Building and Standardizing a School Readiness Battery for Kindergarten Children in light of Classical Theory and Item Response Theory

(A Comparative Study)

MA Thesis Submitted for the Master Degree in Education

(Educational psychology)

By

Nada Abo Seif Mohammad Abo Seif

Demonstrator in Department of Educational Psychology

Under the Supervision of:

Prof.Dr. Ahmad Taha Mohammad
Professor of Educational PsychologyPsychology Ex-Dean ofthe Faculty of Education -Fayoum University

Dr. Aisha Ali Rafallah

Lecturer in Educational Faculty of Education Fayoum University

2018 G./ 1439 H.

Summary of the study

Firstly: The Thesis Title:

"Building and Standardizing a School Readiness Battery for Kindergarten Children in light of Classical Theory and Item Response Theory: A Comparative Study"

Secondly: The Aims of the Study: The study aims at:

1. Building a school readiness battery for kindergarten children.
2. Standardizing a school readiness test battery for kindergarten children according to the Classical Theory and the Item Response Theory.
3. The gradation of the school readiness test battery using Rasch's Model and the Partial Estimation Model.

Thirdly: Problem of the Study: the study problem was determined in building a school readiness battery for kindergarten children inclusive of:

1. The lingual abilities; it includes: (Phonological awareness– Identifying letters – Receptive vocabulary– Expressive Vocabulary).
2. Remembrance Abilities; including: The Rapid Automatic Naming (RAN) (the Working Memory Span – the Inhibition) as executive functions.
3. The numerical ability, inclusive of: (Counting and getting to know numbers and their orders); and standardizing it in the light of the Classical Theory and the Item Response Theory (A Comparative Study).

The problem was also determined in the following main question:

To what extent are the validity, reliability, gradation and the estimation of the individuals' abilities for each raw probable whole mark as well as the amount of information and the criteria of each of the readiness battery tests?

Under this main question come the following sub-questions:

1. To what extent are the validity, reliability, gradation and the estimation of the individuals' abilities for each raw probable whole mark as well as the amount of information and the criteria of the test of Phonological awareness according to the Classical Theory and the Item Response Theory?

2. To what extent are the validity, reliability, gradation and the estimation of the individuals' abilities for each raw probable whole mark as well as the amount of information and the criteria of the test of identifying the letters according to the Classical Theory and the Item Response Theory?

3. To what extent are the validity, reliability, gradation and the estimation of the individuals' abilities for each raw probable whole mark as well as the amount of information and the criteria of the test of Receptive vocabulary according to the Classical Theory and the Item Response Theory?

4. To what extent are the validity, reliability, gradation and the estimation of the individuals' abilities for each raw probable whole mark as well as the amount of information and the criteria of the test of Expressive Vocabulary according to the Classical Theory and the Item Response Theory?

5. To what extent are the validity, reliability, gradation and the estimation of the individuals' abilities for each raw probable whole mark as well as the amount of information and the criteria of the test of the Rapid Automatic Naming (RAN) according to the Classical Theory and the Item Response Theory?

6. To what extent are the validity, reliability, gradation and the estimation of the individuals' abilities for each raw probable whole mark as well as the amount of information and the criteria of the test of Working Memory span according to the Classical Theory and the Item Response Theory?

7. To what extent are the validity, reliability, gradation and the estimation of the individuals' abilities for each raw probable whole mark as well as the amount of information and the criteria of the Inhibition Test according to the Classical Theory and the Item Response Theory?

8. To what extent are the validity, reliability, gradation and the estimation of the individuals' abilities for each raw probable whole mark as well as the amount of information and the criteria of the Numerical Ability Test according to the Classical Theory and the Item Response Theory?

Fourthly, the Study Procedures, including:

1. **Method of study:** The Descriptive Method
2. **The Study Samples,** which are**:**
   a. The pilot samples: They are a group of 13 KG2 male and female children upon whom the study tools were applied to determine the time of each of the battery tests and to examine the degree of clarity of the application instructions of each test

as well as the suitability of the vocabulary items for the children and how clear and correctly phrased they are.

    b. The Basic Sample (Standardization): The basic sample consisted of 199 KG2 male and female children who were randomly chosen from a group of primary schools in Fayoum.

3. The Study Tools:

First: The Test of the Colored Successive Sets, prepared by Ravin (1956)

Second: The School Readiness Test Battery, prepared by the researcher

    <u>The Battery included (8) sub-tests, distributed as follows:</u>

Firstly: The Lingual Abilities Tests:

1. Test of Phonological awareness.
2. Test of identifying letters.
3. Test of Receptive vocabulary.
4. Test of Expressive Vocabulary.

