PSYCHOMETRIC CHARACTERISTICS AND SOCIALVALIDITY

OF THE 3M PRESCHOOL MILESTONES SCALE

SPANISH VERSION

By

Catalina Patricia Morales Murillo

R. A. McWilliam
Professor of Education
(Chair)

Kim O'Kelley Wingate
Associate Professor of Education
(Committee Member)

Tom Buggey
Professor of Education
(Committee Member)

PSYCHOMETRIC CHARACTERISTICS AND SOCIALVALIDITY

OF THE 3M PRESCHOOL MILESTONES SCALE

SPANISH VERSION

By

Catalina Patricia Morales Murillo

A Thesis Submitted to the Faculty of the University of
Tennessee at Chattanooga in Partial Fulfillment
of the Requirements of the Degree of Master of Education: Special Education

The University of Tennessee at Chattanooga
Chattanooga, TN

May 2015

ABSTRACT

Development occurs through the interaction of an individual and the context (e.g. classroom)
where he or she develops. The literature has identified the need for a functional and contextualized tool to
measure children's participation and functionality in classroom routines. We developed the 3M Preschool
Milestone Scale, a developmental scale for children between 3 and 5 years old to be completed by
teachers. Through two studies, this project aims to evaluate the psychometric characteristics and the social
validity of the Spanish version of this tool in six preschools in Valencia, Spain. As well, we evaluate the
importance of the scale's items for assessing functionality and development from the teachers'
perspective. Q-sort analysis was used to identify teachers' groups as they rated the items importance.
Results indicated that the 3M scale has a strong internal consistency and social validity and has items
relevant to child functional assessment.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to:

Dr. McWilliam, for being my thesis chair and graduate student assistantship supervisor, I cannot be more thankful for all his guidance, teachings and support throughout this project. I have grown immensely working with him. I have greater admiration for his work and he has inspired my research and practice approaches. He has a true teacher soul, always teaching his students something new.

Dr. Wingate for being my academic advisor through the past two years, I really appreciate all her advice and teachings and for serving in my thesis committee. I have learned a lot from her experience and work in early intervention, and from her happy and helpful personality. She is always willing to give a hand to whoever needs it; even when she is extremely busy she will put whatever she is doing aside and help with a big smile.

Dr. Buggey for being part of my thesis committee and my professor, he has a heart for working with children with disabilities. I hope to be able to practice with such care and passion.

The Universidad Católica de Valencia (UCV), especially Dr. María Dolores Grau and Dr. Marga Cañadas, and the UCV's Early Learning Center: L'Alqueria, for helping to contact the centers that participated in this research project, without their support this project would have not been possible.

The teachers from the participant centers, without them this project would have not been completed. I am extremely thankful for their questions, time, and enthusiasm as they completed the scales.

Last but not least, I like to thank:

Dr. Tania Boavida, for her support and inside as the project was developed and through its conclusion. I thankful for her friendship and approachable personality and for always been willing to help me and give me valuable advice.

Rosa Fernández Valero, a hardworking professional, her positive personality has been a big encouragement. I thank her for listening.

The Million family, for being my host family for the past three years, they have become a second family to me when far away from home. I thank them for their support, care, and encouragement.

My parents, brothers and sisters, for their love and support throughout my 24 years of life, I am thankful for the laughers and thousand and more great experiences together. I thank them for teaching me to love, be respectful, and work hard.

Pau, for his unconditional and unceasing support thought this entire project, I cannot be more thankful for his words of encouragement and time, especially through the final days of my thesis writing and defense. He is the kindest, most caring and loving man I have ever met. I am a blessed girl.

God, to Him be the glory.

TABLE OF CONTENTS

ix

LIST OF TABLES

x

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Health, well-being, learning, biological, and behavioral processes underline and influence the course of human development (Hertzman & Boyce, 2010). Brain research shows that activation of neurons and brain shaping is in part a result of the experiences young children and infants are exposed to (Dipietro, 2000). Other lines of research have shown that development occurs with the interaction of an individual and the context where he or she develops (Bronfenbrenner & Morris, 1998; Sameroff & Fiese, 2000). Historically, this understanding of human development has had an impact on research conducted in the fields of early childhood education and early childhood special education. This research has resulted in an increased classroom arrangement, children's assessment, and interventions (Adolfsson, Malmqvist, Pless, & Granuld, 2011; Berliner & Rosenshine, 1977; de Kruif & McWilliam, 1999; Dunst, Hamby, Trivette, Raab, & Bruder, 2000; Dunst, Burder, Trivette, & Hamby, 2006; Favell, Reid, & Risley, 1983; Favell & Risley, 1984; Fisher & Berliner, 1985; McConnell, 2000; McWilliam, 1991; McWilliam, 2006; McWilliam, 2011; McWilliam & Bailey, 1992; McWilliam & de Kruif, 1998; McWilliam & Scott, 2001; Wolery, Anthony, & Heckathorn, 1998).

Acknowledging the importance of children's participation and functionality across everyday classroom routines as a means for learning and developmental growth and in response to a need for functional and contextualized assessment in English and Spanish that facilitates identification of children's strengths and weaknesses to spur child functioning in the classroom, this project involved the development of a concise scale designed to allow teachers to rate children's functionality across classrooms routines. Psychometric properties such as internal consistency and social validity, factorial

1

characteristics of this new scale, and teachers' ratings of the importance of the scale and its items are presented.

Theoretical Background

Young children do not learn best through simulated trials in predetermined sessions but through constant interaction with their environment throughout the day and over time (McWilliam & Bailey, 1992). Children's learning is proportional to the amount of time they are engaged with the environment by interacting with caregivers, peers, and materials. As children interact with their environment, Independence, Social relationships, and Engagement have been identified as the three foundations of learning which facilitate acquisition of knowledge (McWilliam, 2006). These three foundations can also be used to measure child's functionality among everyday classroom activities and routines. The level of engagement, independence, and social relationships in classroom routines reflects a child's functioning and also allows a child to continue learning and mastering these skills. McWilliam (2006) has called these three pillars of learning (i.e., the foundations) functional domains. Considering these three functional domains can be helpful in evaluating the functioning of children with or without disabilities and could be integrated into intervention planning.

Engagement

McWilliam (1991) and, later on, McWilliam & Bailey (1992) defined engagement as the amount of time children interact with the environment (peers, teachers, materials) in an appropriate way in relationship to the child's age, ability, and his or her surroundings. Engagement has different levels, and the level of complexity increases as the child grows; thus, an infant's engagement may look different from the engagement of a 5-year-old (McWilliam & de Kruif, 1998; de Kruif & McWilliam, 1999). More time engaged and higher levels of sophistication of engagement are always better (McGarity & Buts, 1984). Following engagement theory, teachers and caregivers strive to provide children with

2

environments that elicit engagement at appropriate times, always striving to help children reach more sophisticated levels of engagement such as persistence (Dunst, 1996; Dunst, Bruder, Trivette, Raab, & McLean, 2001).

McWilliam (2011) indicated that authors like Berliner and Rosenshine (1977), Favell et al. (1983), Favell and Risley (1984), and Fisher and Berliner (1985) have found relationships between children's engagement and children's increased positive interactions with caregivers and peers, thinking and reasoning skills, and positive behaviors (McWilliam, 1991). Thus, more children's engagement results in a better classroom environment and more learning (Chien, 2010; McWilliam and Bailey, 1992).

The classroom's programmatic organization, environmental organization, and instruction can influence a child's engagement and promote his or her learning. The use of checklist training for teachers, classroom arrangement by open spaces and learning zones, assigning all adults to individual roles in a routine, assigning teachers to classroom zones and not the children, planning for smooth transitions from one activity to the other, and using incidental teaching and effective behavior management can help increase child engagement and reduce challenging behaviors (Dunst, McWilliam, & Holbert, 1986). McWilliam (2006) describes engagement as one of three functional domains along with independence and social relationships. These last two, however, are subsets of engagement.

Independence

This foundation of learning is concerned with children's ability to meet their needs and move around the environment without depending on an adult (McWilliam, 2011). Its meaning could differ from one culture to the other, thus considering the culture in which the child grows is crucial to determine the degree of independence of a child. For example, American parents tend to encourage independence more than Chinese parents (King & Bond, 1985). Therefore, American children's degree of independence may differ from Chinese children's degree of independence.

Social relationships

This foundation of learning is concerned with the child's ability to communicate and have appropriate interactions with peers and adults (McWilliam, 2011, p.128). Emotional stability and communication are linked to learning (Shonkoff & Phillips, 2000). In addition, Hamre et al (2014) indicate that children's interactions with teachers allow them to improve in the areas of social, behavioral, and cognitive development. These gains will transcend the preschool years and contribute to the child's future performance in other contexts. (Curby et al., 2009; Hamre & Pianta, 2005; O'Connor & McCartney, 2007). Children's social relationships are shaped by individual characteristics, social exchange, and cultural factors (Chen & French, 2008).

Old Paradigm-New Paradigm of Assessment

Over the past 40 years assessment has been dominated by norm-referenced standardized tests, which are being administered to children by strangers, in unfamiliar places, evaluating behavior and skills which may be irrelevant to a child's functioning in everyday routines (Meisels, Bickel, Nicholson, Xue, & Atkins-Burnett, 2001). Decontextualized assessment is not likely to produce results that represent children's actual functioning in classroom routines. It can encourage teachers to teach to the test and it does not identify a child's true functional needs (Bagnato, 2005; Meisels et al., 2001).

In addition to assessment, practices in early childhood have also come under scrutiny (McWilliam 2010; Odom & Wolery , 2003). Empirical evidence supports the importance and effectiveness of providing services in the child's natural context (Hwang, Chao, & Liu, 2013). Instead of focusing on children's deficits, the emphasis in these new models is on promoting functional skills that facilitate children's participation in daily routines at home, school, or any environment where the child spends most of his or her time.

Some current assessment practices are reverting to the historically valued observation of children and their environments (Bagnato, 2005; Meisels et al., 2001). Research has shown that contextualized

assessment of children's performance contributes to better integration of the curricula by the teachers, higher-order skills and thinking, more positive instruction, and more awareness of students' individual growth and development in the classroom (Meisels et al., 2001).

The paradigm shift started in the late 1980s, when Bronfenbrenner (1989) proposed the ecological systems model. His research supported the idea that children's learning and development occurs throughout constant interaction with their environment.

Ecological Model and Bioecological Theory

In his earliest publications Bronfenbrenner (1975, 1977a, 1979b) explained human development in light of what he called the ecological model. He indicated that human development takes place when there is an interaction between a human being who is growing and changing and the context in which this individual lives and relates with others.

In the beginning, Bronfenbrenner (1977c; 1978; 1979b) presented the environment as divided in different levels, with some levels proximal to the developing individual, having a more direct impact on his or her development. These levels were called the microsystem and the mesosystem. Other levels were more distal, still influencing indirectly an individual's development. These levels were called the exosystem and the macrosystem. Here we will focus on the microsystem and mesosystem levels. The microsystems are the immediate contexts in which the child spends most of his/her time and therefore could most influence a child's development (home, child care, and playground). The interaction between two or more microsystems creates a mesosystem. For example, the interaction between home and childcare is a mesosystem (Bronfenbrenner, 1977; 1978; 1979b).

Later, Bronfenbrenner targeted his research on the role of individual characteristics and how these characteristics affected development. This emphasis resulted in the evolution from the ecological model to the bioecological theory of human development (Bronfenbrenner & Ceci, 1993; Bronfenbrenner & Morris, 1998, 2006).

5

Transactional Model

Sameroff (1983) also stated that child development is the result of the interaction of the child and his or her context, especially the social context. He formulated these ideas as the "transactional model," taking into consideration the interaction between the genetic characteristics of a child (genotype) and the visible expressions of these characteristics (phenotype) and the environment where the child is developing. This model proposes that even when biological damage exists, the right interactions with the environment could have a positive impact on a child's development. A child's developmental achievements are the result of a transaction between the environment and his or her biological constitution, each influencing the other and the child's development equally (Sameroff & Fiese, 2000). The same environment and people may interact differently with different children (Sameroff & Fiese, 1990), which means individualized assessment is needed.

Natural Environments

Natural environments are contexts where children have opportunities to practice existing skills and to learn new ones (McWilliam & Ware, 1994). Dunst et al. (2000) explained that natural environments are contexts where children are provided with natural learning opportunities. Also, as defined by McWilliam (2005), natural environments are not a specific setting or place, these are contexts (cultural, religious, community, school, social, etc.) where children and their families spend most of their time and participate in what the author called life routines.

