

An Examination of Parametric and Nonparametric Dimensionality Assessment Methods with Exploratory and Confirmatory Mode

Hakan Koğar¹

¹ Faculty of Education, Department of Educational Sciences, Akdeniz University, Antalya, Turkey

Correspondence: Hakan Koğar, Faculty of Education, Department of Educational Sciences, Akdeniz University, Antalya, Turkey.

Received: February 13, 2018

Accepted: March 7, 2018

Online Published: March 20, 2018

doi:10.5539/jel.v7n3p148

URL: <https://doi.org/10.5539/jel.v7n3p148>

Abstract

The aim of the present research study was to compare the findings from the nonparametric MSA, DIMTEST and DETECT and the parametric dimensionality determining methods in various simulation conditions by utilizing exploratory and confirmatory methods. For this purpose, various simulation conditions were established based on number of dimensions, number of items, item discrimination levels, sample size and correlation between dimensions values. The performance of dimensionality determining methods based on MSA and factor analysis are similar, yet MSA is more effective in determining the number of dimensions. However, the method of DETECT has displayed a more powerful performance when compared with the other dimensionality methods. Particularly the confirmatory DETECT method could reveal the true dimensionality in conditions of both low discrimination and high discrimination methods. On the other hand, the exploratory DETECT method was affected by discrimination and, thus, could perform well only with high-discrimination items. In conditions where the exploratory dimensionality reduction methods are used to determine the number of dimensions, it is beneficial to confirm this structure by using confirmatory dimensionality reduction methods. For this purpose, using confirmatory DETECT is particularly recommended.

Keywords: unidimensionality, multidimensionality, mokken scaling, DIMTEST, DETECT, factor analysis

1. Introduction

It is of utmost importance to determine the reliability and validity of measurement tools in the fields of education and psychology. When the assumptions for such measurement theories as the Classical Test Theory (CTT) and the Item Response Theory (IRT), developed for this purpose, are examined, it can be maintained that the most outstanding of these assumptions is the concept of dimensionality. Camilli, Wang, & Fesq (1995) define dimensionality as the identification of the number of latent variables deriving from the correlations between the responses to the items within a defined dataset. Measurement theories such as CTT and IRT are developed based on the assumption of unidimensionality. Unidimensionality is defined as a test measuring a single ability and the responses being compatible with the rules of local independence (Nandakumar & Stout, 1993). Today, there are numerous unidimensional IRT models, while the number of multidimensional IRT theories is rather limited. There are two manifest important reasons underlying the preference for unidimensionality. The first reason is the possibility of obtaining a single total score in a unidimensional structure. The second reason is that with the increase in the number of parameters, multidimensional IRT models are more complicated when compared with unidimensional models (Van Abswoude, Van der Ark, & Sijstma, 2004).

As unidimensional IRT models have started to be frequently used in such areas as educational measurements, test equating studies and computerized adaptive tests, concerns related to dimensionality in tests have started to emerge increasingly (Jang & Roussos, 2007). The use of unidimensional IRT models in conditions incompatible with the local independence assumption leads to lower estimations owing to the requirement of standard error values in individuals' ability parameter estimations (Wainer & Wang, 2001). In addition, the traits measured in the fields of education and psychology are generally so complicated that they cannot be grouped under a single dimension. Hence, multidimensionality is considered to be an important concept. In simple-structured multidimensionality, items are explicitly distributed to each dimension to form item sets in which each item points to a latent trait (Stout, Habbing, Douglas, Kim, Roussos, & Zhang, 1996).

Linear factor analysis is traditionally used in analysing the dimensionality conditions of categorical or continuous datasets. Despite being popular, this method can lead to some problems especially when there is a variety in levels of difficulty or when it comes to correlations as regards dichotomously scored items (1-0) (Nandakumar & Stout, 1993). As for the tetrachoric correlation matrix, developed to resolve this problem, some other problems emerge, such as the unidentification of a positive definition and moving away from the normalization assumption (Knol & Berger, 1991). The purpose of the confirmatory factor analysis (CFA) is to evaluate whether or not the assumed correlations among a group of items are supported with the dataset. In CFA, generally the maximum likelihood (ML) technique is utilized as it is selected as the default in many software packages. However, using ML as an estimation technique for datasets in which there are few response categories and which do not show a normal distribution leads to subjectivity in factor loadings, standard errors, chi-square test statistics and goodness of fit indices (Hutchinson & Olmos, 1998).

As an alternative to this method, methods based on the nonparametric item response theory (NIRT) have been developed. NIRT defines the correlation between item pair responses and a continuous latent trait with a nonlinear model and can directly be applied to dichotomously scored items. Hence, there is no need for the tetrachoric correlation matrix. These methods are based on covariances of dichotomous items (Van Abswoude, Van der Ark, & Sijstma, 2004). Three different NIRT-based methods, which can determine dimensionality in two different ways as exploratory and confirmatory as in factor analysis, were selected in line with the purpose of the present study: Mokken Scale Analysis (MSA), Dimensionality Evaluation to Enumerate Contributing Traits (DETECT) and the Dimensionality Test (DIMTEST).

Mokken scale is an IRT model and is similar to the Rasch scaling technique, but it has far fewer assumptions than those existing in the Rasch scaling technique (Meijer, Sijstma, & Smid, 1990). In confirmatory MSA, a pre-determined item set and latent traits are tested in accordance with an identified c value. To this end, an item pair (H_{ij}) and a scalability coefficient (H) (Mokken, 1971) are used. In exploratory MSA, the H scalability coefficient is used in matching K number of item pairs and L number of unidimensional item sets. H_{ij} coefficients of item pairs with the highest positive significance are determined. Subsequently, the remaining items are selected so as to increase coefficient H to its highest level. This item selection method is repeated until there are no remaining items.

DETECT is a method which assigns items with a positive conditional covariance matrix in a multidimensional structured test to the same dimension and the items with a negative conditional covariance matrix to different dimensions. Thus, it enables the items to be structured in different sets. While confirmatory DETECT makes computations based on a structure defined by the user, exploratory DETECT continues with the analyses until the DETECT index reaches its highest point. The DETECT index is an estimation of the test's multidimensionality and is obtained with the distribution of the items to the sets (Jang & Roussos, 2007). There are simulation studies showing that, independent of how the structure is defined, it is DETECT that best defines the true dimensionality structure (Rousson & Ozbek, 2006; Zhang & Stout, 1999).

