Ali Hasmy (2014). Compare Unidimensional & Multidimensional Rasch Modelfor Testwith Multidimensional Construct and Items Local Dependence. *Journal of Education and Learning.* Vol. 8 (3) pp. 187-194.

# Compare Unidimensional and Multidimensional Rasch Model for Test with Multidimensional Construct and Items Local Dependence

Ali Hasmy*
Pontianak Islamic State College

## Abstract

Test in education, generally, constructed based on multi contents. If every contents viewed as a dimension, hence the test also have multi-dimension too, so that the analysis must be done with the multidimensional model. In reality at some selected tests, we can see that dependence exist between items in different contents. Dependence as referred, make assumption about multidimensional test is not valid anymore so that gives us the opportunity to apply unidimensional models in its analysis, including by using Rasch Model. This matter motivated me to do write this paper with aim to: get Unidimensional Rasch Model based on logistics function, get the item parameter estimation and testee parameter estimation together using the Joint Maximum Likelihood Method, and prove that Unidimensional Rasch Model is the better choice than Multidimensional Rasch Model in condition like this.

**Key Words**: *Educational Testing, Rasch Model, Items Local Dependence*

* Ali Hasmy, Islamic Studies Department, Pontianak Islamic State College
Soeprapto Street, No19, Pontianak 78121
E-mail: *ali_hasmy@yahoo.com*

## Introduction

Test in the field of education is oftentimes constructed based on multi contents (Wilson, 2003) and if each content assumed represent one dimension, hence test will oftentimes also have constructed with multi-dimensions and ought to be analyzed with multidimensional model (Xie, 2001). Though test constructed base on single content, by using statistical analysis the dimension is proven sometimes not single caused by dependence between contents. For example in English test which constructed base on grammar content, if after tryout and analyzed statistically, first item ($b_1$) dependence with $b_2$ and $b_3$ so that unionize the items, also with 4th item ($b_4$) until item J th ($b_J$), thereby statistical construct, we can say that the test has a multidimensional construct caused by dimension addition (Linden & Hambleton, 1997). Based on this argument, many educational measurers concluded that a test is a multidimensional phenomenon like Fusco & Dickes (2006), so that applied model to analyze it should be multidimensional too.

Opinion that a test is multidimensional phenomenon can be accepted, but in reality at some selected test, we can see that between content to another actually often not independent. For example in English test which can consist of Grammar, Vocabulary, and Reading Comprehension, we can see that Reading Comprehension related to Grammar and so that Vocabulary. Statistical analysis to the test result often show that there are dependence between items between contents, and make one new group of items. Thereby, that happened is not addition of dimension but reduction of dimension, so that we can be applying uni dimensional model to analyze it. This matter has been proven by Jiao & Kamata in their research entitling, "*Model Comparisons in the Presence of Local Item Dependence*" in the year 2003, using One Parameter Hierarchical Generalized Linear Model (Jiao & Kamata, 2003).

Differ from work of Jiao & Kamata, this paper using Rasch Model with aim to get Unidimensional Rasch Model for analyze test with multidimensional content base construct, where items local dependence exist, to get it's testee ability and item difficulty index parameter, and proof that Unidimensional Rasch Model is the better choice than Multidimensional Rasch Model in condition like this.

## Rasch Measurement Theory

A test when viewed as a multidimensional phenomenon, say, start the 1st dimension until Kth, hence we can arrange a table where testee *i* display in rows, dimension *k* display in columns and each dimension has several items *j*. Each observation which display in cells have a value below:

$$X_{ijk} = \begin{cases} 1, & \text{If the answer is CORRECT} \\ 0, & \text{If the answer is WRONG} \end{cases} \quad , \iota =, i = 1, 2, \ldots I;\ j = 1_k, 2_k, \ldots, J_k; \text{ and } k = 1, 2, \ldots, K.$$