Routines. McWilliam and Scott (2001) have defined routines as times of the day or recurring events. For example, a classroom routine could be washing hands, meal time, nap time, or free-play. Other authors like Coster and Khetani (2008) have defined routine taking into consideration the International Classification of Functioning for Children and Youth (ICF-CY) and its components of Activities and Participation. They explained that a routine is a sequence of simple and individual

6

activities/actions undertaken to accomplish a goal or purpose. For example, eating with a spoon is an action itself, drinking from a cup is another action, asking for more food is an action too. When eating with a spoon, drinking from a cup, and asking for more food are put in a sequence, then, these are done to accomplish a bigger goal like feeding oneself, and it becomes the routine of meal time. Because the McWilliam definition is the "time of day," these definitions largely are aligned with each other.

The extent to which children are able to perform these activities reflects the level of participation. Routines as well as transitions between activities represent opportunities for teaching children (Wolery, et al., 1998). Measuring children's participation in those routines, in natural environments, is necessary to determine children's level of functioning, to identify skills to be learned; and to promote participation, which is important for children's learning, especially for those with a disability (Chien, Rodger, Copley, & Skorka, 2014)

Teacher's Role

Concerns have been expressed about relying on teachers' rating children's performance and skills in the classroom based on their observations in the natural context. These concerns are about the trustworthiness, validity, reliability, the teachers' knowledge of child development, and their ability to distinguish between a student's motivation and actual performance (Hoge & Coladarci, 1989; Salvensen & Undhiem, 1994). However, these concerns have been put to rest by research results indicating that teachers' judgement when assessing their students' learning and skills by using a performance assessment tool and observation in the natural context can be trusted (Meisels et al., 2001).

Teachers and caregivers are in an advantageous position to obtain valuable information on a child's participation during a routine and to evaluate a child's intellectual (engagement), behavioral (independence), and socio-emotional (social relationships) skills (Calfee & Hiebert, 1991; Kenny & Checaluk, 1993). Based on this knowledge, teachers can determine children's existing skills and new skills to be learned in natural occurring classroom routines.

7

Teachers have the opportunity to facilitate naturally occurring learning opportunities for the child to practice and develop new skills and also to help children engage in more sophisticated levels of engagement like persistence and pretend play (McWilliam, 2010). These opportunities are born after observation and careful consideration of children's current and future skills. If the items on the observation tool are functional and used for intervention planning, it can be considered a curriculum-based assessment (CBA).

Curriculum-Based Assessment (CBA)

Teachers use a CBA several times during an intervention or academic period to evaluate where the student stands in relationship to the program objectives (i.e., the "curriculum"). Based on the student's assessment results, the curriculum is adjusted or modified to meet the needs of the student or intervention is provided (Fuchs & Fuchs, 1986). Functional and contextualized assessment, such as a CBA, can contribute to the onset of interventions and modifications to the classroom environment to promote children's acquisition of skills and to keep a record of child improvement over time (Dunst et al. 1986).

Because children learn through repetition through the day and over time (McWilliam & Bailey, 1992) and through interaction with the environment (Bronfenbrenner & Morris 1998; Sameroff & Fiese, 1990, 2000) and considering the benefits of using CBA and functional and contextualized assessment (Chien et al., 2014; Dunst et al., 1986; Fuchs & Fuchs, 1986) a need to develop a tool that measures children's participation in classroom routines was identified (Adolfsson et al., 2011; Bedell & Coster, 2008; Chien et al et al, 2014; Coster & Khetani, 2007; Morris, Kurinczuk, & Fitzpatrick, 2005; Ziviani, Desha, Feeney, & Boyd, 2010).

Tools for Measuring Functionality through EISR

The literature was reviewed to find classroom scales that evaluate children's engagement, independence, and social relationships while participating of classroom routines. Three scales have been

8

identified: Scale for Teacher's Assessment of Routines Engagement (STARE) (McWilliam, 2000, 2011), Classroom Measure of Engagement Independence and Social Relationships (ClaMEISR) (McWilliam, 2014), and Measure of Engagement Independence and Social Relationships (MEISR) (McWilliam, & Younggren, N., 2012). The first two scales have been designed for classroom teachers to rate children based on their experience with them. The STARE is a 45-item, 5- point rating scale, rating from 1= Almost none of the time to 5= Almost all of the time, which is used to rate the amount of time a child is engaged overall during a routine and with adults, peers, and materials The complexity of the child's engagement in that routine is also rated on a 5-point scale. The STARE consists of six of the most common classroom routines (arrival, circle time, centers/free play, teacher-directed activity, and snack/lunch) and has three blank routines for the teachers to incorporate other routines. The teacher must have observed the child for 10 minutes in a routine to rate the child in that routine (Casey & McWilliam, 2007).

The MEISR was designed for children between 0 and 36 months old. It has 382 items and is subdivided into 13 common home routines. Each routine contains engagement, independence, and social-relationship behaviors, and each behavior has an approximate starting age. This scale is completed by a child's caregiver, who rates the competence of the child on this behavior from 1 to 3: 1=child does not perform the behavior yet, 2 = child sometimes performs the behavior, and 3 = child has mastered the behavior or has outgrown the skill. This measurement has been translated into Spanish, and is available for use, but the psychometric properties of this Spanish version have yet to be studied.

The ClaMEISR, like the MEISR, measures the three foundations of learning or functional domains, engagement, independence, and social relationships (EISR), in children between 3 and 5 years old. This scale is completed by the teacher.

Even though all three scales are organized by routines, only two of them measure the three foundations of learning: engagement, independence and social relationships. Of these, only the ClaMEISR could be used with children in the 3 to 5 year old range. Moreover, even though ClaMEISR is

a strong option to measure children's EISR, covers almost all classroom routines, completing it takes a long time because of the number of items. If used for following one child, this would be a realistic option. However, for evaluating all the children in the classroom, it would be difficult for a teacher to complete the ClaMEISRs regularly. In addition, no psychometric characteristics have been studied.

Although there is at least one strong scale for measuring preschoolers engagement, independence, and social relationships (McWilliam, 2014), a need exists for a scale with fewer items that could be completed by the teacher in less than 5 minutes and that still measures engagement, independence, and social relationships for both typically and atypically developing children.

Functional and Contextualized Assessments in Spanish

Functional and contextualized assessments are needed in Spanish. With better health care policies and social development in middle- and low-resource countries, infant and child mortality has decreased. Concomitantly, the infant, toddler, and early-childhood population has increased (Scherzer et al, 2012). Therefore, there is a higher demand for educational and intervention services for this age group than ever before. Latin American countries are in need of functional and contextualized measurements targeting children's development and learning in the classroom.

Functional and Contextualized Assessments in Costa Rica

For this study, the focus will be Costa Rica, which is a developing country, with approximately 11% of its population under 6 years old, according to the 2011 census (Programa Estado de la Nación, 2013). Centers of Education and Nutrition (CEN) and Centers for Children's Integral Support (CINAI) are the main providers of early childhood education services. The main tool for evaluation of children's

development and skills has been the EDIN (Escala del Desarrollo Integral del Niño/ Integral Child

Development Scale). This scale is administered monthly during the first year of life and then annually.

Literature results favoring functional and contextualized evaluations are known in the early

childhood education field in Costa Rica. The importance of a holistic and integrated development is

understood. However, no steps have been taken towards implementing functional or contextualized

evaluations, and the clinical mentality, with an emphasis on deficits, still remains (Programa Estado de la

Nación, 2013). The traditional domains continue to be used for assessment, and therapists, physicians, or

psychologists are responsible for screening and diagnosing children. Only 38% of children under 5 years

old are enrolled in a preschool or early learning program. The Government's Annual Report of the state of

education indicated that there have not been any improvements in the percentage of children of low

performance in one of six domains over the past 3 years. Figure 1.1 shows the percentage of children with

low performance in these domains during 2011.
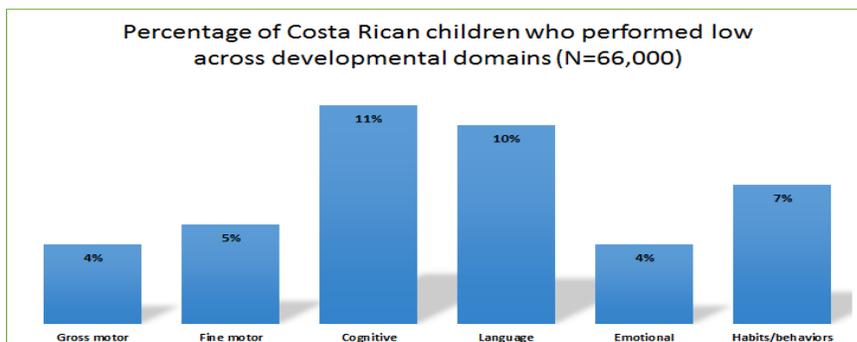
Figure 1. 1 Percentage of Costa Rican children who performed low across developmental domains

It appears the way children are taught needs to change, and this change should be guided by the

evidence (Odom & Wolery, 2003). Independence, social relationships, and most importantly, engagement

can become the core when considering curriculum and assessment of children's learning. The literature

has shown positive correlations between children's engagement and positive interactions with caregivers

and peers, thinking and reasoning skills, and positive behaviors (McWilliam, 1991). Targeting these

11

functional skills through functional and contextualized assessment could help move the curriculum towards recommended practices and inform teachers about children's needs (Wilson & Berne, 1999).

Purpose of This Project

This project aims to present the psychometric characteristics and the social validity of a new developmental scale, the 3M Preschool Milestone Scale, Spanish version, which is constructed to capture functional domains: engagement, independence, and social relationships (McWilliam, 2006, McWilliam, 2011). It was also pursued to understand teachers' perception on the importance of the scale items for evaluating child development and functionality, and their perception on the scale relevance for their classroom routines and possible for future use.

Originally, the psychometric properties of the scale were to be tested with Costa Rican children through collaboration with the CEN-CINAIs. However, due to scheduling conflicts, this was not possible. So, with the collaboration of the Universidad Católica de Valencia, Spain, and teachers from the Early Learning Centers: Sagrado Corazón, Ciutat Artista Faller, Joan Fuster, and L'Alquería, a sample of 366 children was obtained. This collaboration allowed for preliminary testing of the scale with Spanish-speaking teachers.

Research Questions

In summary, this study aimed to answer the following questions:

1. To what extent does the 3M Preschool Milestone scale, Spanish version, have internal reliability?
2. What is the underlying factor structure of the scale?
3. To what extent is the scale sensitive to age differences?
4. To what extent do children's scores vary by routine?
5. To what extent do children with typical development score differently from children with atypical development?

6. To what extent do teachers identify the items on the scale as relevant for evaluating child development?

7. To what extent do children score differently according to their teachers' years of experience or age?

8. To what extend do teachers differed according to their ratings on the items importance?

In order to answer those questions, we have designed two studies. Study 1 deals with the internal consistency and factor structure of the scale. As well, we investigated how do children's scores on the scale differed according to their age, family structure, type of school they attend (i.e. private or charter versus public), nationality, and disability presence. Study 2 deals with the social validity of the scale. In addition, it looked at the teachers' perceptions on the importance of the items for evaluating child development and functionality, as well, we use the Q-sort method, to study how did teachers group according to their ratings on items' importance.

To facilitate reading and understanding of these two studies, the methods, results and discussion of findings are presented separately for each study. First, it is presented Study 1 followed by Study 2. Introduction and Conclusion are shared for both studies

CHAPTER II

STUDY 1

ABSTRACT

This study pertained to the scores on the 3M. It was designed to answer these questions:

1. To what extent does the 3M Preschool Milestone scale, Spanish version, have internal reliability?

2. What is the underlying factor structure of the scale?

3. To what extend does the instrument captures the concepts of functioning?

4. To what extent is the scale sensitive to age differences?

5. To what extent do children's scores vary by routine?

6. To what extent do children with typical development score differently from children with atypical development?

METHODS

Participants

A total of 366 children attending either a public (n = 162) or private (n = 204) early learning center in Spain participated in the study. However, only the children with more than 80% of the scale completed (N = 364) were included in the analysis. Children came from six early learning centers as shown in Table 2.1.