DIMTEST is a method which tests an item group can be modelled within a single dimension. Different from the other models, the number of dimensions in this model cannot be revealed. It is used to test the unidimensionality assumption. An assessment subtest (AT), comprised of similar items in terms of dimensionality, is defined. The other items that are not used in AT are grouped in the partitioning subtest (PT). The dimensional differences of these sets are determined by comparing the sets by various ways.

There is a limited number of studies on methods identifying dimensionality. In a study by Abswoude, Van der Ark, & Sijstma (2004), the findings obtained from the MSA, DETECT and hierarchical cluster analyses (HCA/CCPROX) were compared in various simulation conditions. It was reported that even in conditions where the correlation between traits is 0.80 and there are high discrimination items, there are cases when MSA cannot determine dimensionality. In such conditions, DETECT and HCA/CCPROX could succeed in determining the true dimensionality. Generally the DETECT method has displayed a better performance when compared with the other two methods. Stout et al. (1996) applied the DIMTEST, DETECT and HCA/CCPROX methods to the Law School Admission Test (LSAT). They found that the HCA/CCPROX method is the most effective method in determining the sets into which the dimensions are grouped. Erroneous estimations were made with the DIMTEST and DETECT methods. It was maintained that when these methods are used in combination, a near-to-real framework can be drawn as regards the dimensionality of the test.

The aim of the present research study was to compare the findings from the nonparametric MSA, DIMTEST and DETECT and the parametric dimensionality determining methods in various simulation conditions by utilizing exploratory and confirmatory methods.

2. Method

2.1 Data Simulation Procedures

Datasets compatible with the 2PL model, which is one of the parametric IRT models, were generated. The simulation conditions were made up of 128 cells: 2 (number of latent traits) x 4 (correlations between traits) x 4 (the number of items in the latent traits) x 2 (sample size) x 2 (levels of discrimination).

- *Number of latent traits:* 2 (L2) and 4 (L4) latent variables were identified.

- *The correlation between traits:* Various studies defining the correlation between traits were reviewed. The correlations between traits were identified to be 0.00, 0.25 and 0.50 in a study by Batley & Boss (1993), 0.2, 0.5 and 0.7 in a study by Jiang, Wang & Weiss (2016), 0.10, 0.40 and 0.70 in a study by Van Abswoude, Vermunt, Hemker & Van der Ark (2004); and 0.00, 0.20, 0.40, 0.60, 0.80 and 1.00 in a study by Van Abswoude, Van der Ark, & Sijstma (2004). Since the primary aim of the present study was to examine dimensionality in various conditions, the correlations among the latent traits were simulated to resemble those reported by Van Abswoude, Van der Ark, & Sijstma (2004) as 0.30, 0.70 and 1.00. The reason underlying this criterion is that it defines many strong and weak correlations with 0.00, indicating an independent latent trait and 1.00, indicating unidimensionality.

- *The number of items in latent traits:* With the consideration of the related literature, the number of items in each latent trait was determined as 5 or 20 so that the impact of both parametric and nonparametric techniques could be observed. The item numbers were distributed to the latent traits as short-short (5;5), short-long (5;20), long-short (20;5) and long-long (20;20). The L2 formula for the number of items was determined to be [2:5;5], [2:5;20], [2:20;5] and [2:20;20], while the L4 formula for the number of items was identified as [4:5;5;5;5], [4:5;5;20;20], [4:20;20;5;5] and [4:20;20;20;20]. These formulas explain the structure as follows: For example, in a structure with the formulation of [2:5;20], the “:” sign indicates the number of dimensions of the previous value. In the structure of this example, there are two dimensions. Then there is the item number in each dimension, and the number of items is separated from each other with the sign “;”. In this example, there are five items in the first dimension and 20 items in the second dimension.

- *Sample size:* It is observed that there are 200 samples in simulation studies in which nonparametric techniques are used (Van Abswoude, Van der Ark, & Sijstma, 2004; Van Abswoude et al. 2004). As the present research study adopted parametric techniques as well, a relatively large sample size of 500 participants took part in the study.

- *Levels of item discrimination:* The items in each latent trait were simulated as having low and high discrimination. The bold-faced figures indicate high discrimination items. The [2:5;5], [2:5;20], [2:20;5] and [2:20;20] formulas for L2 indicate the low discriminatory items, while the [2:**5**;5], [2:**5**;20], [2:**20**;5] and [2:**20**;20] formulas for L2 indicate the high discriminatory items. As for the L4 formulas, [4:5;5;5;5], [4:5;5;20;20], [4:20;20;5;5] and [4:20;20;20;20] indicate the low discriminatory items, while [4:**5**;5;5;5], [4:**5**;5;20;20], [4:**20**;20;5;5] and [4:**20**;20;20;20] indicate the high discriminatory items. With the consideration of studies in which low discrimination and high discrimination were conditions of simulation (Chen & Jiao, 2013; Desa, 2012; Dodeen, 2004; McBride, 1977; Olmuş, Nazman, & Erbaş, 2017; Roussos & Stout, 1996; Van Abswoude, van der Ark, & Sijstma, 2004), values were randomly selected from the normal distribution where the mean and variance were 0.5 and 0.1, respectively for low discrimination, and the mean and variance were 2.0 and 0.1, respectively for high discrimination. The difficulty parameter values were selected randomly from the distribution between the intervals -2 and +2.

The data sets were simulated via the MIRTGEN 2.0 software.

2.2 Data Analysis

Parametric (factor analysis) and nonparametric (Mokken Scale Analysis, DIMTEST and DETECT) dimensionality reduction methods were used for two different purposes of exploratory and confirmatory in determining dimensionality in various simulation conditions.

