledge.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: The Guilford Press.
- Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.
- Chubb, J. E., & Moe, T. M. (1990). *Politics, markets, & America's schools*. Washington, DC: The Brookings Institution.
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: The Guilford Press.
- DeVellis, R. F. (2012). *Scale development: Theory and applications* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

- Engelhardt, G., Jr. (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences*. New York, NY: Routledge.
- Finn, J., Caldwell, K., & Raub, T. (2006). Why parents choose charter schools for their children with disabilities. *Journal of Educational Research & Policy Studies*, 6, 91-110.
- Fisher, W. P. (2006). Survey design recommendations. *Rasch measurement transactions*, 20, 1072-1074.
- Forbes, S. G., & Martin, R. A. (2014, April). *What holistic education claims about itself: An analysis of holistic schools' literature*. Paper presented at the 2014 American Educational Research Association's annual conference, San Diego, CA.
- Fryer, R. G., Jr. (2012). *Learning from the successes and failures of charter schools* (Discussion paper 2012-06). Retrieved from The Hamilton Project website http://www.hamiltonproject.org/papers/learning_from_the_successes_and_failures_of_charter_schools/
- Goldstein, H., & Wood, R. (1989). Five decades of item response modelling. *British Journal of Mathematical and Statistical Psychology*, 42, 139-167.
- Grazzini, C. (2004). The four planes of development. *The NAMTA Journal*, 29, 27-61.
- Greeley, A. M. (1982). *Catholic High Schools and Minority Students*. New Brunswick: Transaction Books.
- Gutek, G. L. (2004). Introduction: A biography of Montessori and an analysis of the Montessori method. In G. K. Gutek (Ed.), *The Montessori method* (pp. 1-66). Lanham, MD: Rowman & Littlefield Publishers, Inc.

- Hershey Montessori School Adolescent Community. (2015). *Occupation planning: Objectives for evaluation*. Retrieved from the North American Montessori Teachers' Association website:
<http://www.montessori-namta.org/Curriculum-Downloads>
- Hoffer, T. B. (2000). Catholic school attendance and student achievement: A review of the extension of research. In J. Youniss & J. J. Convey (Eds.). *Catholic schools at the crossroads: Survival and transformation* (pp. 87-112). New York, NY: Teachers College Press.
- Hunt, T. C. (2005). Catholic schools: Yesterday, today, and tomorrow. *Journal of Research on Christian Education, 14*, 161-175.
- Hursh, D. W. (2006). Marketing education: The rise of standardized testing, accountability, competition, and markets in public education. In E. W. Ross & R. Gibson (Eds.), *Neoliberalism and education reform* (pp. 15-34). Cresskill, NJ: Hampton Press, Inc.
- Jordan, G. (2013, February 18). *Charter schools say their teachers will be evaluated the same as district teachers*. Retrieved from the National Public Radio website
<https://stateimpact.npr.org/florida/2013/02/18/charter-schools-say-their-teachers-will-be-evaluated-the-same-as-district-teachers/>
- Kahn, D. (2011). Eight pictures at an exhibition: A Montessori retrospective on the discovery of the adolescent. *Communications, 1-2*, 15-41.
- Kahn, D., & Pendleton, D. R. (2007). *The whole-school Montessori handbook*. Burton, OH: NAMTA.