Table 2. 1 Distribution of children by center

Early learning center	# of Children	%
Alzira-Colégio Público	12	3.3%
C.J.X.Javer	28	7.7%
Ciutat Artista Faller	25	6.9%
Joan Fuster	95	26.1%
L'Alquería	13	3.6%
Sagrado Corazón	191	52.5%
Total	364	100%

Three-hundred and twenty-three (89%) of the 364 children were born in Spain, and 32 (9%) were born in other countries such as: Romania, Portugal, and Colombia, for 9 children (2%) no nationality was reported. The children's ages in months ranged from 34 months to 70 months (M = 53). Twelve children had a disability including autism and language and developmental delays, and 352 were typically developing. Most children came from families with a middle income (N = 294), followed by those coming from lower income (N = 43), and finally children coming from high-income families (N = 27). Teachers reported that 257 of the 364 children lived with both parents, 70 children lived with divorced parents, foster parents, or one parent had died. Twelve children live with single mothers, and two lived with a single father. Data were missing for 22 children, concerning their legal guardian or caretaker.

3M Preschool Milestone Scale

The 3M, for short, is a 25-item, 4-point rating scale, where 1 = not yet, 2 = rarely, 3 = sometimes, and 4 = almost always. This scale measures a child's functioning during specified classroom activities: meals, free play, toileting, art, and teacher-led activity. Each routine has 5 items, each one representing a functional behavior which is developmentally appropriate for a child between the ages of 3 and 5 (see Appendix A). Because this study was conducted in Spanish, the 3M scale was translated. The Spanish version is in Appendix B. Psychometric properties are reported in the results section.

Procedures

Scale Design

Items were selected from (a) the behaviors and functions described in the International Classification of Functionality-Children and Youth (ICF-CY) (World Health Organization, 2007), (b) the Center for Disease Control (CDC) Milestones listings for 3-, 4-, and 5-year-olds (Centers for Disease Control and Prevention, 2014), and (c) the advice of two experts on child development, Robin McWilliam, Ph.D., and Tânia Boavida, Ph.D. The structure and organization of the scale replicated the Measure of Engagement, Independence, and Social Relationships (MEISR) (McWilliam, R. A., & Younggren, N., 2012).

Data Collection

Teachers at each early learning center received a package consisting of an instruction sheet, an informed consent form (one for the child's family and one for the teacher), the 3M Preschool Milestone Scales (one for each child in the classroom), a Q-Sort Matrix Form, and a Social Validity scale. The Q-sort is reported in Study 2. Teachers were asked to fill out a 3M Preschool Milestone Scale for each child who had been in the classroom for more than two weeks. After completing the 3Ms, they were asked to

complete the Social Validity scale and the Q-Sort Matrix Form. Surveys packages were dropped off and picked up in person by the researcher. This concluded the data collection stage.

Data Analysis

  The internal consistency of the 3M Preschool Milestones Scale was investigated, along with a factor analysis to identify the underlying factor structure. The procedure for factor analysis is described below. After these factors were identified, the internal consistency of the factors was analyzed. T-test and ANOVA were used to compare means of different group variables, and correlations were used to determine the relationship among all variables.

  For the exploratory factor analysis (EFA), assumptions regarding sample size, multivariate normality, linearity, and correlation among variables were evaluated to verify the appropriateness of the data for factor analysis (Comrey, 1973; Tabachnick & Fidell, 1989). Varimax rotation was used to carry out the EFA. For analysis, we accepted factors with an eigenvalue greater than 1 (Harman, 1976). Factor analysis (FA) was also used to determine if the instrument was really capturing the concepts of functioning. To analyze the internal consistency of the 3-M scale, we performed Cronbach's alpha, test for homogeneity, the KMO Index, and Bartlett's test of sphericity (Kaiser, 1974).

<div align="center">RESULTS</div>

<div align="center">

3M Preschool Milestone Scale

</div>

Internal Consistency of the Scale

  The scale was found to have strong internal consistency ($\alpha = .94$), which was true also for the scale factors (see Table 2.2).

Table 2. 2 Cronbach's alpha values by 3M factor

Factor	Cronbach's Alpha (α)
Engagement	0.93
Independence and self-expression	0.93
Following directions	0.86
Self-Help	0.76
3M Total	0.94

Exploratory Factor Analysis

Four factors were generated from the exploratory factor analysis (EFA) described above, explaining 71.97% of the variance. The four-factor solution, shown in Appendix G of this document, had a sufficient number of items with factor loadings greater than 0.4, and each factor had items with high loadings (greater than 0.7). Forcing five factors resulted in one very weak factor. Forcing three factors produced a low percentage of variance and too many "double loadings."

Examination of the items in Factor 1, by looking at the items' content, led to the conclusion that this factor measured Engagement. The content of these items was related to child's participation in classroom activities and interaction with peers, by talking or collaborating while playing. This factor explained 21.53% of the variance. Most of the items in this factor were part of Free Play and Meals routines. The following items constituted this factor: 1.3, 1.5, 2.1, 2.2, 2.3, 2.4, and 5.2.

The second factor measured Independence and Self-Expression. The items were related to children's independence while participating in classroom routines and their ability to express themselves by talking or creating something. This factor explained 21.03% of the variance. Most of the items in this factor were part of the Art routine. The following items constituted this factor: 2.5, 3.3, 3.4, 4.1, 4.2, 4.3, 4.4, and 5.3.

The third factor measured Following Directions. The items were related to following the teacher's directions while participating in a routine, especially during teacher-led activities. It explained 13.80% of the variance. Most of the items in this factor were part of Teacher-Led Activities. The following items constituted this factor: 1.4, 4.5, 5.1, 5.4, and 5.5.

The fourth factor measured Self-Help. The items were related to children's ability to take care of personal needs, which are essential for functioning within toileting. It explained 13.80% of the variance. Most of the items in this factor were part of the Toileting routine. The following items constituted this factor: 1.1, 1.2, 3.1, 3.2, and 3.5.

The factors found underline the construct structure of the scale and capture the concepts of functioning indicating that the scale has construct validity.

Relationships Among Factors

Correlations between the scale factors were moderate to large, as shown in Table 2.3. All factors were correlated with the total score, with Independence and Engagement having correlations above .90, showing the total score is most affected by these factors., The two other factors, Self Help and Following Directions, were also moderately to highly associated with the total score, validating the four-factor solution.

Table 2. 3 Correlations between 3M factors

	1	2	3	4
1.Total_3M	..			
2. Engagement	.925**	..		
3. Independence and Self-Expression	.946**	.820**	..	
4. Following Directions	.824**	.695**	.704**	..
5. Self_Help	.707**	.590**	.579**	.543**

Note. ** p = 0,01.

Children's Scores Across the 4 Factors and the Total 3M Mean by Age

Table 2.4 shows that children's 3M total score factor scores tended to increase as the age of the child increased. For the factors, all age groups scored the highest in Self- Help, but differed on the factors at which children scored the lowest. For the 3- and 5-year-olds, the lowest factor was Independence. The 5-year-olds also scored the lowest in Following Directions. For the 4-year-olds, the lowest score was in Following Directions.

19

Table 2. 4 Means and standard deviations by factors and age of the child

	3M Total Mean			Engagement			Independence			Following Directions			Self Help		
	N	M	SD	N	M	SD	N	M	SD	N	M	SD	N	M	SD
3 year olds	93	2.96	0.21	93	2.98	0.90	93	2.43	0.84	93	3.12	0.74	93	3.54	0.64
4 year olds	153	3.48	0.13	153	3.58	0.52	153	3.43	0.48	153	3.42	0.49	153	3.71	0.27
5 year olds	106	3.66	0.11	106	3.76	0.45	106	3.67	0.49	106	3.67	0.42	106	3.81	0.30

In order to have a better visual representation of the behavior of children scores and differences across age groups, Figure 1.2 is included.
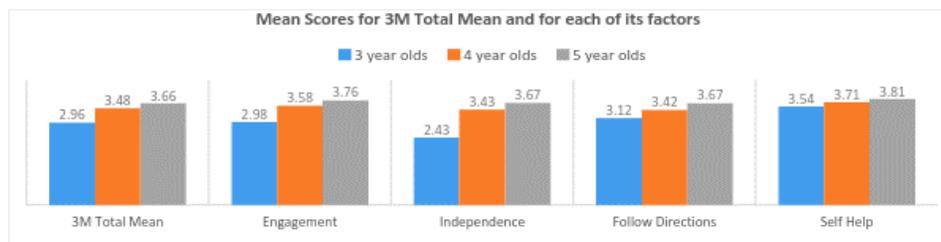
Figure 1. 2 Mean scores for the 3M Total Mean and for each of its factors

Cohen's d coefficients were calculated to determine the effect size for the differences between the factors across age groups. Table 2.5 shows Cohen's d when comparing two age groups' means (e.g. 3- and 4-year-olds) across all the factors and the 3M total mean. All differences between means were moderate to very large. Differences in total scores between 3- and 4-year-olds were almost three standard deviations. When comparing 3- and 5-year-olds, the overall score increased approximately 4 standard deviations, and for the 4- and 5--year olds the score changes were close to 1½ standard deviations.

Table 2. 5 Cohen's d values when comparing age groups across the 3M factors and the Total Mean score

	Cohen's d				
Comparison	3M Total Mean	Engagement	Independence	Following Directions	Self Help
3-4 year olds	2.99	0.83	1.52	0.49	0.36
3-5 year olds	4.29	1.16	1.87	0.93	0.56
4-5 year olds	1.42	0.38	0.51	0.54	0.37

By factor, the standardized difference between the 3- and 4-year-olds in Engagement and Independence were large, whereas in Following Directions and Self-Help they were small. The difference

between 4- and 5-year-olds in Engagement and Self-help was small, and, in Independence and Following Directions, it was moderate.

3M Descriptive Analysis by Routines

Table 2.6 shows that children in the three age groups tended to score lower in Art and higher in Eating. All children tended to score higher in a routine as their age increased.

Table 2. 6 Means and Standard deviation by routines and effects sizes between age groups

	3 year old group		4 year old group		5 year old group		d		
	M	SD	M	SD	M	SD	3 – 4	3 – 5	4 -5
Eating	3.34	.68	3.57	.49	3.78	.41	0.39	0.81	0.47
Free play	2.88	.93	3.46	.57	3.67	.58	0.77	1.05	0.37
Toileting	3.04	.76	3.52	.43	3.71	.48	0.81	1.08	0.42
Art	2.35	.83	3.41	.58	3.58	.66	1.50	1.65	0.27
Teacher-led activ.	3.08	.79	3.53	.52	3.66	.53	0.69	0.88	0.25

The standardized differences in means for the Eating routine were small but noteworthy between 3- and 4-year-olds and between 4-and 5-year-olds. The differences in Free Play and Teacher-Led Activities were moderate between 3- and 4-year-olds and small but noteworthy between 4- and 5-year-olds. Differences in Toileting and Art were large between 3- and 4-year-olds and small but noteworthy between 4- and 5-year-olds. These results suggest the 3M was sensitive to age differences.

Child Demographic Variables

Age

One-way analysis of variance (ANOVA) was conducted to determine the impact of age on the total score and the four factor scores. Table 2.7 shows all five ANOVAs were statistically significant, even after a Bonferroni correction.   Effect sizes reveal age had a highly noteworthy impact on scores for Engagement and the total score and a moderate impact on the other three factors. This analysis confirmed

with inferential statistics that the instrument was sensitive to age. Post-hoc tests Confirmed differences between age groups on all factors and the total score.

Table 2. 7 Anova results for children's age x 3M factors

		N	M	SD	F	df	p	η²
Total_Mean_3M	3 years old	95	2.91	0.68	57.52	2	<.001	0.24
	4 years old	158	3.49	0.45				
	5 years old	111	3.67	0.49				
Engagement	3 years old	95	2.97	0.90	32.73	2	<.001	0.15
	4 years old	158	3.54	0.58				
	5 years old	111	3.70	0.57				
Independence and Self-expression	3 years old	95	2.42	0.84	101.32	2	<.001	0.36
	4 years old	158	3.39	0.54				
	5 years old	111	3.62	0.58				
Follow_Directions_F3	3 years old	95	3.12	0.74	18.42	2	<.001	0.09
	4 years old	158	3.39	0.53				
	5 years old	111	3.62	0.51				
Self_Help_F4	3 years old	95	3.52	0.66	8.05	2	<.001	0.04
	4 years old	158	3.69	0.31				
	5 years old	111	3.77	0.42				

Private and Charter vs. Public Centers

An independent-samples t-test was conducted to compare private and charter centers with public centers. Table 2.8 again shows differences for all scores except Self-Help, although the differences were very small.