Mokken Scale Analysis (MSA): The coefficients for three different criteria were computed. The first criterion was the computation of the H_{ij} coefficients, which revealed the correlation between each $n-1$ item and the n^{th} item, based on the assumption that $n-1$ number of items maintained their place in the measurement instrument for exploratory Mokken scale analysis. If the coefficients of these pair-wise items were statistically significant with a value above 0, then it indicated that the first criterion is met. For confirmatory Mokken scale analysis, all the H_{ij} coefficients were computed. The coefficients of these pair-wise items were aimed to be statistically significant with a value above 0. The second criterion was the expectation of the H_j coefficients obtained for

every item to be above the criterion value $c=0.3$ (Van der Ark, Croon, & Sijstma, 2008). This criterion value was used as the lowest bottom limit (Zijlstra, Van der Ark, & Sijstma, 2011). Finally, the third criterion was the H coefficient, which showed the power of the test. It was based on the alternative $H > c$ hypothesis. In the present study, $c=0.3$ was identified as the bottom limit, so the alternative $H > 0.3$ hypothesis was analyzed. An H coefficient between the 0.3-0.4 interval showed a weak level of scalability, an H coefficient between the 0.4-0.5 interval showed a moderate level of scalability and an H coefficient of 0.5 or above indicated a strong level of scalability (Van der Ark et al., 2008). The Mokken 2.8.9 package in the R software was used.

Dimensionality Test (DIMTEST): It is a nonparametric technique testing whether or not a dataset is unidimensional. Analyses are conducted based on three different datasets: DIMTEST, assessment subtest (AT) and partitioning subtest (PT). As a method of exploratory dimensionality reduction method, in AT, from among J number of items, M number of items with similar traits are grouped together. For this purpose, the correlation coefficients obtained from the tetrachoric correlation matrix are used. As a confirmatory dimensionality reduction method, the items measuring the same trait are defined as AT. PT is composed of J-M—the remaining items (Van Abswoude, Van der Ark, & Sijstma, 2004). In this technique, the significance of the DIMTEST statistics (T) is tested. The T value which is determined to be statistically significant indicates the multidimensionality of the related dataset (Leighton, Gokiart, & Cui, 2007). For this purpose, the DIMTEST 1.0 (William Stout Institute for Measurement, 2006) software was used.

Dimensionality Evaluation to Enumerate Contributing Traits (DETECT): DETECT is a nonparametric technique that makes grouping analysis using the conditional covariance matrix. It is regarded as one of the most powerful techniques in determining dimensionality. The r index, which determines whether or not a dataset has a simple, multidimensional structure, is computed. If this value is 0.80 or above, it is considered to be a strong indication that the related dataset has a simple structure with numerous variables. In the analysis of the dimensionality, the DETECT index is computed. If this value is below 0.20, it indicates either unidimensionality or a weak multidimensionality. A value between 0.20-0.40 indicates between weak multidimensionality to moderate level multidimensionality, a value between 0.40-1.00 indicates between moderate level of multidimensionality to strong multidimensionality, and 1.00 or above shows a strong multidimensionality (Rousson & Ozbek, 2006; Yavuz & Doğan, 2015; Zhang & Stout, 1999). In the present study, conditions where the DETECT index was 0.1 or lower were defined as unidimensionality (Zhang & Stout, 1999). For this purpose, the DIMTEST 1.0 (William Stout Institute for Measurement, 2006) software was used.

Factor Analysis (FA): As an exploratory dimensionality reduction method, the Parallel Analysis was used. After the number of dimensions were identified, the tetrachoric correlation matrix was used to identify the number of items in the dimensions yielded by the exploratory factor analysis. As a confirmatory dimension reduction method, the asymptotic covariance matrix was used as an attempt to confirm the defined structured with the Unweighted Least Squares. For these purposes, the psych 1.7 and lavaan 0.5 packages in the R software were used.

3. Results

The findings obtained from each dimensionality reduction methods were reported. Subsequently, a comparative analysis of the findings were conducted.

According to the exploratory MSA findings (Table 1), in high discrimination items, a unidimensional structure was obtained in conditions where the correlation between dimensions was 0.70 and above. In conditions where the correlation between dimensions was 0.00 and 0.30, findings that were generally very close to the true dimensionality were obtained. In conditions where the correlation between dimensions is 1.00, in both sample sizes, the same findings as those of true dimensionality were obtained. In low discrimination items, the correlation among the dimensions was found to be 1.00, and in four dimensional situations, a unidimensional structure was obtained. In a two-dimensional structure, there was a direct correlation between sample size and number of dimensions. The findings regarding dimensionality in two-dimensional structures resemble the true dimensionality when compared with four-dimensional structures. According to the confirmatory MSA findings (Table 1), all the findings regarding the high discrimination items reflect the true dimensionality. It was found that the number of scalability items was highly limited in low discrimination items. Generally a unidimensional structure is revealed, while in the two-dimensional structure a maximum of two dimensions and in a four-dimensional structure, a maximum of four dimensions are obtained. Scalability could not be achieved in conditions where the correlation between dimensions was 1.00. Only 20% of the simulation conditions identified the number of dimensions accurately; none of the simulation conditions could reflect the true dimensionality structure.