Data structure as mean above can make into single dimension with replace dimension *k* but remain using the same item columns. Each observation which display in every cells have a value,

$$X_{ij} = \begin{cases} 1, & \text{If the answer is CORRECT} \\ 0, & \text{If the answer is WRONG} \end{cases} \qquad i = 1, 2, \ldots, I, j = 1, 2, \ldots, J \tag{1}$$

### Distribution, Logistic Function, and Model

Observations with value at (1) in each cells, $X_{ij}$ ~ *Bernoulli* (1, *p*),with density function,

$$f(x) = p^x q^{1-x} I_{(0,1)}(x), \tag{2}$$

with random variable space *x*, $\Omega x$ = {0,1}, parameter space $\Omega p = \{p \mid 0 \le p \le 1\}$, and $q = 1 - p$ (Mood, Graybill, & Boes, 1987).So that, observations from one testee at entire items, $\sum_{j=1}^{J} X_{ij} \sim B(J, p)$, and observations at an item for entire testee, $\sum_{i=1}^{I} X_{ij} \sim B(I, p)$. Generally, if *n* we use to symbolize *i* = 1, 2, ..., *I*, or *j* = 1, 2, ..., *J*, then we can have $y \sim B(n, p)$ with density function,

$$f(y)=\binom{n}{y}p^y q^{n-y} I_{(0,1,\ldots,n)}(y),$$

with random variable space *y*, $\Omega y$ = {0,1, …, *n*}, parameter space $\Omega p=\{p\,|0\le p\le 1\}$, and *q* = 1 – *p* (Johnson & Albert, 2000; Mood, et al., 1987).

Value $X_{ij}$ as referred to (1) fulfilling the logistics function: $f(z_{ij})=1/(1+\exp(-z_{ij}))$ If $z_{ij}=\alpha+\beta X_{ij}$, (Johnson & Albert, 2000; McCullah & Nelder,1997) then substitute to logistics function, hence will be got the logistics model (Kleinbaum, 1994).

## Multidimensional Rasch Model

### Model

If at test owning construct based on contents then it has multi-dimensions. When P($X_{ij}$) = 1 defined as opportunity of testee *i* to answer the item *j* is correctly determined by the combination from ability of testee *i*, ($\theta_i$), with difficulty index of item *j*, $(\delta_j)$, (Mc Cullah & Nelder, 1997), hence the model (Hoijtink & Vollema, 2001) can be written down as follow:

$$P_j(\theta_i)=\frac{\exp\left(\sum_{k=1}^{K} b_{jk}(\theta_{ik}-\delta_j)\right)}{1+\exp\left(\sum_{k=1}^{K} b_{jk}(\theta_{ik}-\delta_j)\right)} \quad (3)$$

where $P_j(\theta_i)$ is the opportunity of testee *i*answer correctly at item *j*, $b_{jk}$ = 1 is a constant for item discrimination, $\theta_{ik}$ is the ability of testee *i* at $k^{th}$ dimension, $\delta_j$ is the difficulty index of item *j* (Johnson & Albert, 2000; Verguts & Boeck, 2000).

### Basic Assumptions

Two basic assumptions for applying Rasch Measurement Theory are local independence and dimensionality (Hambleton, Swaminathan, & Rogers, 1991, and Xie, 2001). Dimensionality of a test related to subdividing of items (Gierl, Leighton, & Tan, 2006) classified as Content-Based Dimensionality and Statistical Dimensionality. Statistically, testing dimensionality is equal with testing of local independence. This matter because if there be dependence between items between contents, hence the test which initially constructed base on the contents is improvable multi-dimensions statistically. In a condition like this, the dimension will be reduce and make opportunity to apply Unidimensional Rasch Model. While item local independence interpreted that any item couples at one particular test is independence each other to the selected latent trait $\theta$ (Mari& Kotz, 2001 and Verguts & Boeck, 2001), equally the answer of testee *i* at one item is dependence with another (Fusco & Dickes, 2006 and Xie, 2001). Relationship between items only explained passing the conditional relationship with $\theta$.