Table 2. 8 Independent samples t-test results and effect size by Early Learning Center type

		N	M	SD	t	df	p	d
3M Total Mean	Public	160	3.50	0.53	2.975	361.43	.003	-0.31
	Private or Charter	204	3.31	0.65				
Engagement	Public	160	3.55	0.69	2.643	354.32	.009	-0.28
	Private or Charter	204	3.35	0.76				
Independence	Public	160	3.34	0.67	2.858	361.57	.005	-0.30
	Private or Charter	204	3.11	0.88				
Following Directions	Public	160	3.50	0.55	3.073	360.56	.002	-0.32
	Private or Charter	204	3.31	0.66				
Self Help	Public	160	3.69	0.42	.550	361.20	.583	-0.06
	Private or Charter	204	3.66	0.51				

Family Structure

ANOVA was conducted to examine whether the family structure influenced child functioning. Mean scores of children living with both parents were compared to other family structures, consisting of single parents and other type of caregiver. This grouping was done based on literature findings on the influence of having both parents at home versus having one parent in child outcomes. As shown in Table 2.9, the Total Mean 3M and the Engagement and Independence scores were affected by family structure, with other children from other family structures performing higher than children from two-parent households. These differences were small but noteworthy and in the opposite direction from what was expected.

Table 2. 9 Anova Family structure and 3M total mean score and 3M factors

		N	M	SD	F	df	p	η²
Total Mean 3M	Both Parents	257	3.35	0.63	9.74	1	.002	0.03
	Other	85	3.59	0.47				
Engagement	Both Parents	257	3.38	0.77	12.45	1	<.001	0.03
	Other	85	3.70	0.53				
Independence	Both Parents	257	3.15	0.83	11.71	1	.001	0.03
	Other	85	3.49	0.62				
Following Directions	Both Parents	257	3.37	0.66	1.95	1	.164	0
	Other	85	3.47	0.49				
Self Help	Both Parents	257	3.65	0.49	1.19	1	.276	0
	Other	85	3.72	0.42				

Nationality

To determine whether the child's nationality made a difference in his or her 3M score, t tests were conducted between means for Spanish children and means for children of other nationalities. As Table 2.10 shows, for the total, Engagement, and Following Direction scores, Spanish children scored higher than did other-nationality children (large standardized difference for the first two means and moderate for the third).

Table 2. 10 T test between nationalities

		N	M	SD	t**	df**	p**	d
3M Total Mean	Spanish	325	3.43	0.58	2.51	31.90	0.02	0.54
	Other	30	3.05	0.80				
Engagement	Spanish	325	3.49	0.68	3.10	31.36	0.00	0.70
	Other	30	2.89	1.03				
Independence	Spanish	325	3.24	0.79	1.965	33.294	.058	0.39
	Other	30	2.91	0.89				
Following Directions	Spanish	325	3.42	0.59	2.06	31.49	0.05	0.46
	Other	30	3.08	0.87				
Self Help	Spanish	325	3.69	0.44	1.58	31.68	0.12	0.35
	Other	30	3.50	0.63				

Note. **Equal variance is not assumed

Disability vs. No Disability

To determine whether the 3M was sensitive to the presence of a disability in the child, means between children with a disability and those without a disability were compared, using t tests. All differences were statistically significant except for Self-Help, with children without disabilities scoring higher, as expected. All differences between means, including Self Help, were large, when effect sizes were examined, as shown in Table 2.11.

Table 2. 11 Independent Samples t-test Disability x 3M Scale and its factors

		N	M	SD	t	df	p	d
3M Total Mean	No-disability	12	3.53	0.69	2.980	22	.007	1.23
	Disability	12	2.54	0.92				
Engagement	No-disability	12	3.63	0.69	3.620	22	.002	1.50
	Disability	12	2.38	0.98				
Independence	No-disability	12	3.47	0.71	3.012	22	.006	1.25
	Disability	12	2.38	1.04				
Following Directions	No-disability	12	3.38	0.79	2.100	22	.047	0.86
	Disability	12	2.61	0.99				
Self Help	No-disability	12	3.67	0.67	2.050	22	.052	0.85
	Disability	12	2.96	1.00				

DISCUSSION

A need for a functional and contextualized scale in Spanish that facilitates children's assessment of engagement, independence, and social relationships in classroom routines, and completed by the teachers was identified. This first study intended to evaluate the psychometric characteristics of a new functional assessment in its Spanish version, the 3M Milestones Scale Spanish, and children scores in functionality across five classroom routines (i.e. Eating, Free Play, Toileting, Art, Teacher Led Activity) when rated by their teachers.

Scale Level Findings

The 3M Milestones Scale Spanish was found to have a high internal consistency (α=.94), indicating that all the items in the scale are strongly correlated. This internal consistency is comparable to the internal consistency of other renowned scales used in the field such as Parents Evaluation of Developmental Status: Developmental Milestones (PEDS: DM, Brothers, Glascoe, & Robertshaw, 2008) and Ages and Stages Questionnaire (Filgueiras, Pires, Maissonette, & Landeira-Fernandez, 2013; Hornman, Kerstjens, de Winter, Bos, & Reijneveld, 2013).

A factorial analysis of the scale resulted in four factors: engagement, independence, self-expression, Following Directions and Self -Help, which presented high levels of internal consistency (range .76 to .93). The scale's factorial structure reflects in its content the three foundations of learning (engagement, independence, and social relationships) (McWilliam, 2011) indicating that the scale has construct validity for measuring child functioning. The first factor, engagement, contains items related to children's engagement in routines, the items in the independence/self-expression and self-help factors relate to children's independence and social relationships, and finally Following Directions items are related to Social Relationships.

25

Results show that routines tended to have more items of a specific factor. For example most items in Free Play and Eating were engagement items, whereas most of the items in Toileting were self-help. Art items were mostly independence and self-expression items, and finally, items in the teacher led activity routine were mostly following direction items. When looking at the routines and how items belong to a specific factor as well as how structured a routine is, unstructured routines like Free Play was related to engagement, not surprisingly since free play offers children great opportunities to engage in exploration, pretend playing, and other different ways of engagement (McWilliam & de Kruif, 1998; de Kruif & McWilliam, 1999). Items of more structured routines such as Teacher Led Activity were related to Following Directions, this may be explained by the level of involvement that teachers may display have in this routine compared to other routines like Free Play, given that in activities that are led by teachers, teachers are giving children more instructions. Therefore, children's functionality may be related to their ability to follow these directions.

Independence and Engagement factors had the highest correlations with the overall scale, r=.95 and r=.93 respectively; whereas, self-help had the lowest correlation at r= .71. Between factors the highest correlations were between Engagement and Independence, r=.82, followed by Following Directions and Independence (r=.70). Self-Help had the lowest correlations with all the other factors indicating that there is a need for reviewing the items in Self-help and to identify if those items are worded differently or if there are any other differences that may be contributing to lower correlations with the overall scale and other factors.

Child Level Findings

Anova and t-test results when comparing children's total scale mean scores and factors mean scores, after grouping by children's variables, were found to be statistically significant in the following categories: age in years, type of school children's attended, type of family structure (both parents vs. single or foster parents), child nationality, and disability presence.

26

For children's age in years, scores differed across age groups. Older children tended to score higher in the overall mean and for each factor, and these differences were statistically significant. This finding corresponds to our research question regarding differences in 3M scores according to children's age. We found that children did differ in their scores as age increased. These results are attributed to the developmental maturity of children as they age increases and gain more skills.

This was also the case when looking at children's scores by routines. Children scored higher on Eating items, which is a highly functional routine requiring engagement, independence, and social skills, but is not as sophisticated (e.g. Uses words, signs, and/or gestures to communicate needs to the teacher or classmates). The lowest scores were in the Art items. This a more abstract routine which could require sophisticated levels of engagement skills that may still not be attained at 5 years old (e.g. Creates representations of real objects (draws, paints, builds things that resemble real objects).

Other variables also contributed to differences on children's scores Children with both parents tend to score lower than children with single or foster parents. This finding was surprising since literature indicates that children tend to have better outcomes when having both parents at home (Ackerman, D'Eramo, Umylny, Schultz, & Izard, 2001; Carlson & Corcoran, 2001). However, literature also indicates that the influence of having both parents in child outcomes differs across race and socioeconomic status (Brooks-Gunn, Han, & Waldfogel, 2002), which may help to explain our findings. This is because the literature supporting having both parents as positive influences on child outcomes are based on studies using a sample of US families (Battle, 1998; Dunifon & Kowalesky-Jones, 2002; Shaw, Winslow, & Flanagan, 1999). It will be interesting to explore these findings on a sample from Spain to determine if cultural differences may have a similar effect as race and socioeconomic status do. Also, Children born in Spain scored statistically significantly higher in the overall scale and in Engagement and Following Directions than children who were born in another country.

Finally, children with disabilities tended to score lower than children without disabilities, showing that the scale was sensitive in relation to having a disability. This finding goes along with

27

findings in literature that indicate that children with disabilities may lack or have a lower performance of certain skills (Lowenthal, 1992). Nonetheless, these findings must be taken carefully due to the small sample and also that data on the severity of the child's disability was not collected.

3M and Renown Developmental Scales: PEDS: DM, EDIN, and ASQ-3

When looking at the content of the 3M scale, there have been identified some similarities with other developmental and milestone scales, more specifically the Parents Evaluation of Developmental Status: Developmental Milestones (PEDS: DM) (Glascoe & Robertshaw, 2008), and Ages and Stages Questionnaire (Squires & Bricker, 2009). These scales were originally designed in English then translated in Spanish and other languages. Several studies suggest the validity and reliability of the translated version of these scales (Brothers et al., 2008; Filgueiras et al., 2013; Hornman et al., 2013; Juneja, Mohanty, Jain, & Ramji, 2011; Kerstjens et al., 2009; Troude, Squires, L'Hélias, Bouyer, & de La Rochebrochard, 2011). As the 3M Milestones scale, these two scales are completed by caregivers who have knowledge about the child's functionality, however, the caregiver is the parent not the teacher, and have been normed in preschools and clinics. Even though these scales are targeted to evaluate traditional developmental domains, and the 3M scale is based on functional domains, as defined by McWilliam (2006) as Engagement, Independence and Social Relationships (McWilliam, 2011, p.128); the 3M scale domains, are comparable to the content of the traditional domains found in the PEDS: DM and ASQ-3. For example, items that are related to engagement are related to the problem solving domain in ASQ-3 and to the pre-academic/literacy domains in the PEDS: DM scale. The items related to social relationships are comparable to those in the personal-social domain in the ASQ-3 and the socio-emotional domain in the PEDS: DM.

When looking at the age range of the 3M Milestone Scale in comparison to the PEDS: DM and ASQ-3, the 3M Milestones Scale is more specific to be used in preschool classrooms, age range= 3 to 5

28

years old; whereas PEDS: DM and ASQ-3 have large age ranges and are targeted for developmental screening of children, birth to 8 years old and 1 month to 5 ½ years, respectively.

It is important to highlight that although the 3M Milestones Scale could be used to evaluate children's functionality in relationship to functional domains, this scale is meant to be part of curriculum-based assessment and not for developmental screening or to be used as the only curriculum-based assessment. However, it cannot be ignored that the 3M Milestones scale, as part of the curriculum-based assessment, could facilitate the assessment of children's functionality, in the scale four factors. Moreover, this scale is contextualized, is completed by teachers, and in short is needed in English and Spanish.

Our results support that the 3M Milestone Scale could be used to evaluate a child's functioning. Results indicate that the 3M Milestone scale has a strong internal consistency and it discriminates depending on a child's age and if a child has a disability or not. Older children were identified as having a higher functioning than younger ones, and children with disabilities tended to score lower than children without disabilities.

3M and Other Routine-Based Scales

Other functional and contextualized scales like ClaMEISR (McWilliam, 2014) and STARE (McWilliam, 2000, 2011) share the routine-based structure and measure at least one of the functional domains. However, these two tools do not satisfy the need for a short functional and contextualized assessment of functional domains for children between 3 and 5 years old. STARE only measures one functional domain, Engagement, and ClaMEISR measures all three functional domains and has a complete list of behaviors for each routine. Because of its length, it is time consuming, it would be difficult to use it for assessing the functional skills of all the students in a classroom. No psychometric characteristics have been obtained for STARE or ClaMEISR, and there is no data on the ClaMEISR. The 3M Milestones Scale is a short, 25 item scale that can be completed in less than five minutes. Teachers

reported that it took them between 2-3 minutes to complete the scale. In addition, some of the psychometric characteristics of this scale are available.