Table 1. Exploratory and confirmatory MSA dimensionality findings

S	NI & ID	Correlation Between Dimensions							
		Exploratory				Confirmatory			
		0.00	0.30	0.70	1.00	0.00	0.30	0.70	1.00
200	[2:5;5]	1:2	1:2	2:4;2	1:3	0:0	1:2	1:1	1:1
	[2:5;20]	3:8;2;2	1:9	2:3;2	1:11	2:1;4	1:12	0:0	1:1
	[2:20;5]	1:6	2:9;2	1:9	1:11	1:1	1:1	2:1;1	1:1
	[2:20;20]	2:6;6	2:10;9	2:13;2	3:10;2;2	2:1;2	2:4;5	0:0	1:1
	[2:5;5]	2:5;5	1:10	1:10	1:10	2:5;5	2:5;5	2:5;5	2:5;5
	[2:5;20]	2:5;20	2:5;20	1:25	1:25	2:5;20	2:5;20	2:5;20	2:5;20
	[2:20;5]	2:20;5	2:20;5	1:25	1:25	2:20;5	2:20;5	2:20;5	2:20;5
	[2:20;20]	2:20;20	2:20;20	1:40	1:40	2:20;20	2:20;20	2:20;20	2:20;20
	[2:5;5]	1:2	0:0	1:5	1:5	0:0	0:0	1:1	2:1;1
	[2:5;20]	3:7;3;3	3:7;2;2	2:9;2	2:10;2	2:2;1	1:1	1:1	0:0
500	[2:20;5]	2:8;2	2:8;3	4:7;2;3;2	1:9	2:1;1	2:1;1	1:1	1:1
	[2:20;20]	2:8;8	4:7;9;2;2	2:12;2	3:14;2;2	1:2	2:1;1	2:1;1	1:1
	[2:5;5]	2:5;5	1:10	1:10	1:10	2:5;5	2:5;5	2:5;5	2:5;5
	[2:5;20]	2:5;20	1:25	1:25	1:25	2:5;20	2:5;20	2:5;20	2:5;20
	[2:20;5]	2:20;5	2:21;4	1:25	1:25	2:20;5	2:20;5	2:20;5	2:20;5
	[2:20;20]	2:20;20	1:40	1:40	1:40	2:20;20	2:20;20	2:20;20	2:20;20
	[4:5;5;5;5]	1:3	0:0	2:2;2	1:20	0:0	1:1	2:1;2	-
	[4:5;5;20;20]	3:12;6;2	2:7;5	2:13;2	1:50	3:2;10;1	3:2;3;2	1:1	-
	[4:20;20;5;5]	2:6;2	3:8;2;7	2:13;3;2	1:50	2:1;1	4:1;2;2;1	2:3;2	-
	[4:20;20;20;20]	2:8;5	2:7;8	3:13;6;8	1:80	3:1;2;2	4:2;1;1;1	3:3;1;1	-
200	[4:5;5;5;5]	4:5;5;5;5	2:10;10	1:20	1:20	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5
	[4:5;5;20;20]	4:5;5;20;20	3:24;20;5	1:50	1:50	4:5;5;20;20	4:5;5;20;20	4:5;5;20;20	4:5;5;20;20
	[4:20;20;5;5]	4:20;20;5;5	3:25;20;4	1:50	1:50	4:20;20;5;5	4:20;20;5;5	4:20;20;5;5	4:20;20;5;5
	[4:20;20;20;20]	4:20;21;20;19	4:38;20;19;2	1:80	1:80	4:20;20;20;20	4:20;20;20;20	4:20;20;20;20	4:20;20;20;20
	[4:5;5;5;5]	1:3	2:2;2	2:3;2	1:20	2:1;1	1:1	1:2	-
	[4:5;5;20;20]	2:7;8	3:8;9;2	4:12;2;2;3	1:50	2:1;1	1:1	1:2	-
	[4:20;20;5;5]	2:8;8	2:9;9	2:14;2	1:50	1:2	1:2	2:2	-
	[4:20;20;20;20]	4:6;8;8;8	5:7;9;6;5;2	4:18;8;3;2	1:80	4:1;3;1;1	3:1;1;1	3:2;1;1	-
	[4:5;5;5;5]	4:5;5;5;5	2:10;10	1:20	1:20	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5
	[4:5;5;20;20]	4:5;5;20;20	3:25;20;5	1:50	1:50	4:5;5;20;20	4:5;5;20;20	4:5;5;20;20	4:5;5;20;20
500	[4:20;20;5;5]	4:20;20;5;5	4:20;20;5;5	1:50	1:50	4:20;20;5;5	4:20;20;5;5	4:20;20;5;5	4:20;20;5;5
	[4:20;20;20;20]	4:20;20;20;20	3:40;20;20	1:80	1:80	4:20;20;20;20	4:20;20;20;20	4:20;20;20;20	4:20;20;20;20

Note. SS: Sample Size, NI & ID: Number of Item & Item Discrimination, boldface indicates highly discriminating items.

According to the exploratory DETECT findings (Table 2), in low discrimination items, the true dimensionality structured as two dimensions could not be revealed in any of the simulation conditions. The number of dimensions obtained varied between three and five. A direct correlation was observed between sample size and the estimated number of dimensions; that is, as the sample size became larger, the number of dimensions increased. As for the true dimensionality structured as four dimensions, it was observed that the true dimensionality was revealed in most of the simulation conditions. There was also an increase in the total number of scaled items. It was found that in low discrimination items, a higher number of dimensionality estimations were made in an correlation between dimensions of 0.70 and above in small samples and an correlation between dimensions of 1.00 in large samples. This situation is far from being an estimation of a unidimensional structure. As for the true dimensionality structured in four dimensions, unidimensionality was found in an correlation between dimensions of 1.00 and in the other conditions of correlation between dimensions, it was found that a structure resembling the true dimensionality could be revealed. According to the DETECT findings (Table 2), it was observed that the impact of a low or high discrimination had a low impact on the findings. DETECT was the least affected method by discrimination. In all simulation conditions, unidimensionality was identified in correlation between dimensions of 1.00 and in the other conditions, it was observed that structures resembling the true dimensionality could be revealed. This points to the power of the confirmatory DETECT.

