Utilizing ordinal alpha reliability on modification of Mark H. Davis’s empathy rate instrument (case study in Department of Mathematics FMIPA UI)

F A Doko, R Setiadi and S M Soemartojo
Department of Mathematics, Faculty Mathematics and Natural Sciences (FMIPA), Universitas Indonesia, Depok 16424, Indonesia
Corresponding author’s email: rianti@sci.ui.ac.id

Abstract. Most psychology research requires the existence of an instrument to obtain the required information. An instrument using Likert scale needs to know its reliability and validity. However, the use of an instrument is not necessarily suitable for particular population. This can be seen from the fulfillment of reliability, construct validity, item validity, and content validity of the instrument. Unreliable and invalid instruments due to cultural differences in the population. Therefore, a modification of the instrument is needed such that better information can be obtained to the relevant population. Because in this study using ordinal scale data, then the reliability will be measured by the coefficient ordinal alpha with polychoric correlation. In this study, modifications will be made to Mark H. Davis' rate empathy instrument consisting of 45 indicators into an instrument consisting of only 15 indicators. Modified measuring tools proved reliable, have valid content, valid constructs, and valid items. The methods to be used are factor analysis, Spearman correlation, polychoric correlation, and ordinal alpha reliability.

Keywords: Alpha ordinal, factor analysis, likert scale, polychoric correlation

1. Introduction
Nowadays research has become common thing in many fields of knowledge for both social and science. One of the fundamental things needed in doing psychological research is the questionnaire. The questionnaire is one way to collect the required data from the respondents.

There are several ways of measuring variables in research such as using a physical measuring instrument or asking a respondent with a question. In social science, often find latent variables where the variables cannot be measured directly. The latent variable is measured using several indicators which are a reflection of these latent variables and can be measured directly. One way to measure latent variables with multiple indicators is using Likert scale.

In using Likert scale there are two things that must be checked, they are the reliability and validity. Reliability is to measure how far the indicators have reliable, consistent, stability to measure latent variables. The results of measurement are reliable if in several times the measurement of the same subject group will give relatively similar results [1]. There are several ways to estimate the reliability of the measuring instrument such as test retest, split half, Cronbach’s alpha, theta, omega and alpha ordinal. This research used ordinal data, therefore the reliability used in this research is ordinal alpha reliability.
The ordinal alpha is a method for estimating the reliability coefficients specific to ordinal data so it is appropriate to apply this method in estimating the reliability of the empathy level indicators [2]. Validity is to measure how far the accuracy of a measuring instrument in performing its measuring function. There are several types of validity such as construct validity, item validity, and content validity. For certain data, reliability, construct validity, item validity, and content validity are often not being fulfilled. This is due to differences in culture of the study respondents. One way to overcome this case is to make modifications on measuring instrument by adjusting the indicators that exist based on data from respondents. The modification must fulfill reliability, construct validity, item validity, and content validity in accordance with the original measuring instrument.

This research will use the empathy rate instrument made by Mark H. Davis because the measuring instrument has proved to be reliable and valid for the student population of the Faculty of Psychology in Texas. However, the empathy rate instrument made by Mark H. Davis needs to be modified to apply on students of Mathematics FMIPA UI because the construct validity of the measuring instrument is not fulfilled. This is caused by the difference of population used in the research causing the difference of culture which can influence empathy rate of respondent in the population.

At this time there has been no research that discusses the modification of empathy rate measuring instrument made by Mark H. Davis for Mathematics students in Universitas Indonesia. Therefore, to answer this need, this research needs to be done.

Problems:
1) How to choose indicators to be used for measuring empathy rate as a modified measuring instrument
2) How to check the reliability of the modified empathy measuring instrument by using the ordinal alpha coefficient
3) How to check the construct validity, item validity, and the content validity of the modified empathy rate measuring instrument

Objectives:
1) Selecting indicators to be used on the modified empathy rate measuring instrument
2) Examining the reliability of the modified empathy rate measuring instrument by using the ordinal alpha coefficient
3) Checking the construct validity, item validity, and content validity of the modified empathy rate measuring instrument

Scope of problem:
This research has just been analyzed using data were collected by distributing questionnaires to 170 mathematics students of FMIPA UI who have been selected as sample.

2. Experimental Method
Population in this research are students of Department of Mathematics FMIPA, Universitas Indonesia in generation of 2012–2016 with active status of academic year 2017/2018. The number of students is 513 people. The sample was taken using Simple Random Sampling (SRS) as many as 170 students were 76 men and 94 women. Methods used in this research are Pearson correlation, Spearman correlation, polychoric correlation, factor analysis, validity, and ordinal alpha reliability.