Secondly: Tests of Memorization Abilities:

5. The Rapid Automatic Naming Test.
6. Working Memory Span Test.
7. The Inhibition Test.

Thirdly: The Numerical Ability Tests:

8. The Numerical Ability Test.

4. The Statistical Techniques:

In the Test Classical Theory:

1. The item's difficulty coefficient
2. The item's discrimination coefficient
3. The explorational factorial analysis
4. The validity of the vocabulary items (the test's internal consistency)
5. Reliability coefficients (Alpha Cronbach's Coefficient – Kuder-Richardson21 Coefficient).

In the Item Response Theory

1. Rasch's Model.
2. Partial Estimation Model.

The researcher used the statistical programs Winsteps 3.67, IRTPRO for and Windows 2.1 in conducting the necessary statistical processing of Rush's Model and the Partial Estimation Model in the Item Response Theory. The researcher also implemented IBM SPSS 22.0 in performing the necessary processes in the Classical Theory and preparing the files for the

analysis of the Item Response Theory, to answer the study's questions.

5.**Steps of Conducting the Study:** The researcher performed a set of procedures during this study that may be summed up as follows:

1. Perusal of previous studies in the field of the subject of the present study and writing the theoretical framework of the study.

2. Building a school readiness test battery.

3. Applying the school readiness battery in its primal form on 13 male and female children who represent the pilot sample on each child individually.

4. Presenting the school readiness battery on referees and modifying suggestions.

5. Applying the test of colored successive sets and the school readiness battery on 199 male and female children who represent the basic sample.

6. Inserting data and preparing the files of the statistical programs.

7. statistical analysis in the light of the classical Theory and the Item Response Theory.

8.Interpreting the results, setting recommendations and the suggested researches.

The most Important Findings of the Study:

1. There are high indicators of classical validity and reliability in the test of Phonological awareness; for the Alpha value was equal to (0.89) and the value of Kuder-Richardson (20) Coefficient was (0.872); also, (5) vocabulary items were deleted from the test, according to the classical indicators of validity and reliability. Also, the test of Phonological awareness has high indicators of validity and reliability, according to Rasch's Model; for the value of validity coefficient of Kuder-Richardson (20) for estimating abilities was (0.83) and (0.95) for estimating difficulties. Also, (6) vocabulary items were omitted from the test according to Rasch's Model. The test was graded and the extent of the difficulty covered by the items was from -2.38 to 1.61 logit; and the estimations of the individuals' abilities for each probable raw mark range from (-5.6) and (5.4) logit; and the account of the information amount in the test of Phonological awareness. At the end, the test criteria were made; through the account of the percentage ranks and the T-marks, according to the Classical Theory and Rasch's Model.

2. There are high indicators of classical validity and reliability in the test of identifying the letters; for the value of Alpha was equal to 0.743, and the value of Kuder-Richardson (20) Coefficient was (0.726); no vocabulary items were omitted from the test, according to the classical indicators of validity and reliability. The test of identifying letters has acceptable validity and reliability indicators, according to Rasch's Model. For the value of Kuder-Richardson (20) Coefficient for estimating abilities was (0.17) and (0.88) for estimating difficulties. no vocabulary items were omitted from the test, according to Rasch's Model. The test was graded and the extent of the difficulty covered by the items was from -0.77 to 1.71 logit; and the estimations of the individuals' abilities for each probable raw mark range from (-3.34) and (3.47) logit; and the account of the information amount in the test of identifying letters. At the end, the test criteria were made; through the account of the percentage ranks and the T-marks, according to the Classical Theory and Rasch's Model.

3. There are high indicators of classical validity and reliability in the test of Receptive vocabulary; for the value of Alpha was equal to 0.776, and the value of Kuder-Richardson (20) Coefficient was (0.781); no vocabulary items were omitted from the test, according to the classical indicators of validity and reliability. The test of Receptive vocabulary has acceptable validity and reliability indicators, according to Rasch's Model. For the value of Kuder-Richardson (20) Coefficient for estimating abilities was (0.64) and (0.94) for estimating difficulties. no vocabulary items were omitted from the test, according to Rush's Model. The test was graded and the extent of the difficulty covered by the items was from -4.15 to 1.83 logit; and the estimations of the individuals' abilities for each probable raw mark range from (-5.62) and (4.95) logit; and the account of the information amount in the test of. At the end, the test criteria were made; through the account of the percentage ranks and the T-marks, according to the Classical Theory and Rasch's Model.