- Kena, G., Aud, S., Johnson, F., Wang, X., Zhang, J., Rathbun, A., Wilkinson-Flicker, S., & Kristapovich, P. (2014). *The Condition of Education 2014* (NCES 2014-083). Retrieved from the U.S. Department of Education, National Center for Education Statistics website: <http://nces.ed.gov/pubsearch>.
- Kline, P. (1998). *The new psychometrics: Science, psychology, and measurement*. New York, NY: Routledge.
- Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions*, 7, 328.
- Linacre, J. M. (1998). Structure in Rasch residuals: Why principal components analysis (PCA)? *Rasch Measurement Transactions*, 12, 636.
- Linacre, J. M. (2000). Comparing “partial credit models” (PCM) and “Rating Scale Model” (RSM). *Rasch Measurement Transactions*, 13, 768.
- Linacre, J. M. (2012a). *Winsteps Rasch tutorial 2*. Retrieved from the Winsteps website: <http://www.winsteps.com/tutorials.htm>
- Linacre, J. M. (2012b). *Winsteps Rasch tutorial 3*. Retrieved from the Winsteps website: <http://www.winsteps.com/tutorials.htm>
- Linacre, J. M. (2015a). *A user's guide to Winsteps Ministep Rasch-model computer programs*. Retrieved from <http://www.winsteps.com/manuals.htm>
- Linacre, J. M. (2015b). *Ministep Version 3.80.1* [Computer software]. Retrieved from <http://www.winsteps.com/ministep.htm>
- Lipman, P. (2011). *The new political economy of urban education: Neoliberalism, race, and the right to the city*. New York, NY: Routledge.
- Ludwig, H. (2011). Editorial. *Communications*, 1-2, 1-14.

- Martin, R. A. (2002, April). *Alternatives in education: An exploration of learner-centered, progressive, and education*. Paper presented at the American Educational Research Association annual meeting, New Orleans, LA.
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50, 370-396.
- Maslow, A. H. (1987). *Motivation and personality*. New York, NY: Longman.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.
- Mayes, C. & Williams, E. (2013). *Nurturing the whole child*. Lanham, MD: Rowman & Littlefield Education.
- Miller, J. P. (2010). *Whole child education*. Toronto, ON: University of Toronto Press.
- Miller, R. (1990). *What are schools for? Holistic education in American culture*. Brandon, VT: Holistic Education Press.
- Montessori Foundation. (2015). *Frequently asked questions*. Retrieved from <http://www.montessori.org/FAQ>
- Montessori, M. (1973). *From childhood to adolescence: Including Erdkinder and the function of the university*. New York, NY: Schocken Books.
- Montessori, M. (2011a). Principles and practice in education. *Communications*, 11-12, 50-60.
- Montessori, M. (2011b). The physical and psychological development of the adolescent. *Communications*, 11-12, 67-72.
- Montessori, M. (2011c). The adolescent – a “social newborn.” *Communications*, 11-12, 73-78.

- Nardi, P. M. (2006). *Doing survey research: A guide to quantitative methods*. Boston, MA: Pearson Education, Inc.
- North American Montessori Teachers' Association (2015a). *Curriculum downloads*. Retrieved from <http://www.montessori-namta.org/Curriculum-Downloads>
- North American Montessori Teachers' Association. (2015b). *Montessori assessment outline*. Retrieved from <http://www.montessori-namta.org/Montessori-Assessment-Outline>
- North American Montessori Teachers' Association. (2016). *School directory*. Retrieved from <http://www.montessori-namta.org/School-Directory>
- O'Neill, Y., Peabody, M., Tan, R. J. B., & Du, Y. (2013). How much drift is too much?. *Rasch Measurement Transactions*, 27, 1423-1424.
- Ostini, R., & Nering, M. L. (2005). *Polytomous Item Response Theory Models* (Ch. 3.). Thousand Oaks, CA: Sage Publications, Inc. Retrieved from https://us.sagepub.com/sites/default/files/upm-binaries/5197_Ostini_Chapter_3.pdf
- Parker, D. E. (2007). *Navigating the social/cultural politics of school choice: "Why do parents choose Montessori?" A case study* (A doctoral dissertation). Retrieved from <http://libres.uncg.edu/ir/>
- Pottish-Lewis, P. (2013). Standardized tests: Help or hindrance. *AMI/USA Journal*, Winter, 3-8.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: University of Chicago Press. (Reprinted from 1960).