## Unidimensional Rasch Model as Alternative when Multidimensional Test has Items Local Dependence

### Model

Related with Kleinbaum's opinion about the odds value, hence it can be interpreted in two meaning. **First**, if the logistics model with one predictor used to express opportunity of testee answer wrongly, hence by writing down opportunity of testee to answer the correctly as,

$$1-P(z_{ij})=P_j(\theta_i)=1-\frac{1}{1+\exp(-z_{ij})}=\frac{1+\exp(-z_{ij})}{1+\exp(-z_{ij})}-\frac{1}{1+\exp(-z_{ij})}=\frac{\exp(-z_{ij})}{1+\exp(-z_{ij})}$$

and if $z_{ij} = \alpha + \beta X_{ij}$ then $P_j(\theta_i) = \dfrac{\exp(-(\alpha + \beta X_{ij}))}{1 + \exp(-(\alpha + \beta X_{ij}))}$ .If second part of $\beta X_{ij}$ replaced with $\theta_i$ , symbolize ability testee *i* to answer item *j* correctly, in multiplicative model hence

$$P_j(\theta_i) = \frac{\exp(-(\alpha\theta_i))}{1 + \exp(-(\alpha\theta_i))}. \tag{4}$$

With adding $\delta_j$ , symbolize difficulty index of item *j*, where $\delta_j$ having the character to reduction $\theta_i$ , thereby writing wholly is

$$P_j(\theta_i) = \frac{\exp(-(\alpha\theta_i - \alpha\delta_j))}{1 + \exp(-(\alpha\theta_i - \alpha\delta_j))}. \tag{5}$$

If $-\alpha\delta_j$ replaced with $\beta$ ,we got $z_{ij} = \alpha\theta_i + \beta$ .

Equation (5) also can be written as

$$P_j(\theta_i) = \frac{\exp(-\alpha(\theta_i - \delta_j))}{1 + \exp(-\alpha(\theta_i - \delta_j))}, \tag{6}$$

if $-\alpha$ (intercept) is the item discrimination index (Cox & Snell, 1996) specified equal to 1 as difference of value between testee capability to answer correctly (valuable 1) and the unable to answer correctly (valuable 0), hence got the matching one which used by Embretson & Reise (2000), Hambleton & Swaminathan (1985), Hambleton, Swaminathan & Rogers (1991), Johnson & Albert (2000), and von Davier & Carstensen (2007)

$$P_j\left(\theta_i\right) = \frac{\exp\left(\theta_i - \delta_j\right)}{1 + \exp\left(\theta_i - \delta_j\right)}, \tag{7}$$

which equal with (3) when altered directly by changing the index of *k*= 1.

**Second**, if the logistics model in this case used to express opportunity of testee to answer correctly, hence by doing resettlement the logistics model with use the $z_{ij}$= $\alpha + \beta X_{ij}$ , where the part which load the random variable random excluded from equation, hence $P(z_{ij}) = P_j(\theta_i) = 1/(1 + \exp(-(\alpha)))$ .With process like before, we got the Unidimensional Rasch Model pursuant to Baker's opinion (2001), Fox (2010), Hulin, Drasgow, & Parsons (1983), and Wainer, Bradlow, & Wang (2007)which basically the same as (7)

$$P_j\left(\theta_i\right) = \frac{1}{1 + \exp\left(\theta_i - \delta_j\right)}. \tag{8}$$

## Parameters Estimation

Steps to appraise the parameter ability of testee and item difficulty index with Joint Maximum Likelihood Method is writing down again (1) base on (2) in the form as follow,

$$\left[P_j\left(\theta_i\right)\right]^{X_{ij}} \left[Q_j\left(\theta_i\right)\right]^{1-X_{ij}} = \begin{cases} P_j\left(\theta_i\right) \; if \quad X_{ij} = 1 \\ Q_j\left(\theta_i\right) \; if \quad X_{ij} = 0 \end{cases} \tag{9}$$

where $P_j(\theta_i)$ is opportunity of correct answer for the item *j* with selected testee ability, $\theta_i$, while $Q_j(\theta_i) = 1 - P_j(\theta_i)$ is opportunity of wrong answer, and $X_{ij}$ = 1 which indicate the correct answer of testee *i* for item *j* and $X_{ij}$ = 0 which indicate the wrong answer.