Limitations

Among the limitations of this project, only the internal consistency of the scale and its four factors were researched. Other forms of reliability were not studied, like test-retests or inter-rater reliability. Only Face validity was obtained for the scale. Other forms of validity like construct or criterion-related validity of the scale need to be researched. Another limitation has to do with the generalization of the results. The sample was limited to children from six early childhood centers in Valencia Spain, and most of the children, as reported by teachers, came from families with income levels that were medium, meaning that they did not struggle financially nor have a high financial status. Therefore, any generalization from these project's results must be done carefully considering that these s may not be representative for all children's population between 3 and 5 years old.

Implications

Future lines of research need to address the remaining types of reliability and validity of the scale, in order to provide strong evidence to support that the scale is reliable, dependable and it measures what it is supposed to measure. As other steps could be related to norming the scale and reviewing the quality of the items. In addition, it would be interesting to run a RASH analysis to evaluate if the scale has adequate easy and difficult items for measuring child functioning of children between 3 and 5 years old. In relation to children's scores, results could be used to determine which are the children's functional strengths and weakness, and to develop intervention plans to address any weakness along with continuing to strengthen children's existing skill.

CHAPTER III

STUDY 2

ABSTRACT

This study was designed to study teachers' viewpoints about the importance of the 3M items. Whereas factor analysis, the "R" method, looks for correlations among items, Q looks for correlations among participants. The sorting consists of statistical methods to find groups of participants with shared ways of thinking. This study aimed to answer the following questions:

1. To what extent do teachers identify the items on the scale as relevant for evaluating child development?

2. To what extent do children score differently according to their teachers' years of experience or age?

3.   To what extend do teachers differed according to their ratings on the items importance?

METHODS

Participants

Twenty teachers participated in Study 1 by completing the 3M scale for children. From those 20 teachers, 95% of the teachers (N=19) filled out the Social Validity Scale and the Q-Sort Matrix; however, only 75% (N=15) teachers were included in the Q sort. One teacher did not complete the Q-Sort Matrix, and four were excluded because more than 5 of the items (> 20%) were not sorted. Teachers worked in six schools, as shown in Table 3.1.

Table 3. 1 Number of teachers by early learning center

Early Learning Center	#	%
Alzira-Colégio Público	1	5%
C.A.Xjaver	3	15%
Ciutat Artista Faller	3	15%
Joan Fuster	4	20%
L'Alquería	2	10%
Sagrado Corazón	7	35%
Total	20	100%

Teacher age ranged from 27 to 53 years (M= 41.30). Their years of experience working with children varied from 6 to 28 years (M= 15.64). Fifty percent of the teachers (N=10) reported having high knowledge in child development, 25% (N=5) reported having moderate knowledge, and 25% (N=5) did not answer this question. All teachers had at least an associate's degree. Sixty percent of the teachers (N=12) had an associate's degree, 5% (N=1) a licensure, 5% (N=1) a master's degree, and 30% (N=6) of the teachers did not answered this question. Teachers were contacted through the Catholic University of Valencia and through professionals who worked at these centers.

Measures

Social Validity

The Social Validity Scale (see Appendix C) was a five-item, multiple-choice questionnaire, with a 4-point rating scale, 1=Not Important/ Definitely not/ Not probable at all/ Not relevant and 4= Very Important/ Definitely Yes/ Very probable at all/ Really relevant. It was used to understand the social validity of the 3M. It included items such as, Do the activities used in this scale match those that take place in your classroom? And How relevant is the information in the scale in relationship to child development and functioning?

Q-Sort Matrix

This instrument was designed to collect data for performing a Q-sort without having to have cards with the items or to use a computerized program (see Appendix D for the Q-Sort Matrix). In Q sorts, the items are arrayed in a quasi-normal distribution to maximize the robustness of the factor analysis. First, the instrument has two circles with thirteen blanks on each, for teachers to classify items on the 3M as less important or more important. One circle had to have 13 items and the other 12. Second, teachers were asked to organize the items by importance in the columns provided in a matrix like the one shown in Figure 3.1. Each column has a value that increases from left to right, from 1 to 8, with each cell on a column having the same value, although teachers were unaware of this valuation. The first column on the far left contains the item perceived as the least important and is assigned a value of 1, and the column in the far right, which contains the most important item, is assigned a value of 8. Each teacher organized the items in these columns, from the least important to the most important item. Third, teachers provided demographic information such as years of teaching experience and level of education.

33

Ítems Menos Importantes ⟶ Ítems Más Importantes
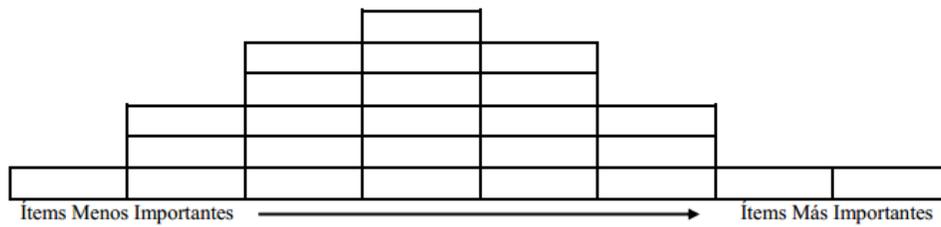
Figure 3. 1 Q-Sort

Procedures

The researcher met with the administrators and teachers at participating preschools. At this meeting the researcher explained the project, how to complete the scales, and answered questions. Also, she hand out the informed consent forms for the teachers and for participating children. Each teacher was given a survey package including (1) 3M Preschool Milestone Scales, one per student in their classroom, (2) one Social Validity Form, (3) a Q-Sort Matrix Form, (4) instructions sheet, and (5) an envelope for teachers to put and seal the completed surveys and signed informed consent forms to be returned to the researcher. Teachers were asked to complete the Social Validity Scale and Q-Sort Matrix Form only after they had completed their students 3M Preschool Milestone Scales. Teachers followed the directions described on the instructions sheet to complete the scales. After data collection was completed, envelopes containing the surveys packages were collected by the researcher, who was the only person who had access to the data. The researcher entered the data into Excel data analysis package and then copied into SPSS 20.0 to perform the data analysis.

Data Analysis

Exploratory factor analysis (EFA), using SPSS 20.0, was used to conduct the Q analysis, with participants listed as "items." The purpose of this analysis was not to generate factors but to group teacher according to their ratings of the scale's items importance. After Varimax rotation, we began by investigating factors with an eigenvalue greater than 1 (Harman, 1976).

34

RESULTS

Social Validity Scale

Teachers reported that having knowledge about their children's development was very important (M=3.86, SD=.36, on a 4-point scale). Of the 14 teachers who completed the Social Validity Scale, 64% reported they probably would use the 3M Preschool Milestone scale to evaluate children's functionality in the future. All teachers reported that it was easy or really easy to complete the scale. Twelve teachers reported that the scale routines were relevant to their classroom routines. Finally, all teachers reported that the content of the scale was relevant to very relevant for measuring child development and functionality.

Q Analysis

A Q analysis was performed to identify teacher groups according to their patterns of organizing the 25 items of the 3M scale from the least important to the most important. The exploratory factor analysis (Q analysis) resulted in five factors that explained 77.26% of the variance, but, when analyzing the components of the rotated solution, one of the factors had fewer than three teachers with coefficients > .40, so a four-factor solution was forced. Four factors explained 69.66% of the variance, but the fourth factor again did not have more than 3 teachers with acceptable factor loadings. Finally, forcing a three-factor solution explained 60.69% of the variance and had a sufficient number of teachers with factor loadings > .40. Each factor had "items" greater than 0.7.

Appendix H shows the results for the final Q analysis. Examining the groups led us to conclude that Group 1 consisted of least experienced-lower 3M teachers, Group 2 was more experienced-middle 3M teachers, and Gourp 3 was the most experienced-higher 3M teachers.

The following describes each group in more detail. Examination of 3M scores by group showed that the groups differed statistically significantly (see Table 3.2) on the total score, Independence, and Self-Help; alpha is .01 because of the Bonferroni correction. Although Engagement and Following

Directions did not achieve statistical significance, effect sizes showed the group differences to be noteworthy but small. In fact, all differences were small.

Group 1- Least experienced-lower 3M teachers: Teachers grouped in this factor shared having the fewest years of experience compared to the other teachers who participated in the study. Teachers' age ranged from 8 to 20 years of experience (M=13.17). This factor explained 12.51% of the variance. Table 3.3 shows this group had the lowest mean scores for the total 3M and each of the factors.

Group 2-More experienced-middle 3M teachers: Teachers grouped in this factor were in the middle range of years of experience compared to the other teachers who participated in the study. Teachers' age ranged from 6 to 28 years of experience (M=16.5). This factor explained 23.93% of the variance. Table 3.3 shows this group had the middle mean scores for the total 3M and each of the factors.

Group 3- Most experienced-Highest 3M teachers: Teachers grouped in this factor all shared having the most years of experience compared to the other teachers who participated in the study. Teachers' age ranged from 19 to 22 years of experience (M=20.5). This factor explained 24.26% of the variance. Table 3.3 shows this group had the highest mean scores for the total 3M and each of the factors.

Table 3. 2 Anova Teacher Factors vs. 3M factors

		N	M	SD	F	df	p	η²
Total Mean 3M	Least experienced	77	3.20	0.70	5.79	2	.003	0.04
	More Experienced	118	3.36	0.63				
	Most Experienced	55	3.58	0.54				
Engagement	Least experienced	77	3.29	0.88	4.64	2	.011	0.03
	More Experienced	118	3.33	0.75				
	Most Experienced	55	3.67	0.62				
Independence	Least experienced	77	3.02	0.85	6.00	2	.003	0.04
	More Experienced	118	3.18	0.81				
	Most Experienced	55	3.50	0.63				
Following Directions	Least experienced	77	3.21	0.70	4.12	2	.017	0.02
	More Experienced	118	3.39	0.66				
	Most Experienced	55	3.53	0.53				
Self Help	Least experienced	77	3.38	0.64	12.20	2	.000	0.08
	More Experienced	118	3.73	0.40				
	Most Experienced	55	3.65	0.44				

Groups and Importance Ratings of 3M Routines

The total mean and standard deviation for each routine was calculated. Teachers tended to give more or less importance to items in a routine according to their years of experience. Toileting was rated the most important by teachers with the most experience (a routine that requires independence and self-help), For Art and Teacher-Led Activities (possibly more structured routines) it was teachers with more experience who rated these routines as the most important. Finally, for Free Play and Eating routines (less structured routines) were teachers with less experience who rated those routines as the most important. Table 3.8 shows the means and standard deviations for all the routines and by teacher factor according to their experience.

Cohen's ds have been run to determine the effect sizes of the differences in the mean scores of teachers rating of the routines' items importance. Results are shown on Table 3.3.

Table 3. 3 Cohen's d values comparing teacher group factors

	Cohen's d				
	Eating	Free Play	Toileting	Art	Teacher Led Activ.
Least experienced - More experienced teachers	-0.76	-2.11	0.63	1.19	2.72
Least experienced - Most experienced teachers	-0.64	-1.22	1.34	0.40	-0.23
More experienced - Most experienced teachers	-0.12	0.87	0.86	-0.73	-1.24

Comparisons Across Teacher Experience Groups

When comparing Least experienced teachers to the More experienced teachers, effect size of the differences on the routine means were large for Free Play, Art, and Teacher Led Activities, and moderate for Eating and Toileting. When comparing Least experienced teachers to the Most experienced teachers effect size on the differences of the routine means were large for Free Play and Toileting, and moderate for Eating and small for Art and Teacher Led Activities. Finally, when comparing More experienced teachers to the Most experienced, effect size on the differences of the routine means were large for Free Play, Toileting, and Teacher Led Activities, moderate for Art, and small for Eating. Figure 3.2 illustrates the mean scores for each routine across the different groups by teachers' experience.