4. There are high indicators of classical validity and reliability in the test of Expressive Vocabulary; for the value of Alpha was equal to 0.831, and the value of Kuder-Richardson (20) Coefficient was (0.702); (5) vocabulary items were omitted from the test, according to the classical indicators of validity and reliability. The test of Expressive Vocabulary has acceptable validity and reliability indicators, according to Rush's Model. For the value of Kuder-Richardson (20) Coefficient for estimating abilities was (0.79) and (0.98) for estimating difficulties. One vocabulary item was omitted from the test, according to Rasch's Model. The test was graded and the extent of the difficulty covered by the items was from (-6.38) and (4.07) logit; and the estimations of the

individuals' abilities for each probable raw mark range from (-6.9) to (6.28) logit; and the account of the information amount in the test of Expressive Vocabulary. At the end, the test criteria were made; through the account of the percentage ranks and the T-marks, according to the Classical Theory and Rasch's Model.

5. There are high indicators of classical validity and reliability in the test of Rapid Automatic Naming; for the value of Alpha was equal to 0.96; no vocabulary item was omitted from the test, according to the classical indicators of validity and reliability. The test of Rapid Automatic Naming has acceptable validity and reliability indicators, according to the Partial Estimation Model. For the value of marginal fixation coefficient for estimating responses was (0.96) and (0.96) for estimating the ability for each raw mark. No vocabulary item was omitted from the test, according to the partial estimation Model. The test was graded and the extent of the difficulty covered by the items was from (-0.03) and (0.22) logit; and the estimations of the individuals' abilities for each probable raw mark range from (-٢.824) to (3.039) logit; and the account of the information amount in the test of Rapid Automatic Naming. At the end, the test criteria were made; through the account of the percentage ranks and the T-marks, according to the Classical Theory and the partial estimation Model.

6. There are acceptable indicators of classical validity and reliability in the test of Working Memory span; for the value of Alpha was equal to 0.449; one vocabulary item was omitted from the test, according to the classical indicators of validity and reliability. The test of Working Memory span has relatively acceptable validity and reliability indicators, according to the Rasch's Model. For the value of marginal fixation coefficient for estimating responses was (0.27) and (0.265) for estimating the ability for each raw mark. No vocabulary item was omitted from the test, according to Rasch's Model. The test was graded and the extent of the difficulty covered by the items was from (-5.85) and (3.3) logit; and the estimations of the individuals' abilities for each probable raw mark range from (-2.33) to (1.285) logit; and the account of the information amount in the test of Working Memory span. At the end, the test criteria were made; through the account of the percentage ranks and the T-marks, according to the Classical Theory and the Rasch's Model.

7. There are high indicators of classical validity and reliability in the Inhibition Test; for the value of Alpha was equal to 0.925; the value of Kuder-Richardson fixation coefficient (20) was 0.921. No vocabulary item was omitted from the test, according to the classical indicators of validity and reliability. The Inhibition Test has relatively acceptable validity and reliability indicators, according to the Rasch's Model. For

the value of marginal fixation coefficient   for estimating abilities was (0.62) and (0.72) for estimating the difficulties. No vocabulary item was omitted from the test, according to Rush's Model. The test was graded and the extent of the difficulty covered by the items was from (-5.85) and (3.3) logit; and the estimations of the individuals' abilities for each probable raw mark range from (-1.94) to (1) logit; and the estimations of the individuals' abilities for each probable raw mark range from (-4.72) to (4.6) logit and the account of the information amount in the Inhibition Test. At the end, the test criteria were made; through the account of the percentage ranks and the T-marks, according to the Classical Theory and the Rasch's Model.

8. There are high indicators of classical validity and reliability in the Numerical Ability Test; for the value of Alpha was equal to 0.928; the value of Kuder-Richardson fixation coefficient (20) was 0.911. No vocabulary item was omitted from the test, according to the classical indicators of validity and reliability. The Numerical Ability Test has relatively high validity and reliability indicators, according to the Rush's Model. For the value of the value of Kuder-Richardson fixation coefficient (20) for estimating abilities was (0.80) and (0.97) for estimating the difficulties. Three vocabulary items were omitted from the test, according to Rasch's Model. The test was graded and the extent of the difficulty covered by the items was from (-2.54) and (4.1) logit; and the estimations of the individuals' abilities for each probable raw mark range from (-5.26) to (6.1) logit; and the account of the information amount in the Numerical Ability Test. At the end, the test criteria were made; through the account of the percentage ranks and the T-marks, according to the Classical Theory and the Rasch's Model.