- Rathunde, K. (2001). Montessori education and optimal experience: A framework for new research. *The NAMTA Journal*, 26, 11-43.
- Royal, K. D., & Elahi, F. (2011). Psychometric properties of the Death Anxiety Scale (DAS) among terminally ill cancer patients. *Journal of Psychosocial Oncology*, 29, 359-371.
- Schilpzand, E. J., Sciberras, E., Efron, D., Anderson, V., & Nicholson, J. M. (2015). Improving survey response rates from parents in school-based research using a multi-level approach. *PLoS One*, 10, e0126950.
- Sick, J. (2010). Assumptions and requirements of Rasch measurement. *SHIKEN: JALT Testing & Evaluation SIG Newsletter*, 14, 23-29.
- Standing, E. M. (1998). *Maria Montessori: Her life and work*. New York, NY: PLUME.
- Taggart, G. (2001). Nurturing spirituality: A rationale for holistic education. *International Journal of Children's Spirituality*, 6, 325-339.
- Tornar, C. (2011). The secret of adolescence. *Communications*, 1-2, 113-120.
- Toland, M. D. (2014). Practical guide to conducting an item response theory analysis. *The Journal of Early Adolescence*, 34, 120-151.
- Tristan, A. (2006). An adjustment for sample size in DIF analysis. *Rasch Measurement Transactions*, 20, p. 1070-1071.
- United States Department of Education (2016). *Every Student Succeeds Act (ESSA)*. Retrieved from <http://www.ed.gov/essa>

United States Department of Health and Human Services, Office of Head Start. (2013).

Improving teacher-child interactions: Using the CLASS in Head Start preschool programs. Retrieved from

<http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/docs/using-the-class.pdf>

United States Department of Health and Human Services, Office of Head Start. (2014).

Head Start professional development Classroom Assessment Scoring System

(CLASS) – webcast [video file]. Retrieved from

<http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/eecd/Assessment/Classroom%20Assessment%20Scoring%20System%20%28CLASS%29/ClassroomAssessment.htm>

Vitiello, V. E., & Hadden, D. S. (2014). *CLASS system implementation guide.* Retrieved from the Teachstone Training website:

<http://info.teachstone.com/class-implementation-guide>

Weinberg, D. R. (2011). Montessori, Maslow, and self-actualization. *Montessori Life*, 23(4), 16-21.

Welch, C. J., & Miller, T. R. (1995). Differential item functioning in direct writing assessments: Problems and an example. *Journal of Educational Measurement*, 32, 163-178.

Willis, G. B. (1999). *Cognitive interviewing: A "how to" guide.* Paper presented at the 1999 Meeting of the American Statistical Association.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis.* Chicago, IL: MESA Press.

Wright, B. D., & Masters, G. N. (1990). Computation of OUTFIT and INFIT statistics.

Rasch Measurement Transactions, 3, 84-85. Retrieved from

<http://www.rasch.org/rmt/rmt34e.htm>

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago, IL: MESA Press.

Zarybnisky, E. M. (2010). *A ray of light: A mixed-methods approach to understanding*

why parents choose Montessori education (A doctoral dissertation). Retrieved

from <http://digitalcommons.unl.edu/>

VITA

Anthony Philip Setari

PLACE OF BIRTH Sunrise, Florida

EDUCATION

Master of Arts TEACHING – SECONDARY EDUCATION
The Johns Hopkins University, Dept. of Teaching and Learning
Baltimore, Maryland
May 2011

Bachelor of Arts WRITING & LINGUISTICS – CREATIVE WRITING
Georgia Southern University, Dept. of Writing & Linguistics
Statesboro, Georgia
May 2009

PROFESSIONAL EXPERIENCE

Graduate Research Assistant

Dept. of Educational Policy Studies & Evaluation – eLearning Innovation Initiative Grant
University of Kentucky
Lexington, KY
June 2015 – Present

Educational Assessment and Accountability Intern

Performance Assessment of Competency Education (PACE) Accountability System
National Center for the Improvement of Educational Assessment, Inc.
Dover, NH
June 2015 – August 2015

Accreditation and Evaluation Intern

Montessori High School of Kentucky
Lexington, KY
January 2015 – May 2015

Research Assistant

Kentucky Statistical Analysis Center
Kentucky Justice and Public Safety Cabinet
Frankfort, KY
August 2014 – May 2015