Using (9) we made the likelihood function to individually response that is $L(\theta_1, \theta_2, \ldots, \theta_j \left| X_{ij} \right.) = \prod_{i=1}^{I} \prod_{j=1}^{J} \left[ P_j\left(\theta_j\right) \right]^{X_{ij}} \left[ Q_j\left(\theta_i\right) \right]^{1-X_{ij}}$ and then the logarithm is $l = \sum_{i=1}^{I} \sum_{j=1}^{J} \left[ X_{ij} In P_j\left(\theta_i\right) + \left(1 - X_{ij}\right) In Q_j\left(\theta_i\right) \right]$ .If we derive to $\theta_i$ and equal the result with 0, we got $\frac{\partial l}{\partial \theta_i} = \sum_{j=1}^{J} \left\{ \left[ X_{ij} - P_j\left(\hat{\theta}_i\right) \right] \frac{P_j\left(\hat{\theta}_i\right)'}{P_j\left(\hat{\theta}_i\right) Q_j\left(\hat{\theta}_i\right)} \right\} = 0$ .

To estimating testee ability parameter, because $P_j(\theta_i)Q_j(\theta_i) > 0$, we can simplify it to $\sum_{j=1}^{J} X_{ij} = \sum_{j=1}^{J} P_j(\theta_i)$ and then,

$$t_i = \sum_{j=1}^{J} P_j\left(\theta_i\right), \; i = 1, \ldots, I \qquad (10)$$

where $t_i = \sum_{j=1}^{J} X_{ij}$ is the total score for testee *i*. To estimating item difficulty index parameter,

$$s_j = \sum_{i=1}^{I} P_j\left(\theta_i\right), j = 1, \ldots, J \qquad (11)$$

where $s_j = \sum_{i=1}^{I} X_{ij}$ is the total wrong answer for item *j*.

## Compare the Models Using Real Data

In this research, real data from Pontianak Islamic State College English Test with multi dimensional construct based on contents like grammar (10 items), vocabulary (3 items) and reading comprehension (7 items) used. This test used to select 114 candidates to enter this college.

Three step shave done to analyze the data: **first**, found dependence between items there with using Cochran's Test, and Breslow-Day Test for the assumption of odds homogeneity. The result proved that local items dependence exist 10 from 121 item pairs between contents or 8,3%.

**Second**, testee and item parameters found using Prox. Method (Miller, 2004). The result showed on Table 1.

Table 1. Item difficulty

| Items | Difficulty ($\delta$) | | Category | |
|---|---|---|---|---|
| | Unidimensional | Multidimensional | Unidimensional | Multidimensional |
| 1 | 0,37650 | 0,14010 | Moderate | Moderate |
| 2 | 0,00625 | – 0,21158 | Moderate | Moderate |
| 3 | 0,07691 | – 0,14359 | Moderate | Moderate |
| 4 | 0,45682 | 0,21481 | Moderate | Moderate |
| 5 | 0,71592 | 0,45155 | **High** | **Moderate** |
| 6 | 1,65988 | 1,24555 | High | High |
| 7 | 0,45682 | 0,21481 | Moderate | Moderate |
| 8 | – 0,93731 | – 1,15577 | Low | Low |
| 9 | – 0,19847 | – 0,41080 | Moderate | Moderate |
| 10 | – 0,13128 | – 0,34506 | Moderate | Moderate |
| 11 | – 0,77803 | – 1,35151 | Low | Low |
| 12 | 0,85851 | 1,09295 | High | High |
| 13 | 0,26049 | 0,25856 | Moderate | Moderate |
| 14 | 0,53990 | 1,02566 | **Moderate** | **High** |
| 15 | 0,11278 | 0,56864 | Moderate | Moderate |
| 16 | – 0,84158 | – 0,46591 | **Low** | **Moderate** |
| 17 | 1,01261 | 1,52904 | High | High |
| 18 | – 1,76642 | – 1,50572 | **Very Low** | **Low** |
| 19 | – 1,19743 | – 0,85970 | Low | Low |
| 20 | – 0,68286 | – 0,29201 | **Low** | **Moderate** |