MEAN SCORES FOR TEACHERS' RATING OF ROUTINE ITEMS' IMPORTANCE IN RELATIONSHIP TO THEIR YEARS OF EXPERINCE

■ Less experienced teachers ■ More experienced teachers ■ Most experienced teachers

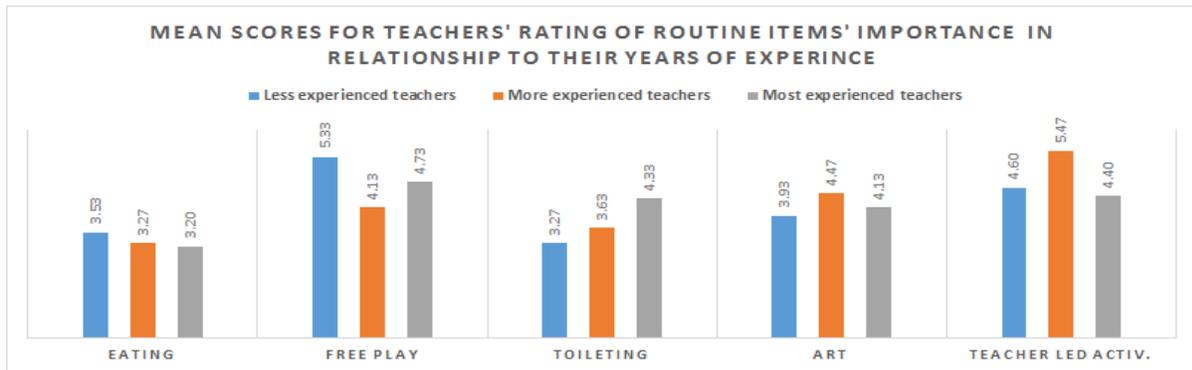|	EATING	FREE PLAY	TOILETING	ART	TEACHER LED ACTIV.
Less experienced teachers	3.53	5.33	3.27	3.93	4.60
More experienced teachers	3.27	4.13	3.63	4.47	5.47
Most experienced teachers	3.20	4.73	4.33	4.13	4.40

Figure 3. 2 Mean scores of teacher's rating of routine items' importance in relationship to teachers' years of experience

In conclusion, results of this study showed that three ways of thinking about children's functioning, as defined by the 25 items on the 3m, existed. The three ways corresponded with teachers' experience levels and their tendency to score the 3M higher or lower.

DISCUSSION

Social Validity Scale

Results from the social validity scale show that the scale has face validity. Teachers reported that the content of the scale was relevant to very relevant for measuring child development and child functioning. In addition, it was found that the scale is easy to complete and it takes less than 5 minutes per child (2-3 minutes). Identifying teachers' interests in child development, and determining that more than half of the teachers would use the scale for children's future assessment is promising. Future research should strive for validating the scale and norming children's performance according to their age in order to make this instrument even more useful for teachers to identify strengths and weakness in their students' functional development.

Teachers' Ratings on 3M Items Importance

The two items more frequently rated as more important, 5.1 and 5.5, are part of the scale factor Following Directions and part of the routine Teacher Led Activity. Indicating that teachers tended to see children's attention and ability to comply with teachers requests as important aspects for evaluating child's functionality, most of the second group of items selected more frequently as more important were part of the Engagement factor. It could be inferred that teachers consider engagement also important for evaluating functionality. It is important to highlight that these items were mainly oriented towards engagement with peers through communication, collaboration, and symbolic play. Since teachers were asked to categorize items and do it so that these two groups have an even number of items, there could be items that were considered important or less important and were assigned to the other group in order to keep the number of items even on each group. However, when looking at the average rating of each item when organized in the Q-Sort template, item 5.1: "Pays attention to the teacher for more than 5 minutes, when he/she is speaking to the group," had the highest mean. Followed by item 2.3: "Collaborates with peers while playing (e.g. negotiates play roles), and 5.5: "Follows teacher's rules and instructions." Item 1.4: "Picks up the table (throws trash away and picks up after him/herself) by own initiative" had the lowest mean of all. These results show that there is a consistency in the way teachers saw more important and less important items by groups and the way they accommodate each item in the Q-Sort template from the least important to the most important.

Findings in Q analysis

Q-sort analysis of teachers rating of items from least to most important resulted in three teacher groups. Teachers on each group had in common the number of years of experience. An interesting finding was that children's scores differed significantly according to their teachers' years of experience. For example, children with the most experienced teachers tended to score the highest for the Total 3M Mean, Engagement, Independence, and Following Directions, and children from more experienced teachers

scored the highest in Self-help. None of the children of less experienced teachers scored the highest in any of the factors or overall scale. These findings are supported by other literature which showed that as teachers' years of experience increase so did their students' performance (Avalos, 1985). These results may indicate that teachers with more experience may have more understanding of how to present curricula in a way that fosters children's engagement and natural learning opportunities. More learning opportunities facilitate practice of existing skills and acquisition of new ones (McWilliam & Ware, 1994) which may explain our findings.

Another interesting finding was that when averaging teachers' ratings of importance of items by routines, the average of importance for each routine differed depending on the teachers' years of experiences. Most-structured routines like art and teacher led activity were rated with higher importance by more experienced teachers, more-structured routines were rated as more important by the most experienced teachers, and finally, non-structured routines like eating and free-play were rated as more important by less structured teachers.

Limitations

As for study 1, among the limitations for this study are the lack of other forms of validity besides construct and social validity.

Implications

Future lines of research could concentrate on revising the scale items according to their importance for evaluating child functioning and development as reported by teachers. Also, the scale and items could be used as a guide for teachers to evaluate children's functionality across classroom routines. Results from this study could be used by teachers to adjust their routings in relationship to their children's functional strengths and weaknesses.

CHAPTER IV

CONCLUSION

A need for a short, functional, evidence-based, and contextualized instrument, measuring functional domains, in Spanish, was identified. The study was designed to develop a new functional scale and evaluate some of its psychometric properties. It was also designed to evaluate the social validity through teachers' perceptions of the items' importance of each item for evaluating functional development. This Q sort resulted in three groups of teachers. Results indicate that the 3M Preschool Milestone Scale scores had strong internal consistency of the overall scale and its factors. Findings also revealed that the scale discriminate between children's performance at different ages, as well as between children with and without disabilities. The sorting of the teachers on their importance ratings showed that each group had different lengths of experience. Teacher's groups rated the 3M items importance differently according to their years of experience. Further studies are needed to strengthen the reliability and validity of scale scores. Nonetheless, findings from this project indicate that the 3M scale could be used for assessment of children's functionality.

REFERENCES

Ackerman, B. P., D'Eramo, K. S., Umylny, L., Schultz, D., & Izard, C. E. (2001). Family structure and the externalizing behavior of children from economically disadvantaged families. Journal of Family Psychology, 15, 288–300.

Adolfsson, M., Malmqvist, J., Pless, M., & Granuld, M. (2011). Identifying child functioning from an ICF-CY perspective: everyday life situations explored in measures of participation. Disability and Rehabilitation, 33(13-14), 1230-1244.

Avalos, B. (1985). Training for better teaching in the third world: Lessons from research. Teaching and teacher education, 1(4), 289-299.

Bagnato, S. J. (2005). The authentic alternative for assessment in early intervention: An emerging evidence-based practice. Journal of Early Intervention, 28(1), 17-22.

Battle, J. J. (1998). What beats having two parents? Education outcomes for African American students in single- versus dual-parent families. Journal of Black Studies, 28, 783–801.

Bedell, G., & Coster, W. (2008). Measuring participation of school-aged children with traumatic brain injuries: considerations and approaches. The Journal of head trauma rehabilitation, 23(4), 220-229.

Berliner, D.C., & Rosenshine, B. (1977). The acquisition of knowledge in the classroom. In R. Spito & W. Montague (Eds.), Schooling and the acquisition of knowledge. Mahwah, NJ: Lawrence Erlbaum Associates.

Bronfenbrenner, U y Morris, P. A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), Handbook of child psychology. (5th Ed.). Vol.1: Theoretical models of human development. New York: John Wileys and Sons.

Bronfenbrenner, U. (1977c). Toward an experimental ecology of human development. American Psychologist, 32, 513–531.

Bronfenbrenner, U. (1978). Who needs parent education?. The Teachers College Record, 79(4), 767-787.

Bronfenbrenner, U. (1979b). The ecology of human development: experiments by nature and design. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1989, April). The developing ecology of human development: paradigm lost or paradigm regained. In biennial meeting of the Society for Research in Child Development, Kansas City, MO.

Bronfenbrenner, U., & Ceci, S. J. (1993). Heredity, environment, and the question" How?": A first approximation.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. Handbook of child psychology.

Brooks-Gunn, J., Han, W., & Waldfogel, J. (2002). Maternal employment and child cognitive outcomes in the first three years of life: The NICHD Study of Early Child Care. Child Development, 73, 1052–1072.

Brothers, K. B., Glascoe, F. P., & Robertshaw, N. S. (2008). PEDS: developmental milestones--an accurate brief tool for surveillance and screening. Clinical Pediatrics, 47(3).

Calfee, R., & Hiebert, E. (1991). Teacher assessment of achievement. Advances in Program Evaluation (vol.1, pp. 103-131). JAI Press.

Carlson, M. J., & Corcoran, M. E. (2001). Family structure and children's behavioral and cognitive outcomes. Journal of Marriage and Family, 63, 779–792.

Casey, A. M., & McWilliam, R. A. (2007). The STARE: The Scale for Teachers' Assessment of Routines Engagement. Young Exceptional Children, (11), 2-15.

Chen, X., & French, D. C. (2008). Children's social competence in cultural context. Annu. Rev. Psychol., 59, 591-616.

Chien, C. W., Rodger, S., Copley, J., & Skorka, K. (2014). Comparative content review of children's participation measures using the international classification of functioning, disability and health–children and youth. Archives of physical medicine and rehabilitation, 95(1), 141-152.

Chien, N. C., Howes, C., Burchinal, M., Pianta, R. C., Ritchie, S., Bryant, D. M., ... & Barbarin, O. A. (2010). Children's classroom engagement and school readiness gains in prekindergarten. Child Development, 81(5), 1534-1549.

Comrey, A.L. (1973). A First Course in Factor Analysis. , Academic Press, NY.

Coster, W., & Khetani, M. A. (2008). Measuring participation of children with disabilities: issues and challenges. Disability & Rehabilitation, 30(8), 639-648.

Curby, T. W., Lo Casale-Crouch, J., Konold, T. R., Pianta, R., Howes, C., Burchinal, M., … Barbarin, O. (2009). The relationships observed pre-K classrooms quality profiles to children's academic achievement and social competence. Early Education and Development, 20, 346-372.

de Kruif, R. E. L., & McWilliam, R. A. (1999). Multivariate relationships among developmental age, global engagement, and observed child engagement. Early Childhood Research Quarterly, 14, 515–536.

DiPietro, J. A. (2000). Baby and the brain: advances in child development. Annual review of public health, 21(1), 455-471.

Dunifon, R., & Kowaleski-Jones, L. (2002). Who's in the house? Race differences in cohabitation, single parenthood, and child development. Child Development, 73, 1249–1264.

Dunst, C. J. (1996). Early intervention in the USA: Programs, models, and practices. In M. Brambring, H. Rauh, & A. Beelmann (Eds.), Early childhood intervention: Theory, evaluation, and practice (pp. 11–52). Berlin, Germany: de Gruyter.

Dunst, C. J., Bruder, M. B., Trivette, C. M., Raab, M., & McLean, M. (2001). Natural learning opportunities for infants, toddlers, and preschoolers. Young Exceptional Children, 4(3), 18-25.

Dunst, C. J., McWilliam, R. A., & Holbert, K. (1986). Assessment of preschool classroom environments. Assessment for Effective Intervention, 11(3-4), 212-232.

Dunst, C.J., Burder, M.B., Trivette, C, & Hamby, D. (2006). Everyday activity settings, natural learning environments, and early intervention practices. Journal of Policy and Practice in Intellectual Disabilities, 3 (1) 3–10.

Dunst, C.J., Hamby, D., Trivette, C.M., Raab, M. & Bruder, M.B. (2000) Every family and community life and children's naturally occurring learning opportunities. Journal of Early Intervention, 23, 151-164.

Favell, J.E., & Risley, T.R. (1984, November). Organizing living environments for developmentally disabled persons. Workshop presentation at the 18th annual Association for the Advancement of Behavior Therapy Convention, Philadelphia, PA.

Favell, J.E., Reid, D.H., & Risley, T.R. (1983, December). Organizing living environments for developmentally disabled persons. Workshop presentation at the Congress on Behavior Therapy and 17th annual Association for the Advancement of Behavior Therapy Convention, Washington, DC.