Graduate Teaching Assistant/Graduate Research Assistant

Dept. of Educational Policy Studies & Evaluation/College of Education Evaluation
Center
University of Kentucky
Lexington, KY
June 2014 – May 2015

Graduate Research Assistant
School of Human Environmental Sciences
University of Kentucky
Lexington, KY
August 2011 – May 2014

Teach for America – Language Arts Teacher
Arundel Elementary/Middle School #164
Baltimore City Public School System
Baltimore, MD
August 2009 – May 2011

PUBLICATIONS

Peer Reviewed Journal Articles

Setari, A. P., & Setari, R. R. (Accepted for publication). Trends in Catholic school minority enrollment and higher education entrance over the recession. *Journal of Catholic Education*.

Setari, A., & Zimmerman, J. N. (2013). TANF and SNAP participation fluctuations during the Great Recession and implications for Family and Consumer Sciences Cooperative Extension Services. *The Forum for Family and Consumer Issues (FFCI)*, 18 (2). On-line: <http://www.ncsu.edu/ffci/index.php>

Hosier, A. F., & **Setari, A.** (2013). Stand up to falling: A multi-component fall prevention Extension program. *The Forum for Family and Consumer Issues (FFCI)*, 18 (1). On-line: <http://www.ncsu.edu/ffci/index.php>

Manuscripts in Review

Setari, A. P., Lee, J., & Bradley, K. D. (Revise and resubmit). A psychometric approach to the validation of a student evaluation of teaching instrument.

Reynolds, K., **Setari, A. P., & Bradley, K. D.** Evaluating graduate student experiences: A walk through the creation of a higher education exit survey.

Professional Reports

Setari, A. P., & Marion, S. (2015). *New Hampshire's Performance Assessment for Competency Education (PACE): Theory of action and evaluation guide*. Report prepared for the National Center for the Improvement of Educational Assessment, Inc.

Evans, C. M., & **Setari, A. P.** (2015). *Feedback from New Hampshire school districts that implemented PACE in 2014-2015*. Policy brief prepared for the National Center for the Improvement of Educational Assessment, Inc.

Setari, A. P. (2015). *Montessori High School of Kentucky Student Data Document*. Report prepared for the Montessori High School of Kentucky.

- Setari, A. P., & Stanek, M.** (2015). *Rasch analysis of the Kentucky Department of Juvenile Justice's Intensive Aftercare selection instrument*. Report prepared for the Kentucky Dept. of Juvenile Justice.
- Setari, A.** (2014). *Statistical analyses of fire experience and risk data*. Report prepared for the University of Kentucky's Dept. of Curriculum and Instruction.
- Vail, A., Smith, C. K., Zimmerman, J. N., **Setari, A.**, & Manseau, R. (2013). *2014 Building strong families for Kentucky*. University of Kentucky, School of Human Environmental Sciences: Lexington, KY. On-line: <http://HES.UKY.EDU/StrongFamilies>
- Moore, N., Smith, C. K., & **Setari, A.** (2013). *Building strong families for Kentucky: Partners in food and health*. Briefing report prepared for the 2013 Kentucky Family Impact Seminar. On-line: <http://HES.UKY.EDU/FIS>
- Vail, A., Stephenson, L., Zimmerman, J. N., **Setari, A.**, & Manseau, R. (2012). *2013 Building strong families for Kentucky*. University of Kentucky, School of Human Environmental Sciences: Lexington, KY. On-line: <http://HES.UKY.EDU/StrongFamilies>
- Vail, A., Stephenson, L., Zimmerman, J., **Setari, A.**, & Davis, J. (2011). *2012 Building strong families for Kentucky*. University of Kentucky, School of Human Environmental Sciences: Lexington, KY. On-line: <http://HES.UKY.EDU/StrongFamilies>

PRESENTATIONS

International

- Setari, A.** (2013, June). *Economic indicators of expected older adult population growth*. Poster presented at the 20th International Association of Gerontology and Geriatrics World Congress of Gerontology and Geriatrics. Seoul, Korea.
- Hosier, A., & **Setari, A.** (2013, June). *Stand Up to Falling: Evaluating the impact of a fall prevention program on older adults*. Paper presented at the 20th International Association of Gerontology and Geriatrics World Congress of Gerontology and Geriatrics. Seoul, Korea.