Table 1 has shown that 10 (50%) items on the test have moderate difficulty index, 4 (20%) items have high difficulty index (hard items), 5 (25%) items have low difficulty index (easy items) and, 1 (5%) item has very low difficulty index (very easy item) in unidimensional analysis. While in multidimensional analysis12 (60%) items have moderate difficult, 4 (20%) items are difficult, and 4 items (20%) is easy items. The composition is make sense.

**Third**, total errors $\sum P_j(\theta_i)$ from $\sum X_{ij}$ per testee using Unidimensional Rasch Modeland Multidimensional Rasch Modelhas found and compared. $P_j(\theta_i)$ from unidimensional and multidimensional analysis can be seen on Table 2 and Table 3.

Table 2. $P_j(\theta_i)$ from Unidimensional Analysis

| Observed Total Scores | Testee Ability ($\theta$) | Category |
|---|---|---|
| 1 | – 3,28020 | Very Low |
| 2 | – 2,44778 | Very Low |
| 3 | – 1,93240 | Very Low |
| 4 | – 1,54438 | Low |
| 5 | – 1,22389 | Low |
| 6 | – 0,94392 | Low |
| 7 | – 0,68963 | Low |
| 8 | – 0,45170 | Moderate |
| 9 | – 0,22355 | Moderate |
| 10 | 0 | Moderate |
| 11 | 0,22355 | Moderate |
| 12 | 0,45170 | Moderate |
| 13 | 0,68963 | High |
| 14 | 0,94392 | High |
| 16 | 1,54438 | High |
| 18 | 2,44778 | Very High |

Table 3. $P_j(\theta_i)$ from Multidimensional Analysis

| Dimension | Total Scores | Testee Ability ( $\theta$ ) | Category |
|---|---|---|---|
| 1 | 1 | – 2,34263 | Very Low |
| 1 | 2 | – 1,47803 | Low |
| 1 | 3 | – 0,90337 | Low |
| 1 | 4 | – 0,43230 | Moderate |
| 1 | 5 | 0 | Moderate |
| 1 | 6 | 0,43230 | Moderate |
| 1 | 7 | 0,90337 | High |
| 1 | 8 | 1,47803 | High |
| 1 | 9 | 2,34263 | Very High |
| 2 | 1 | – 0,85857 | Low |
| 2 | 2 | 0,85857 | High |
| 3 | 1 | – 2,12380 | Very Low |
| 3 | 2 | – 1,08610 | Low |
| 3 | 3 | – 0,34099 | Moderate |
| 3 | 4 | 0,34099 | Moderate |
| 3 | 5 | 1,08610 | High |
| 3 | 6 | 2,12380 | Very High |

Using Wilcox on Sign Rank Test to compare the residuals from unidimensional and multidimensional analysis based on Table 2 and Table 3, proved that Unidimensional Rasch Model have a significant lower residual than Multidimensional Rasch Model with p-value < 0,05. So, applying Unidimensional Rasch Model in the case like this is better than Multidimensional Rasch Model.