Filgueiras, A., Pires, P., Maissonette, S., & Landeira-Fernandez, J. (2013). Psychometric properties of the Brazilian-adapted version of the Ages and Stages Questionnaire in public child daycare centers. Early human development, 89(8), 561-576.

Fisher, C. W., & Berliner, D. C. (1985). Perspectives on instructional time. Addison-Wesley Longman Ltd.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: a meta-analysis. Exceptional children.

Glascoe, F., & Robertshaw, N. (2008) PEDS Developmental Milestones: A Tool for Surveillance and Screening (Professionals' Manual). Nolensville, TN: PEDStest.com, LLC. Retrieved from: www.pedstest.com

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure?. Child development, 76(5), 949-967.

Hamre, B., Hatfield, B., Pianta, R., & Jamil, FM. (2014). Evidence for general and domain-specific elements of teacher–child interactions: Associations with Preschool Children's Development.Child Development, 85(3), 1257-1274.

Harman, H.H. (1976). Modern factor analysis. Chicago. IL: University of Chicago Press.

Hertzman, C., & Boyce, T. (2010). How experience gets under the skin to create gradients in developmental health. Annual review of public health, 31, 329-347.

Hoge, R, D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. Review of Educational Research, 59, 297-313.

Hornman, J., Kerstjens, J. M., de Winter, A. F., Bos, A. F., & Reijneveld, S. A. (2013). Validity and internal consistency of the Ages and Stages Questionnaire 60-month version and the effect of three scoring methods. Early human development, 89(12), 1011-1015.

Hwang, A. W., Chao, M. Y., & Liu, S. W. (2013). A randomized controlled trial of routines-based early intervention for children with or at risk for developmental delay. Research in developmental disabilities, 34(10), 3112-3123.

Juneja, M., Mohanty, M., Jain, R., & Ramji, S. (2012). Ages and stages questionnaire as a screening tool for developmental delay in Indian children. Indian pediatrics, 49(6), 457-461.

Kaiser, H. F. (1974). An index of factorial simplicity. Psychometrika, 39(1), 31-36.

Kenny, D. T., & CheKaluk, E. (1993). Early reading performance: a comparison of teacher-based and test-based assessments. Journal of Learning Disabilities, 26, 227-236.

Kerstjens, J. M., Bos, A. F., ten Vergert, E. M., de Meer, G., Butcher, P. R., & Reijneveld, S. A. (2009). Support for the global feasibility of the Ages and Stages Questionnaire as developmental screener. Early human development, 85 (7), 443-447.

King & Bond. (1985). The Confucian parading of man: a sociological view. In V: TSENG and D.Y.H. VW (EDS).

Lowenthal, B. (1992). Functional and developmental models: a winning early intervention combination. Infant-Toddler Intervention, 2, 161-168.

McConnell, S. R. (2000). Assessment in early intervention and early childhood special education building on the past to project into our future. Topics in Early Childhood Special Education, 20(1), 43-48.

McGarity, J. R., & Butts, D. P. (1984). The relationship among teacher classroom management behavior, student engagement, and student achievement of middle and high school science students of varying aptitude. Journal of Research in Science Teaching, 21(1), 55-61.

McWilliam, R. A. (1991). Targeting teaching at children's use of time: perspectives on preschoolers' engagement. TEACHING Exceptional Children, 23(4), 42-43.

McWilliam, R. A. (2000). Scale for teacher's assessment of routines engagement. Engagement of every child in the preschool classroom, 168-170.

McWilliam, R. A. (2005). Foundations for learning in a modern society. Paper presented at the Encontro Internacional Diferenciação do Conceito à Práctica, Universidade Católica Portuguesa, Porto, Portugal.

McWilliam, R. A. (2006). What happened to service coordination? Journal of Early Intervention, 28, 166-168.

McWilliam, R. A. (2010). Routines-based early intervention: Strategies for supporting young children with disabilities. Baltimore, MD: Brookes Publishing.

McWilliam, R. A. (2011). Scale for teacher's assessment of routines engagement. Siskin Children's Institute, Chattanooga, TN.

McWilliam, R. A. (2014). Classroom Measure of Engagement, Independence, and Social Relationships (ClaMEISR). Siskin Children's Institute, Chattanooga, TN.

McWilliam, R. A., & Bailey, D. B. (1992). Promoting engagement and mastery. In D. B. Bailey, & M. Wolery (Eds.), Teaching infants and preschoolers with disabilities (2nd ed., pp. 229–256). New York, NY: Merrill.

McWilliam, R. A., & Casey, A. M. (2008). Engagement of every child in the preschool classroom. Baltimore, MD: Paul H. Brookes Publishing Company.

McWilliam, R. A., & De Kruif, R. E. L. (1998). E-Qual III: children's engagement codes. Chapel Hill, NC: Frank Porter Graham Child Development Center, University of North Carolina at Chapel Hill.

McWilliam, R. A., & Scott, S. (2001). A support approach to early intervention: a three-part framework. Infants & Young Children, 13(4), 55-62.

McWilliam, R. A., & Ware, W. B. (1994). The reliability of observations of young children's engagement an application of generalizability theory. Journal of Early Intervention, 18(1), 34-47.

McWilliam, R. A., & Younggren, N. (2012). Measure of Engagement, Independence, and Social Relationships (MEISR).

Meisels, S. J., Bickel, D. D., Nicholson, J., Xue, Y., & Atkins-Burnett, S. (2001). Trusting teachers' judgments: a validity study of a curriculum-embedded performance assessment in kindergarten to grade 3. American Educational Research Journal, 38(1), 73-95.

Morris, C., Kurinczuk, J. J., & Fitzpatrick, R. (2005). Child or family assessed measures of activity performance and participation for children with cerebral palsy: a structured review. Child: care, health and development, 31(4), 397-407.

O'Connor, E., & McCartney, K. (2007). Examining teacher-child relationships and achievement as part of an ecological model of development. American Education Research Journal, 44, 340.

Odom, S. L., & Wolery, M. (2003). A unified theory of practice in early intervention/early childhood special education evidence-based practices. The Journal of Special Education, 37(3), 164-173.

Programa Estado de la Nación. (2013). Cuarto informe estado de la educación. San José, Programa Estado de la Nación. Retrived on February 10, 2015, from: http://www.estadonacion.or.cr/estado-educacion/educacion-informe-ultimo

Salvesen, K. A., & Undhiem, J. O. (1994). Screening for learning disabilities. Journal of Learning Disabilities, 27, 60-66.

Sameroff, A. J. (1983). Development systems: Contexts and evolution. Handbook of child psychology: formerly Carmichael's Manual of child psychology/Paul H. Mussen, editor.

Sameroff, A. J., & Fiese, B. H. (2000). Transactional regulation: the developmental ecology of early intervention. Handbook of early childhood intervention, 2, 135-159.

Sameroff, A.J. y Fiese, B.H. (1990). Transactional regulation and early intervention. EnS.J. Meisels y J.P. Shonkoff (Eds.). Handbook of early childhood intervention (pp.119-149). Cambridge: Cambridge University Press.

Scherzer, A. L., Chhagan, M., Kauchali, S., & Susser, E. (2012). Global perspective on early diagnosis and intervention for children with developmental delays and disabilities. Developmental Medicine & Child Neurology, 54(12), 1079-1084.

Shaw, D. S., Winslow, E. B., & Flanagan, C. (1999). A prospective study of the effects of marital status and family relations on young children's adjustment among African American and European American families. Child Development, 70(3), 742-755.

Shonkoff, J. P., & Phillips, D. A. (2000). From neurons to neighborhoods: The science of early childhood development. National Academy Press, 2101 Constitution Avenue, NW, Lockbox 285, Washington, DC 20055.

Squires, J., Bricker, D., & Twombly, E. (2009). Ages & Stages Questionnaires. Baltimore, Maryland, 257-182.

Tabachnick, B. G., & Fidell, L. S. (1989). Using Multivariate Statistics (2nd ed.). NY: Harper and Row.

Troude, P., Squires, J., L'Hélias, L. F., Bouyer, J., & de La Rochebrochard, E. (2011). Ages and Stages Questionnaires: feasibility of postal surveys for child follow-up. Early human development, 87(10), 671-676.

Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. Review of research in education, 173-209.

Wolery, M., Anthony, L., & Heckathorn, J. (1998). Transition-based teaching: Effects on transitions, teachers' behavior, and children's learning. Journal of Early Intervention, 21, 117–131.

World Health Organization (Ed.). (2007). International Classification of Functioning, Disability, and Health: Children & Youth Version: ICF-CY. World Health Organization.

Ziviani, J., Desha, L., Feeney, R., & Boyd, R. (2010). Measures of participation outcomes and environmental considerations for children with acquired brain injury: A systematic review. Brain Impairment, 11(02), 93-112.

APPENDIX A

INITIAL IRB APPROVAL LETTER

MEMORANDUM

TO:	Catalina Morales Murillo	**IRB #14-095**
	Dr. Robin McWilliam	
	Dr. Tom Buggey	
FROM:	Lindsay Pardue, Director of Research Integrity	
	Dr. Bart Weathington, IRB Committee Chair	
DATE:	June 19, 2014	
SUBJECT:	IRB #14-095: Psychometric Characteristics and Social Validity of the 3M Preschool Milestone Scale	

The Institutional Review Board has reviewed and conditionally approved your application and assigned you the IRB number listed above. The following requirement is necessary before research may begin:

> ➢ A letter of participation from CEN-CINAIs must be obtained before this research study can be approved.

Please provide a copy of the requested document to the IRB at instrb@utc.edu before beginning your research.

You must include the following approval statement on research materials seen by participants and used in research reports:

> ***The Institutional Review Board of the University of Tennessee at Chattanooga (FWA00004149) has approved this research project #14-095.***

Since your project has been deemed exempt, there is no further action needed on this proposal unless there is a significant change in the project that would require a new review. Changes that affect risk to human subjects would necessitate a new application to the IRB committee immediately.

Please remember to contact the IRB Committee immediately and submit a new project proposal for review if significant changes occur in your research design or in any instruments used in conducting the study. You should also contact the IRB Committee immediately if you encounter any adverse effects during your project that pose a risk to your subjects.
For any additional information, please consult our web page http://www.utc.edu/irb or email us at: instrb@utc.edu .

Best wishes for a successful research project.

APPENDIX B

IRB APPROVAL LETTER FOR CHANGES IN INITIAL IRB

MEMORANDUM

TO: Catalina Morales Murillo   **IRB #14-095**
                Dr. Robin McWilliam

FROM: Lindsay Pardue, Director of Research Integrity
                Dr. Bart Weathington, IRB Committee Chair

DATE: October 1, 2014

SUBJECT: IRB #14-095: Psychometric Characteristics and Social Validity of the 3M Preschool
                Milestone Scale

The Institutional Review Board has reviewed and approved the following changes for the IRB project
listed below:

- Data collection center changed to La Alqueria in Valencia, Spain.

You must include the following approval statement on research materials seen by participants and used
in research reports:

> *The Institutional Review Board of the University of Tennessee at Chattanooga (FWA00004149) has
> approved this research project #14-095.*

Please remember that you must complete a Certification for Changes, Annual Review, or Project
Termination/Completion Form when the project is completed or provide an annual report if the project
takes over one year to complete. The IRB Committee will make every effort to remind you prior to your
anniversary date; however, it is your responsibility to ensure that this additional step is satisfied.

Please remember to contact the IRB Committee immediately and submit a new project proposal for
review if significant changes occur in your research design or in any instruments used in conducting the
study. You should also contact the IRB Committee immediately if you encounter any adverse effects
during your project that pose a risk to your subjects.

For any additional information, please consult our web page http://www.utc.edu/irb or email
instrb@utc.edu

Best wishes for a successful research project.

Page 1 of 1

APPENDIX C

3M PRESCHOOL MILESTONE SCALE

ENGLISH VERSION

3M Preschool Milestone Scale
(Age group 3-5 years old)
Catalina Morales Murillo
University of Tennessee at Chattanooga
R. A. McWilliam
Siskin Children's Institute

Date of Birth (MM/DD/YEAR): _____

Relationship of Person Completing this Form, to the Child: _____

Ethnicity: ( ) Caucasian ( ) African American ( ) Hispanic ( ) Other: _____

Age: _____ Time (in years) attending this center: _____ Does the child have a disability? __Yes __No   Type: _____

Instructions: Please rate the child's level of functioning in performing the following behaviors in each activity. If the child does not perform the behavior, circle 1 (NOT YET). If the child performs the behavior rarely, circle 2 (SELDOM). If the child performs the behavior several times but is not consistent, circle 3 (OFTEN). If the child performs the behavior almost all the time, circle 4 (MOSTLY ALWAYS).