National

- Setari, A. P., & Bradley, K. D.** *Rasch validation analysis of a Montessori high school teacher evaluation survey*. Poster presented at the 2016 American Educational Research Association's Annual Meeting for SIG 123: Survey research in education. Washington, DC.
- Setari, A. P., & Stanek, M.** (2015, November). *Rasch analysis of the Kentucky Department of Juvenile Justice's Intensive Aftercare selection instrument*. Poster presented at the Association for Public Policy Analysis and Management's 2015 Fall Research Conference. Miami, FL.
- Setari, A. P., Egger, J. O., & Bradley, K. D.** (2015, August). *Development of a project-based learning course evaluation tool using item response theory*. Poster presented at the American Psychological Association's 2015 Annual Convention. Toronto, ON.

Setari, A., & Setari, R. R. (2013, November). *Rasch analysis of a financial literacy measurement tool*. Poster presented at the 2013 Association for Financial Counseling and Planning Education Annual Research and Training Symposium. Greenville, SC.

Setari, A., Vail, A., Stephenson, L., Zimmerman, J., & Henken, K. (2013, June). *Connecting with policy makers: "Building strong families" profiles*. Educational session at the American Association of Family & Consumer Sciences 104th Annual Conference & Expo. Houston, TX.

Setari, A. P., & Zimmerman, J. N. (2012, July). *Trends in public assistance usage and unemployment during the current recession: An exploratory look at Kentucky*. Paper presented at the 75th Annual Meeting of the Rural Sociological Society. Chicago, IL.

Regional

Setari, A. P., Bridenbaugh, S., & Bradley, K. D. (2016, February). *Validation of a college preparedness scale for academic enhancement program evaluation*. Paper presented at the 39th Annual Conference of the Eastern Educational Research Association. Hilton Head Island, SC.

Setari, A. P., & Bradley, K. D. (2016, February). *Continuing the validation process of examining a Montessori high school student evaluation of teaching tool*. Poster presented at the 39th Annual Conference of the Eastern Education Research Association. Hilton Head Island, SC.

Setari, A. P. (2015, October). *Developing a Montessori high school teacher evaluation: Challenges and possibilities*. Paper presented at the 2015 Midwestern Educational Research Association's Annual Conference. Evanston, IL.

Setari, A. P., Egger, J. O., & Bradley, K. D. (2015, October). *Rasch validation analysis of the reduced Integrated Music Interest Survey for project-based learning courses*. PowerPoint presented at the 2015 Ohio River Valley Objective Measurement Seminar. Lexington, KY.

AWARDS

Conference

Graduate Student Award (2016) – *American Educational Research Association, SIG 123: Survey Research in Education*

Three Minute Thesis Best Paper Award (2016) – *Eastern Educational Research Association*

Three Minute Thesis People's Choice Award (2016) – *Eastern Educational Research Association*

National

Segal AmeriCorps Education Award (2011) - *AmeriCorps*

Baltimore Corps Member (2009-2011) – *Teach for America*

U.S. Presidential Volunteer Service Award (2007) – *Corporation for National & Community Service*

University

John Edwin Partington and Gwendolyn Gray Partington Scholarship (2015-Present) - *University of Kentucky, College of Education*

Lyman T. Johnson Fellowship (2012-2014) - *University of Kentucky, The Graduate School*

U.K. Graduate School Tuition Scholarship for Teach for America Alumni (2011-2012) –
University of Kentucky, The Graduate School

University Honors Program Scholarship (2006-2009) – *Georgia Southern University, University Honors Program*

Anthony Philip Setari