Table 2. Exploratory and confirmatory DETECT dimensionality findings

S S	Correlation Between Dimensions							
	Exploratory				Confirmatory			
	0.00	0.30	0.70	1.00	0.00	0.30	0.70	1.00
[2:5;5]	5:1;2;2;1;4	4:1;2;3;4	3:3;4;3	4:2;2;3;3	2:5;5	2:5;5	1:10	1:10
[2:5;20]	4:7;2;11;5	4:2;7;6;10	4:8;11;2;4	4:4;5;8;8	2:5;20	2:5;20	1:25	1:25
[2:20;5]	5:5;8;5;1;6	3:8;6;11	4:6;10;5;4	5:7;4;6;5;3	2:20;5	2:20;5	1:25	1:25
²⁰⁰ [2:20;20]	5:3;15;4;16;2	3:20;4;16	3:15;11;14	4:12;9;15;4	2:20;20	2:20;20	1:40	1:40
[2:5;5]	2:5;5	2:5;5	2:5;5	2:7;3	2:5;5	2:5;5	2:5;5	1:10
[2:5;20]	2:5;20	2:5;20	2:6;19	3:9;8;8	2:5;20	2:5;20	2:5;20	1:25
[2:20;5]	2:20;5	2:20;5	3:13;7;5	3:11;7;7	2:20;5	2:20;5	2:20;5	1:25
[2:20;20]	2:20;20	2:20;20	2:20;20	4:13;9;8;10	2:20;20	2:20;20	2:20;20	1:40
[2:5;5]	4:2;2;2;4	4:2;2;3;3	5:2;1;4;2;1	4:3;4;2;1	2:5;5	2:5;5	1:10	1:10
[2:5;20]	4:6;3;14;2	5:7;3;12;1;2	4:8;12;3;2	5:5;5;8;4;3	2:5;20	2:5;20	1:25	1:25
[2:20;5]	3:6;13;6	5:4;3;11;1;6	5:2;9;7;5;2	4:10;6;5;4	2:20;5	2:20;5	2:20;5	1:25
⁵⁰⁰ [2:20;20]	4:19;16;3;2	4:4;18;3;15	5:6;11;12;4;7	5:16;6;6;8;4	2:20;20	2:20;20	2:20;20	1:40
[2:5;5]	2:5;5	2:5;5	2:5;5	2:5;5	2:5;5	2:5;5	2:5;5	1:10
[2:5;20]	2:5;20	2:5;20	2:5;20	3:6;8;11	2:5;20	2:5;20	2:5;20	1:25
[2:20;5]	2:20;5	2:20;5	2:20;5	3:9;7;9	2:20;5	2:20;5	2:20;5	1:25
[2:20;20]	2:20;20	2:20;20	2:20;20	4:12;9;13;6	2:20;20	2:20;20	2:20;20	1:40
[4:5;5;5;5]	5:2;4;6;4;4	5:3;3;6;3;5	4:6;4;6;4	1:20	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5	1:20
[4:5;5;20;20]	4:22;7;20;1	4:9;10;18;13	4:8;10;16;16	1:50	4:5;5;20;20	4:5;5;20;20	4:5;5;20;20	1:50
[4:20;20;5;5]	4:5;16;10;19	3:11;17;22	4:19;17;7;7	1:50	4:20;20;5;5	4:20;20;5;5	4:20;20;5;5	1:50
²⁰⁰ [4:20;20;20;20]	4:16;25;22;17	4:25;20;19;16	5:20;12;13;19;16	1:80	4:20;20;20;20	4:20;20;20;20	1:80	1:80
[4:5;5;5;5]	4:5;5;5;5	4:5;5;5;5	3:5;10;5	1:20	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5	1:20
[4:5;5;20;20]	3:5;25;20	3:10;20;20	3:10;20;20	1:50	4:5;5;20;20	4:5;5;20;20	4:5;5;20;20	1:50
[4:20;20;5;5]	4:20;20;5;5	3:20;20;10	3:20;20;10	1:50	4:20;20;5;5	4:20;20;5;5	4:20;20;5;5	1:50
[4:20;20;20;20]	4:20;20;20;20	4:20;20;20;20	4:20;20;20;20	1:80	4:20;20;20;20	4:20;20;20;20	4:20;20;20;20	1:80
[4:5;5;5;5]	5:1;6;6;2;5	4:4;6;4;6	5:4;1;4;5;6	1:20	4:5;5;5;5	4:5;5;5;5	1:10	1:20
[4:5;5;20;20]	4:10;3;19;18	4:8;4;21;17	4:18;19;6;7	1:50	4:5;5;20;20	4:5;5;20;20	4:5;5;20;20	1:50
[4:20;20;5;5]	4:2;22;19;7	5:21;2;2;18;7	4:8;9;19;14	1:50	4:20;20;5;5	4:20;20;5;5	4:20;20;5;5	1:50
⁵⁰⁰ [4:20;20;20;20]	4:18;22;22;18	5:20;15;23;20;2	4:10;18;23;29	1:80	4:20;20;20;20	4:20;20;20;20	1:80	1:80
[4:5;5;5;5]	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5	1:20	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5	1:20
[4:5;5;20;20]	3:10;20;20	4:5;5;20;20	3:10;20;20	1:50	4:5;5;20;20	4:5;5;20;20	4:5;5;20;20	1:50
[4:20;20;5;5]	3:20;20;10	3:20;20;10	3:20;20;10	1:50	4:20;20;5;5	4:20;20;5;5	4:20;20;5;5	1:50
[4:20;20;20;20]	4:20;20;20;20	4:20;20;20;20	4:20;20;20;20	1:80	4:20;20;20;20	4:20;20;20;20	4:20;20;20;20	1:80

Note. SS: Sample Size, NI & ID: Number of Item & Item Discrimination, boldface indicates highly discriminating items.

In the exploratory and confirmatory DIMTEST findings (Table 3), dimensionality values of 10-item structures were not obtained. The software that was utilized could run analyses with a minimum of 15 PT items and four AT items. Thus, structures with the number of items below 19 cannot be tested (Fay, 2012). According to the exploratory and confirmatory DIMTEST findings, generally unidimensionality was found in correlation between dimension of 1.00, yet independent of item discrimination, multidimensional structures were observed in the other simulation conditions. This situation shows that DIMTEST is an important method in identifying whether or not a dataset is unidimensional, without being impacted by discrimination.