2.1. Pearson correlation
The correlation used mostly between the two variables is the Pearson correlation coefficient [3]. Pearson correlation is calculated using continuous data and calculated based on the value of existing data [4]. The following is a formula for calculating the Pearson correlation coefficient [4]:

\[ r = \frac{\sum_{i=1}^{n} X_i Y_i - n \bar{X} \bar{Y}}{\left(\sum_{i=1}^{n} X_i^2 - n \bar{X}^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} Y_i^2 - n \bar{Y}^2\right)^{\frac{1}{2}}} \]

where \( r \) denotes the Pearson correlation coefficient, \( n \) denotes the number of data pairs \((X, Y)\).
To know that there is a significant correlation between \(X\) dan \(Y\), it is necessary to test the hypothesis for Pearson correlation with \(H_0\) dan \(H_1\) as follows:

\(H_0\) : No correlation between \(X\) and \(Y\)
\(H_1\) : There is correlation between \(X\) and \(Y\)

The test statistic used for this test is:

\[
T = r \sqrt{\frac{n-2}{1-r^2}}
\]  

(2)

If \(X\) dan \(Y\) are normally distributed then it can be shown that \(T\) is distributed \(t\) with degrees of freedom \((n-2)\). \(H_0\) will be rejected if \(T\) is greater than \(t(\alpha/2, n-2)\) or if \(T\) is smaller than \(-t(\alpha/2, n-2)\) [3].

2.2. Spearman correlation

Besides using Pearson correlation, this research also uses Spearman correlation that is linear correlation based on rank of two paired variable. Spearman correlation coefficient is part of nonparametric statistics so that the assumption of distribution can be ignored, but the data used at least ordinal scale [5].

The formula for Spearman's correlation with ties is as follows [4]:

\[
r^* = \frac{\sum_{i=1}^{n} R(X_i)R(Y_i) - n \left(\frac{n+1}{2}\right)^2}{\left(\sum_{i=1}^{n} R(X_i)^2 - n \left(\frac{n+1}{2}\right)^2\right)^{1/2}\left(\sum_{i=1}^{n} R(Y_i)^2 - n \left(\frac{n+1}{2}\right)^2\right)^{1/2}}
\]  

(3)

where \(r^*\) denotes the Spearman correlation coefficient, \(R(X_i)\) denotes the rank of variable \(X_i\), \(R(Y_i)\) denotes the rank of variable \(Y_i\), \(n\) denotes the number of data pairs \((X,Y)\). As for the Spearman correlation coefficient in the absence of ties are as follows [4]:

\[
r^* = 1 - \frac{6 \sum_{i=1}^{n} [R(X_i) - R(Y_i)]^2}{n(n^2-1)} = 1 - \frac{6T}{n(n^2-1)}
\]  

(4)

To know that there is a significant correlation between \(X\) dan \(Y\), it is necessary to test hypothesis for Spearman correlation with \(H_0\) dan \(H_1\) are as follows:

\(H_0\) : No correlation between \(X\) and \(Y\)
\(H_1\) : There is correlation between \(X\) and \(Y\)

The test statistic used for this test is:

\[
T = \sum_{i=1}^{n} [R(X_i) - R(Y_i)]^2.
\]  

(5)

\(H_0\) is rejected at the level of significance \(\alpha\) if the value of \(T\) is less than \(\frac{\alpha}{2}\) quantile in the table of Hotelling-Pabst-Test-Statistics [4].

2.3. Polychoric correlation

This research also uses polychoric correlation which is one method to measure the relationship between two random variables that are ordinal. Because Pearson correlation can only be used to measure the linear relationship between two continuous random variables, so it cannot be used to look for correlations of non-continuous ordinal variables.
2.4. Factor analysis
Modification of measuring instruments in this study will use factor analysis that is a statistical method used to reduce a number of initial variables that are quite a number into several new variables called factors, where the number is less to facilitate interpretation. In addition, the factors formed can adequately explain the information contained in the initial variables.

In factor analysis, large variables are grouped into a number of factors that have similar properties and characteristics in each factor. Grouping is done on the basis of the correlation of the initial variables and subsequently assigns the highly correlated variables in one factor and other variables with a relatively lower correlation placed on the other factors.

Furthermore, the discussion in this paper will be limited to the EFA approach, where researchers have not had the knowledge or theories that make up the structure of the factors that will be formed. EFA process, beginning with the relationship between variables which will then look for the underlying factors [2]. The procedure in performing this analysis is to form the correlation matrix of the initial variables, factor extraction, factor rotation and factor interpretation.

2.5. Validity
After the process of factor analysis, then will be checked the validity. A test is said to have high validity if the measuring instrument performs the measuring function appropriately or provides a measuring result that is suitable for the purpose of the measurement.

There are three kinds of validity, they are construct validity, item validity and content validity. When measuring instrument meets all three types of validity, then it can be said that the measuring instrument has been able to express exactly the actual characteristics or circumstances of the measured object. The following will be explained about the types of validity.

First is the construct validity that is the suitability of grouping indicators with the construction or dimension of latent variables. The construct validity can be checked using factor analysis.

Second is the item validity that states that all the indicators are together measuring the same thing. An indicator is said to be a valid item if the correlation of the indicator with the value of the variable it formed is significant. In this case the correlation used is Spearman correlation.

Third is the content validity that is how far the indicators in the measuring instrument can reflect the characteristics of subject. Content validity is determined by agreement of experts.

2.6. Ordinal alpha reliability
Reliability is a measure that indicates that the instrument used in the study has reliability as a measuring instrument. This is measured through the consistency of a measurement result against the same respondent in different times, as long as the aspects measured in the respondent's self are not changed [6].