## Conclusion

Three conclusions can be made from this research: **first**, Unidimensional Rasch Model can be represented with two basically the same equation (7 & 8), **second**, we can get testee parameters and item parameters with solving these two equations $t_i = \sum_{j=1}^{J} P_j\left(\hat{\theta}_i\right)$, $i$ = 1, ..., $I$ and $s_j = \sum_{i=1}^{I} P_j\left(\hat{\theta}_i\right)$, $j$ = 1, ..., $J$, and **third**, Unidimensional Rasch Model proved better than Multidimensional Rasch Model in this case.

## References


Baker, F. B. (2001).*Basic item response theory*. USA: ERIC Clearinghouse on Assessment and Evaluation.

Cox, D. R., & Snell, E. J. (1996). *Analysis of binary data*. Second Edition. London, UK: Chapman & Hall.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologist*. Mahwah, NJ: Lawrence Erlbaum, Inc.

Fox, J. P. (2010). *Bayesian item response modeling*: *Theory and applications*. New York, NY: Springer.

Fusco, A. & Dickes, P. (2006). Rasch model and multidimensional poverty measurement. *An integrated research infrastructure in the Socio-economic Science* (*IRISS*) *Working Paper Series*. *2*: 1-18.

Gierl, M. J., Leighton, J. P., & Tan, X. (2006). Evaluating DETECT classification accuracy and consistency when data display complex structure. *Paper Presented at Annual Meeting of the National Council on Measurement in Education* (*NCME*). San Francisco.

Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory*: *Principles and applications*. Boston, MA: Kluwer Nijhoff Publishing.

Hambleton R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications, Inc.

Hoijtink, H., & Vollema, M. (2001). Contemporary extensions of the Rasch model. *Paper*. Department of Methodology and Statistics, University of Utrecht, Netherlands.

Jiao, H., & Kamata, A. (2003). Model comparisons in the presence of local item dependence. *Paper Presented atAnnual Meeting of the American Educational Research Association*, Chicago.

Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory*: *Application to psychological measurement*. Homewood, IL: The Dorsey Professional Series.

Johnson, V. E., & Albert, J. H. (2000).*Ordinal Data Modeling*. Second Corrected Printing. New York, NY: Springer-Verlag.

Kleinbaum, D. G. (1994).*Logistic regression*: *A self-learning text*. New York, NY: Springer-Verlag.

Mari, D. D., & Kotz, S. (2001). *Correlation and dependence*. London, UK: Imperial College Press.

McCullah, P., & Nelder, J. A. (1997). G*eneralized linear models*. Second Edition Reprinted. London, UK: Chapman & Hall.

Miller, W., G., (2004), *Statistics and measurement using Open stat package*, http://Statpages.org/ javasta2.html, downloaded at January 13th2007.

Mood, A. M., Graybill, F. A., & Boes, D. C. (1987). *Introduction to the Theory of Statistics*. Third Edition. Auckland, NZ: McGraw-Hill Book Company.

van der Linden, W. J., & Hambleton, R. K.. (1997). *Handbook of modern item response theory*. New York, NY: Springer.

Verguts, T. & Boeck, P. D. (2000). A Rasch model for detecting learning while solving an intelligencetest.*Applied Psychological Measurement*, *24*(2): 151-162.

--------------. (2001). Some Mantel-Haenszel test of Rasch model assumptions. *British Journal of Mathematical and Statistical Psychology*, *54*: 21-37.

von Davier, M., & Carstensen, C. H. (2007). *Multivariate and mixture distribution Rasch models*: *Extension and applications*. New York, NY: Springer.

Wainer, H., Bradlow, E. T., & Wang, X. (2007). *Testlet response theory and its applications*. Cambridge, UK: Cambridge University Press.

Wilson, M. (2003). On choosing a model for measuring. *Methods of Psychological Research Online*. *8*(3), 1-22.

Xie, Y. (2001). Dimensionality, dependence, or both. *Paper*. University of California.