Table 3. Exploratory and confirmatory DIMTEST dimensionality findings

SS	NI & ID	Correlation Between Dimensions							
		Exploratory				Confirmatory			
		0.00	0.30	0.70	1.00	0.00	0.30	0.70	1.00
200	[2:5;5]	-	-	-	-	-	-	-	-
	[2:5;20]	X	+	X	+	X	+	+	+
	[2:20;5]	X	X	X	X	+	X	X	+
	[2:20;20]	X	X	+	X	X	X	X	+
	[2:5;5]	-	-	-	-	-	-	-	-
	[2:5;20]	X	X	X	X	X	X	X	+
	[2:20;5]	X	X	X	+	X	X	X	+
	[2:20;20]	X	X	X	X	X	X	X	+
500	[2:5;5]	-	-	-	-	-	-	-	-
	[2:5;20]	X	X	X	+	X	X	+	+
	[2:20;5]	X	X	X	+	X	X	X	X
	[2:20;20]	X	X	X	X	X	X	X	+
	[2:5;5]	-	-	-	-	-	-	-	-
	[2:5;20]	X	X	X	X	X	X	X	+
	[2:20;5]	X	X	X	X	X	X	X	+
	[2:20;20]	X	X	X	X	X	X	X	+
200	[4:5;5;5;5]	X	+	+	+	X	X	+	+
	[4:5;5;20;20]	X	X	X	+	X	X	X	+
	[4:20;20;5;5]	X	X	X	+	X	X	X	+
	[4:20;20;20;20]	X	X	X	+	X	X	X	+
	[4:5;5;5;5]	X	X	X	+	X	X	X	+
	[4:5;5;20;20]	X	X	X	+	X	X	X	+
	[4:20;20;5;5]	X	X	X	+	X	X	X	+
	[4:20;20;20;20]	X	X	X	+	X	X	X	+
500	[4:5;5;5;5]	X	+	+	+	X	X	+	+
	[4:5;5;20;20]	X	X	X	+	X	X	X	+
	[4:20;20;5;5]	X	X	X	+	X	X	X	+
	[4:20;20;20;20]	X	X	X	+	X	X	X	+
	[4:5;5;5;5]	X	X	X	+	X	X	X	+
	[4:5;5;20;20]	X	X	X	+	X	X	X	+
	[4:20;20;5;5]	X	X	X	+	X	X	X	+
	[4:20;20;20;20]	X	X	X	+	X	X	X	+

Note. SS: Sample Size, NI & ID: Number of Item & Item Discrimination, - indicates no scaling, + indicates unidimensionality, X indicates multidimensionality, boldface indicates highly discriminating items.

According to the exploratory factor analysis findings, in true dimensionality structured as two-dimensional with low discrimination items, a small sample size and an correlation between dimensions of 0.70 and above, findings similar to the true dimensionality could be obtained. However, in other conditions dimensionality values ranging between three to five dimensions were obtained. As sample size increased, so did the number of dimensions increased too. In true dimensionality structured as four-dimensional with a large sample size, it could be observed that true dimensionality could be obtained in some conditions. In correlation between dimension of 1.00, the true unidimensionality was identified. In high discrimination items, it was found that a two-dimensional structure was estimated as being three- or four-dimensional, while the four-dimensional structure could reveal the true dimensionality in some simulation conditions. According to the confirmatory factor analysis findings, it was observed that in low discrimination items, scalability could not be realized in most simulation conditions, and in conditions where scalability was possible, a unidimensional structure was revealed. However, with high discrimination items in true two-dimensional structures, it was found that the defined dimensionality values (including the correlation between dimensions of 1.00) were obtained to reflect the true dimensionality. However, a unidimensional structure was found in a four-dimensional structure with an correlation between dimensions of 1.00. In the other simulation conditions, the true dimensionality was found to be revealed.

Table 4. Exploratory and confirmatory factor analysis dimensionality findings

S	NI & ID	Correlation Between Dimensions							
		Exploratory				Confirmatory			
		0.00	0.30	0.70	1.00	0.00	0.30	0.70	1.00
	[2:5;5]	-	-	5:1;1;1;3;2	1:3	-	-	-	-
	[2:5;20]	5:5;2;2;2;2	4:6;3;2;1	2:8;4	2:7;4	-	-	-	-
	[2:20;5]	3:7;1;4	3:4;6;2	1:9	1:11	1:3	1:5	1:5	2:10;5
200	[2:20;20]	5:5;5;5;2;7	3:7;8;2	2:10;6	1:15	1:9	-	-	-
	[2:5;5]	4:4;2;3;1	3:5;3;2	3:5;3;2	2:6;4	2:5;5	2:5;5	2:5;5	2:5;5
	[2:5;20]	3:5;11;9	3:10;10;5	3:11;10;4	2:15;10	2:5;25	2:5;25	2:5;25	2:5;25
	[2:20;5]	3:11;9;5	3:10;10;5	3:13;9;3	2:13;12	2:20;5	2:20;5	2:20;5	2:20;5
	[2:20;20]	4:13;11;8;8	4:12;12;8;8	4:13;13;8;6	2:16;18	2;20;20	2;20;20	2;20;20	2;20;20
	[2:5;5]	2:1;2	6:1;1;1;1;1	1:3	4:1;1;3;1	-	1:4	1:5	-
	[2:5;20]	3:6;4;3	6:5;2;2;2;3	5:4;3;2;2;1	2:8;3	-	1:5	-	-
	[2:20;5]	3:4;3;2	5:5;3;1;2;2	3:5;4;1	2:5;3	-	1:5	-	1:5
500	[2:20;20]	3:9;5;3	2:7;6	5:7;6;2;1;1	3:12;2;1	-	-	-	1:13
	[2:5;5]	3:5;3;2	3:5;3;2	3:5;3;2	2:6;3	2:5;5	2:5;5	2:5;5	2:5;5
	[2:5;20]	3:12;7;6	3:12;8;5	3:8;12;5	2:13;10	2:5;25	2:5;25	2:5;25	2:5;25
	[2:20;5]	3:12;7;5	3:10;8;5	3:12;8;5	2:11;11	2:20;5	2:20;5	2:20;5	2:20;5
	[2:20;20]	4:11;12;8;8	4:8;8;9;11	4:11;12;8;9	2:24;16	2;20;20	2;20;20	2;20;20	2;20;20
	[4:5;5;5;5]	3:2;1;2	5:1;3;2;1;1	1:4	1:20	-	-	-	-
	[4:5;5;20;20]	6:7;5;2;1;2;4	5:6;7;2;3;1	1:18	1:50	-	2:5;13	-	-
	[4:20;20;5;5]	3:7;3;3	6:8;5;2;3;2;1	3:8;8;2	1:50	1:3	1:7	-	-
200	[4:20;20;20;20]	5:10;9;8;7;6	5:9;7;5;8;3	3:12;8;5	1:80	-	-	-	-
	[4:5;5;5;5]	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5	1:20	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5	1:20
	[4:5;5;20;20]	4:5;5;20;20	3:10;20;20	2:25;25	1:50	4:5;5;20;20	4:5;5;20;20	4:5;5;20;20	1:50
	[4:20;20;5;5]	4:20;20;5;5	3:20;20;10	2:25;25	1:50	4:20;20;5;5	4:20;20;5;5	4:20;20;5;5	1:50
	[4:20;20;20;20]	4:20;20;20;20	4:20;20;20;20	3:20;40;20	1:80	4:20;20;20;20	4:20;20;20;20	4:20;20;20;20	1:80
	[4:5;5;5;5]	4:3;2;3;1	5:1;4;1;1;1	3:3;2;1	1:20	-	-	-	-
	[4:5;5;20;20]	5:8;7;3;2;1	3:9;7;2	2:9;8	1:50	1:1	-	-	-
	[4:20;20;5;5]	3:7;8;1	3:8;9;1	4:7;5;2;1	1:50	-	-	-	-
500	[4:20;20;20;20]	4:7;8;5;7	4:8;10;7;7	3:9;8;6	1:80	-	-	-	-
	[4:5;5;5;5]	4:5;5;5;5	4:5;5;5;5	3:9;5;6	1:20	4:5;5;5;5	4:5;5;5;5	4:5;5;5;5	1:20
	[4:5;5;20;20]	3:5;20;50	3:5;20;50	2:25;25	1:50	4:5;5;20;20	4:5;5;20;20	4:5;5;20;20	1:50
	[4:20;20;5;5]	4:20;20;5;5	3:20;20;9	2:22;28	1:50	4:20;20;5;5	4:20;20;5;5	4:20;20;5;5	1:50
	[4:20;20;20;20]	4:20;20;20;20	4:20;20;20;20	2:40;40	1:80	4:20;20;20;20	4:20;20;20;20	4:20;20;20;20	1:80