CIRCLE the number that best represents the child's stage performing the behavior described in each statement.

Eating	Not yet	Seldom	Often	Mostly Always
1.1 Uses fork and spoon to stab and scoop food	1	2	3	4
1.2 Drinks from cup without spilling content	1	2	3	4
1.3 Initiates communication with peers	1	2	3	4
1.4 Clears table after eating (throws away trash/ puts away food containers) without been prompted.	1	2	3	4
1.5 Uses words, signs, and/or gestures to express needs to the teacher and peers	1	2	3	4

Free play	Not yet	Seldom	Often	Mostly Always
2.1 Engages in pretend play by acting out scenarios	1	2	3	4
2.2 Independently chooses and obtains accessible materials	1	2	3	4
2.3 Cooperates with peers while playing (e.g., negotiates roles)	1	2	3	4
2.4 Talks to peers using understandable language	1	2	3	4
2.5 Shows empathy towards other people's feelings	1	2	3	4

54

Toileting	Not yet	Seldom	Often	Mostly Always
3.1 Urinates in potty with no accidents	1	2	3	4
3.2 Washes his/her hands after using the potty	1	2	3	4
3.3 Uses zipper, snap, or buttons	1	2	3	4
3.4 Dresses and undresses without assistance	1	2	3	4
3.5 Goes into bathroom independently or asks for permission by using words or signs	1	2	3	4

Art	Not yet	Seldom	Often	Mostly Always
4.1 Responds to 3-step instructions from the teacher	1	2	3	4
4.2 Makes representational art (draws, paints, or builds things to look like real objects)	1	2	3	4
4.3 Uses scissors independently	1	2	3	4
4.4 Talks about his or her art product in full sentences	1	2	3	4
4.5 Waits for his or her turn to use materials without getting upset	1	2	3	4

Teacher-led (include circle time, morning meeting)	Not yet	Seldom	Often	Mostly Always
5.1 Attends to teacher when he or she is talking to the group for periods of time longer than 5 minutes	1	2	3	4
5.2 Participates in group activities that involve communication, by using full sentences	1	2	3	4
5.3 Jumps by lifting both feet from the ground	1	2	3	4
5.4 Imitates teacher's gestures and movements while singing songs	1	2	3	4
5.5 Follows rules and teacher's requests	1	2	3	4

Please scan and return this scale to Catalina-MoralesMurillo@mocs.utc.edu

APPENDIX D

3M PRESCHOOL MILESTONE SCALE

SPANISH VERSION

3M Escala de Pilares de Desarrollo en Preescolar
(Edades: 3-5 años)
Catalina Morales Murillo
University of Tennessee at Chattanooga
R. A. McWilliam
Siskin Children's Institute

Fecha de Nacimiento (Día/Mes/Año): _____

Relación de la Persona Completando este Instrumento con el Niño: _____

Edad del Niño:_____ Nacionalidad: _____

Encargados☐ Madre soltera☐ Padre Soltero☐ Padre y Madre☐ Otro: _____ Ingresos económicos mensuales: _____

Discapacidad: _____ Si o _____ No Tipo: _____

Instrucciones: Por favor indique el nivel de funcionamiento del niño cuando presenta los siguientes comportamientos en cada actividad. Si el niño no presenta el comportamiento, encierre con un círculo el 1 (Todavía No). Si el niño presenta el comportamiento de vez en cuando, encierre con un círculo el 2 (Raramente). Si el niño presenta el comportamiento varias veces, encierre con un círculo el 3 (Algunas Veces). Si el niño presenta el comportamiento casi todo el tiempo, encierre con un círculo el 4 (Casi Siempre).

ENCIERRE CON UN CIRCULO el número que mejor representa la etapa en la que el niño se encuentra al realizar el comportamiento descrito en cada oración.

1. Tiempos de Comida	Todavía No	Raramente	Algunas Veces	Casi Siempre
1.1 Come utilizando el tenedor y la cuchara	1	2	3	4
1.2 Bebe de un vaso sin derramar el contenido	1	2	3	4
1.3 Inicia conversaciones con compañeros y compañeras	1	2	3	4
1.4 Recoge la mesa (tira la basura y recoge los platos) por iniciativa propia	1	2	3	4
1.5 Usa palabras, señas, y /o gestos para expresar necesidades al maestro/a o compañeros (as)	1	2	3	4

2. Juego libre	Todavía No	Raramente	Algunas Veces	Casi Siempre
2.1 Participa en juego simbólico mediante la representación de escenarios	1	2	3	4
2.2 Independientemente selecciona y obtiene materiales accesibles	1	2	3	4
2.3 Colabora con compañeros y compañeras al jugar (e.g. negocia roles de juego)	1	2	3	4
2.4 Habla con compañeros (as) usando un lenguaje comprensible	1	2	3	4
2.5 Muestra empatía hacia los sentimientos de los demás	1	2	3	4

57

3. Utilizando el Baño	Todavía No	Raramente	Algunas Veces	Casi Siempre
3.1 Usa el inodoro sin accidentes	1	2	3	4
3.2 Se lava las manos después de usar el inodoro	1	2	3	4
3.3 Puede abrir y cerrar cremalleras, broches y botones	1	2	3	4
3.4 Se viste y desviste sin asistencia	1	2	3	4
3.5 Va al baño independientemente o pide permiso para ir al baño por medio de palabras o señas	1	2	3	4

4. Arte	Todavía No	Raramente	Algunas Veces	Casi Siempre
4.1 Responde a instruciones de tres pasos dadas por el/la maestro/a	1	2	3	4
4.2 Crea representaciones de objetos reales (dibuja, pinta, o construye cosas que son similares a objetos reales)	1	2	3	4
4.3 Usa tijeras independientemente	1	2	3	4
4.4 Habla acerca del projecto de arte que ha creado en oraciones completas	1	2	3	4
4.5 Espera, sin enfadarse, para utilizar materiales que otros (as) estén usando	1	2	3	4

5. Actividades dirigidas por la/el maestro(o) (actividad en un círculo, actividades en la mañana)	Todavía No	Raramente	Algunas Veces	Casi Siempre
5.1 Presta atención a la maestra/maestro durante más de 5 minutos cuando este le está hablando al grupo	1	2	3	4
5.2 Participa en las actividades grupales que requieren conversación, usando oraciones completas	1	2	3	4
5.3 Salta levantando ambos pies del suelo	1	2	3	4
5.4 Durante canciones, imita gestos y movimientos que el/la maestro (a) hace mientras cantan canciones	1	2	3	4
5.5 Sigue las reglas e instrucciones de la maestra/o	1	2	3	4

Por favor escanear este documento y enviarlo a Catalina-MoralesMurillo@mocs.utc.edu

APPENDIX E

SOCIAL VALIDITY SCALE (SPANISH)

Formulario Comprensibilidad / Escala de Validez Social
Catalina Morales Murillo
Universidad de Tennessee en Chattanooga
Robin McWilliam
Siskin Children's Institute

Instrucciones:
Por favor, responda las siguientes preguntas usando la escala Likert.

1. ¿Qué tan importante es tener información sobre el progreso del desarrollo de su estudiante?

No es importante	Algo importante	Importante	Muy importante
1	2	3	4

2. ¿Qué tan relevante es el contenido de esta escala en relación a su comprensión del desarrollo del niño?

No es relevante	Algo relevante	Relevante	Muy relevante
1	2	3	4

3. ¿Qué relevancia tienen las rutinas en esta escala en relación a sus rutinas de la clase?

No es importante	Algo importante	Importante	Muy importante
1	2	3	4

4. ¿Fue fácil completar esta escala?

No, en absoluto	Un poco fácil	Fácil	Muy fácil
1	2	3	4

5. ¿Qué posibilidades hay de utilizar esta escala en el futuro?

Nada probable	Algo probable	Probable	Muy probable
1	2	3	4

60

APPENDIX F

Q-SORT MATRIX FORM

Plantilla para la Organización de los Ítems del Menos Importante al Más Importante en la Detección de Desfases en el Desarrollo Funcional de los niños (as) De la 3M escala de Pilares de Desarrollo en Prescolar

Q-Sort

Catalina Morales Murillo

Universidad de Tennessee en Chattanooga

Robin McWilliam

Siskin Children's Institute

Por favor utilizando la 3M Escala de Pilares de Desarrollo en Prescolar, divida los items en la escala en dos grupos: el primer grupo con los items que usted considera son los menos importantes para detectar desfases en el desarrollo del niño (a) y el segundo grupo con los items que usted considera son los más importantes para detectar desfases en el desarrollo del niño (a).

1. Escriba el número del ítem sobre la línea en el círculo al cual usted considera pertenece el ítem (Por ejemplo: "1.1" para indicar que es el ítem: "Come utilizando el tenedor y la cuchara.", o "3.1" para indicar que es el ítem: "Usa el inodoro sin accidentes."). En uno de los círculos quedará una línea en blanco.

2. Escriba el número de los ítems que usted considera menos importantes en los cuadros hacia la izquierda y los más importantes en los cuadros hacia la derecha. Los cuadros en el centro de la figura pueden que tengan ítems que son menos y más importantes. El cuadro en la esquina izquierda debe contener el ítem menos importante de todos y el cuadro de la esquina derecha debe contener el ítem más importante de todos.

Ítems Menos Importantes ⟶ Ítems Más Importantes

APPENDIX G

FACTORIAL WEIGHTS FOR THE 3M 4 FACTOR SOLUTION

Factorial weights for the 3M 4 factor solution

	Factors			
	Engagement	Independence and Self-expression	Following Directions	Self-Help
Item 1.1	.537	.201	.199	.493
Item 1.2	.348	.048	.249	.599
Item 1.3	.766	.356	.199	.259
Item 1.4	.312	.169	.468	.457
Item 1.5	.795	.066	.190	.289
Item 2.1	.682	.472	.305	.179
Item 2.2	.728	.351	.339	.032
Item 2.3	.683	.450	.303	.159
Item 2.4	.672	.417	.231	.231
Item 2.5	.363	.560	.408	.063
Item 3.1	.359	.307	.032	.738
Item 3.2	-.089	.083	.185	.780
Item 3.3	.152	.762	.258	.269
Item 3.4	.179	.704	.228	.389
Item 3.5	.371	.428	-.062	.666
Item 4.1	.382	.617	.349	.051
Item 4.2	.348	.712	.280	.177
Item 4.3	.287	.818	.111	.105
Item 4.4	.563	.589	.298	.167
Item 4.5	.120	.493	.625	.166
Item 5.1	.256	.175	.808	.129
Item 5.2	.627	.387	.440	.199
Item 5.3	.503	.536	.159	.354
Item 5.4	.485	.270	.577	.170
Item 5.5	.335	.332	.738	.184

APPENDIX H

FACTORIAL WEIGHTS FOR TEACHERS' 3 FACTOR SOLUTION

Factorial weight for teachers' 3 factor solution

	Factors		
	Least experienced	More experienced	Most experienced
Teacher 2	.601	.409	-.037
Teacher 3	.137	.360	.707
Teacher 4	.328	-.123	-.579
Teacher 6	.227	.853	.338
Teacher 10	.395	.560	.116
Teacher 13	.719	.307	.348
Teacher 14	.901	.067	-.047
Teacher 15	.790	.174	-.010
Teacher 16	.214	.777	.287
Teacher 7	.440	.245	-.120
Teacher 5	.366	.808	.015
Teacher 19	.096	-.150	.616
Teacher 20	.325	.493	.052
Teacher 21	-.069	.693	-.319
Teacher 22	.705	.210	.457

VITA

Catalina Patricia Morales Murillo was born in Alajuela, Costa Rica, to parents Pablo Enrique Morales Vargas and Rita Irene Murillo Salazar. She is the oldest of a family of eleven children. Catalina did her elementary studies at María Vargas Rodríguez school, Ciruelas, Alajuela, and her high school at Liceo de Nicoya, Guanacaste, Costa Rica. In 2008, she moved to Rome, Georgia, United States, to pursue a degree in psychology and to learn English. In 2012, she graduated from Berry College, with a Bachelor's of Science in Psychology. Then in 2013, she began pursuing her Master's degree in Special Education at the University of Tennessee, Chattanooga. She currently serves as Senior Student Graduate Assistant at Siskin Children's Institute, where she is passionate about learning more about research and the Routines-Based Model.