Note. SS: Sample Size, NI & ID: Number of Item & Item Discrimination, boldface indicates highly discriminating items.

The findings in various simulation conditions of dimensionality are summarized in Table 5. It was accepted that the true dimensionality in an correlation between dimensions of 1.00 would be unidimensional. Findings related to three different situations are addressed: situations in which the number of dimensions could not be found, situations in which the number of dimensions was identified but all the items could not be scaled under the related dimension, and situations where the true dimensionality was met. In each exploratory and confirmatory conditions, 64 different simulation conditions were examined. For exploratory analyses with low discrimination, an equal number of true dimensionality estimations were made for all the methods. In terms of identifying the number of dimensions accurately, MSA and DETECT have been identified to be more effective when compared with the factor analysis methods. For confirmatory analyses with low discrimination, the DETECT method stands out to have a high performance. True dimensionality was obtained in 52 of the 64 conditions. The other methods could not reveal the true dimensionality in any condition. In exploratory analyses with high discrimination, it was found that MSA and DETECT were more effective in revealing the true dimensionality when compared to the factor analysis methods. In confirmatory analyses with high confirmation, the DETECT method could reveal in all conditions the true dimensionality. It can be claimed that the methods of factor analysis and MSA can reveal the true dimensionality to a high degree too.

Table 5. A summary of dimensionality findings

		Low discrimination		High discrimination	
		Exploratory	Confirmatory	Exploratory	Confirmatory
MSA	Analyzes failed	-	8	-	-
	Correct number of dimensionality	19	17	3	-
	True dimensionality	8	-	35	48
DETECT	Analyzes failed	-	-	-	-
	Correct number of dimensionality	16	-	1	-
	True dimensionality	8	52	44	64
Factor Analysis	Analyzes failed	2	49	-	-
	Correct number of dimensionality	11	-	-	-
	True dimensionality	8	-	20	56

4. Conclusions

Parametric (factor analysis methods) and nonparametric (MSA, DIMTEST and DETECT) dimensionality reduction methods, which could test dimensionality as exploratory and confirmatory in various simulation conditions, were compared. For this purpose, various simulation conditions were established based on number of dimensions, number of items, item discrimination levels, sample size, and correlation between dimensions values. Nonparametric dimension reduction methods are frequently used as mentioned in the related literature; however, there are no or very few studies comparing parametric and nonparametric methods. Furthermore, there is a limited number of studies on confirmatory dimensionality reduction methods.

It was concluded that item discrimination has a great impact on determining dimensionality. It has a serious effect particularly on MSA and factor analysis as dimensionality reduction methods. It was also concluded that only with an increase in discrimination did the likelihood of true dimensionality estimations increase. In dimensionality reduction methods based on DETECT and factor analysis with conditions where the number of dimensions was identified as four, a higher rate of true dimensionality was found. The higher the number of dimensions, the higher the possibility of making an accurate estimation. The impact of the sample size, however, is at a limited level.

Another conclusion is that different from the other methods, instead of identifying the number of dimensions, DIMTEST, which can reveal unidimensionality and multidimensionality without being affected by item discrimination, in 0.70 and lower correlation between dimensions conditions. In addition, it can determine unidimensionality in conditions with an correlation between dimensions of 1.00. In research studies where the purpose is not to determine the number of dimensions, but to test the unidimensionality assumption or whether or not there is a homogeneous dataset, it is believed that DIMTEST would be an appropriate test to use.

The dimensionality determining performance of methods based on MSA and factor analysis are similar, yet MSA is more effective in determining the number of dimensions. However, the method of DETECT has displayed a more powerful performance when compared with the other dimensionality methods. Particularly the confirmatory DETECT method could reveal the true dimensionality in conditions of both low discrimination and high discrimination methods. On the other hand, the exploratory DETECT method was affected by discrimination and, thus, could perform well only with high-discrimination items.

Finally, it was also concluded that confirmatory methods, when compared with exploratory methods, could reveal the true dimensionality at a higher degree. In terms of theory or practice, when there is information about the structure of a dataset, it is recommended that instead of the exploratory dimensional methods, the confirmatory dimensionality reduction methods should directly be used. In conditions where the exploratory dimensionality reduction methods are used to determine the number of dimensions, it is beneficial to confirm this structure by using confirmatory dimensionality reduction methods. For this purpose, using confirmatory DETECT is particularly recommended.

References

- Batley, R. M., & Boss, M. W. (1993). The effects on parameter estimation of correlated dimensions and a distribution-restricted trait in a multidimensional item response model. *Applied Psychological Measurement*, 17(2), 131-141. <https://doi.org/10.1177/014662169301700203>

- Camilli, G., Wang, M. M., & Fesq, J. (1995). The effects of dimensionality on equating the law school admission test. *Journal of Educational Measurement*, 32(1), 79-96. <https://doi.org/10.1111/j.1745-3984.1995.tb00457.x>
- Chen, Y. F., & Jiao, H. (2013). Does model misspecification lead to spurious latent classes? An evaluation of model comparison indices. *New Developments in Quantitative Psychology* (pp. 345-355). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-9348-8_22
- Desa, Z. N. (2012). *Bi-factor Multidimensional Item Response Theory Modeling for Subscores Estimation, Reliability, and Classification* (Doctoral dissertation, University of Kansas).
- Dodeen, H. (2004). The relationship between item parameters and item fit. *Journal of Educational Measurement*, 41(3), 261-270. <https://doi.org/10.1111/j.1745-3984.2004.tb01165.x>
- Fay, D. (2012). *Sample size and test length minima for DIMTEST with conditional covariance-based subtest selection*. Arizona State University.
- Hutchinson, S. R., & Olmos, A. (1998). Behavior of descriptive fit indexes in confirmatory factor analysis using ordered categorical data. *Structural Equation Modeling*, 5, 344-364. <https://doi.org/10.1080/10705519809540111>
- Jang, E. E., & Roussos, L. (2007). An Investigation into the Dimensionality of TOEFL Using Conditional Covariance—Based Nonparametric Approach. *Journal of Educational Measurement*, 44(1), 1-21. <https://doi.org/10.1111/j.1745-3984.2007.00024.x>
- Jiang, S., Wang, C., & Weiss, D. J. (2016). Sample size requirements for estimation of item parameters in the multidimensional graded response model. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00109>
- Knol, D. L., & Berger, M. P. (1991). Empirical comparison between factor analysis and multidimensional item response models. *Multivariate Behavioral Research*, 26(3), 457-477. https://doi.org/10.1207/s15327906mbr2603_5
- Leighton, J. P., Gokiert, R. J., & Cui, Y. (2007). Using exploratory and confirmatory methods to identify the cognitive dimensions in a large-scale science assessment. *International Journal of Testing*, 7(2), 141-189. <https://doi.org/10.1080/15305050701193520>
- Luecht, R. M. (2004). *Mirtgen* (Version 2). Greensboro, NC.
- McBride, J. R. (1977). Some properties of a Bayesian adaptive ability testing strategy. *Applied Psychological Measurement*, 1(1), 121-140. <https://doi.org/10.1177/014662167700100119>
- Meijer, R. R., Sijtsma, K., & Smid, N. G. (1990). Theoretical and empirical comparison of the Mokken and the Rasch approach to IRT. *Applied Psychological Measurement*, 14(3), 283-298. <https://doi.org/10.1177/014662169001400306>
- Mokken, R. J. (1971). *A Theory and Procedure of Scale Analysis: With Applications in Political Research* (Vol. 1). Walter de Gruyter. <https://doi.org/10.1515/9783110813203>
- Nandakumar, R., & Stout, W. (1993). Refinements of Stout's procedure for assessing latent trait unidimensionality. *Journal of Educational Statistics*, 18(1), 41-68.
- Olmuş, H., Nazman, E., & Erbaş, S. (2017). An Evaluation of the Two Parameter (2-PL) IRT Models Through a Simulation Study. *Gazi University Journal of Science*, 30(1), 235-249.
- Roussos, L. A., & Ozbek, O. Y. (2006). Formulation of the DETECT population parameter and evaluation of DETECT estimator bias. *Journal of Educational Measurement*, 43(3), 215-243. <https://doi.org/10.1111/j.1745-3984.2006.00014.x>
- Roussos, L. A., & Stout, W. F. (1996). Simulation studies of the effects of small sample size and studied item parameters on SIBTEST and Mantel-Haenszel type I error Performance. *Journal of Educational Measurement*, 33(2), 215-230. <https://doi.org/10.1111/j.1745-3984.1996.tb00490.x>
- Stout, W., Habing, B., Douglas, J., Kim, H. R., Roussos, L., & Zhang, J. (1996). Conditional covariance-based nonparametric multidimensionality assessment. *Applied Psychological Measurement*, 20(4), 331-354. <https://doi.org/10.1177/014662169602000403>

- Van Abswoude, A. A., van der Ark, L. A., & Sijtsma, K. (2004). A comparative study of test data dimensionality assessment procedures under nonparametric IRT models. *Applied Psychological Measurement*, 28(1), 3-24. <https://doi.org/10.1177/0146621603259277>
- Van Abswoude, A. A., Vermunt, J. K., Hemker, B. T., & van der Ark, L. A. (2004). Mokken scale analysis using hierarchical clustering procedures. *Applied Psychological Measurement*, 28(5), 332-354. <https://doi.org/10.1177/0146621604265510>
- Van der Ark, L. A., Croon, M. A., & Sijtsma, K. (2008). Mokken scale analysis for dichotomous items using marginal models. *Psychometrika*, 73(2), 183. <https://doi.org/10.1007/s11336-007-9034-z>
- Wainer, H., & Wang, X. (2001). Using a new statistical model for testlets to score TOEFL. *ETS Research Report Series*, 2001(1). <https://doi.org/10.1002/j.2333-8504.2001.tb01851.x>
- William Stout Institute for Measurement. (2006). *Nonparametric dimensionality assessment package DIMPACK (Version 1.0) [Computer software]*. St. Paul, MN: Assessment Systems Corporation.
- Yavuz, G., & Doğan, N. (2015). Test Boyutluluğunun Analizinde Kullanılan Programların İncelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(2). <https://doi.org/10.21031/epod.18173>
- Zhang, J., & Stout, W. (1999). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika*, 64(2), 213-249. <https://doi.org/10.1007/BF02294536>
- Zijlstra, W. P., van der Ark, L. A., & Sijtsma, K. (2011). Outliers in questionnaire data: Can they be detected and should they be removed?. *Journal of Educational and Behavioral Statistics*, 36(2), 186-212. <https://doi.org/10.3102/1076998610366263>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).