Due to this research using Likert scale where the data is in ordinal scale, Gadermann et al. [7] introduced the alpha coefficients for ordinal data commonly referred to as ordinal alpha. Ordinal alpha is the appropriate reliability coefficient for ordinal data. The ordinal alpha coefficient uses polychoric correlation in calculating its reliability.

The following is the formula of the coefficient of ordinal alpha reliability [7]:

\[
\alpha = \frac{K\bar{r}^*}{(1 + (K - 1)\bar{r}^*)}
\]

where \( K \) denotes the number of indicators, \( \bar{r}^* \) denotes the average polychoric correlation among indicators.

3. Results and discussion
The instrument to be modified in this study was an empathy rate questionnaire created by Mark H. Davis and consisted of 45 statement indicators. The measuring instrument for this rate of empathy uses a
Likert scale of 1–5. Next will be checked first the construct validity, the item validity and reliability based on the data obtained. If the construct validity of the measuring instrument is not met, it will be modified against the measuring instrument. Modified gauges will be checked again the construct validity, item validity, content validity, and reliability. The following steps are analysis of the data used to answer the research objectives.

3.1. Checking the construct validity of the measuring instrument before modification

Based on data obtained by researchers from 45 indicators of revelation that has been disseminated to 170 students of the Department of Mathematics FMIPA Universitas Indonesia in generation of 2012–2016, conducted Exploratory Factor Analysis (EFA). In the process, indicators 5 and 21 are omitted from measuring instruments because they do not fulfill the Measure of Sampling Adequancy (MSA) requirements. Then, from the remaining 43 indicator variables, conducted Exploratory Factor Analysis (EFA) and obtained the KMO value of 0.741, the significance value based on Bartlett test is 0.000 and based on the value of MSA generated for each indicator has been fulfilled that is greater than 0.5 so it is concluded that the terms of the EFA have been met. Then, from the EFA process obtained the results of the grouping as follows table 1.

From the results above shows that from 43 indicators will form 14 independent factors. But according to Mark H. Davis, the level of empathy is only built by 4 factors. If it done Confirmatory Factor Analysis (CFA) by determining the 4 factors, it will be found that the measuring instrument is not construct valid. Therefore, the researcher will modify the measuring instrument and then check the construct validity, item validity, content validity, and its reliability. Modification of indicators from Mark H. Davis measuring instrument produces 15 indicators. These 15 indicators were obtained by selecting the indicators on the factors formed from the EFA and adjusting to the 4 factors created by Mark H. Davis. Indicators 22, 31 and 45 are not included in the resulted modification instrument because the measured reliability value of the measuring instrument will be less than 0.6. The following results are obtained as follows figure 1.

3.2. Checking the construct validity of the measuring instrument after modification

The modified measuring instrument will be checked for the construct validity using Exploratory Factor Analysis (EFA). The value of KMO is 0.714, the significance value based on the Bartlett test is 0.000 

| Table 1. Factors from EFA process |
| Factor | Indicator |
|--------|----------|
| 1 | 15, 17, 22, 38, 45 |
| 2 | 3, 9, 14, 25, 33 |
| 3 | 8, 16, 26, 31, 36, 39 |
| 4 | 7, 28, 32, 43 |
| 5 | 20, 27, 30, 34 |
| 6 | 18, 19, 44 |
| 7 | 4, 11, 37 |
| 8 | 13, 29 |
| 9 | 2, 40 |
| 10 | 1, 6, 12 |
| 11 | 10, 41 |
| 12 | 24, 35 |
| 13 | 23 |
| 14 | 42 |

Figure 1. Modification chart of measuring instrument
and based on the value of MSA generated for each indicator has fulfilled that is greater than 0.5 so it is concluded that the assumption of EFA has been fulfilled. Then, from the EFA process obtained the results of the grouping as follows (table 2).

From the results on table 2 shows that from 15 indicators, formed 4 independent factors. Adjusting to the journal Mark H. Davis, then factor 1 is perspective taking, factor 2 is personal distress, factor 3 is fantasy and factor 4 is empathic concern. Thus, it can be concluded that the measuring instrument construct of the 15 new indicators is in accordance with the construct of the measuring instrument made by Mark H. Davis.

3.3. Checking content validity of the measuring instrument after modification
Some experts who often use a Likert scale provide a statement that the indicators on the modified instruments have been content valid.

3.4. Checking item validity of the measuring instrument after modification
Check the item validity of the new measuring instrument using Spearman's correlation test. The item validity is obtained by finding the correlation of the indicator with the value of the variable it formed. The results that obtained are all the p-value are 0.000 so all indicators have a significant correlation to the variables that are formed that is perspective taking I, personal distress I, fantasy I, and empathic concern I. So, it can be concluded that the indicators have valid items.

3.5. Checking reliability of the measuring instrument after modification
Check the reliability of the new measuring instrument using the ordinal alpha and the test-retest method. With ordinal alpha method, the results obtained are all latent variables have the value of alpha ordinal coefficients greater than 0.6 so it can be concluded that measuring instrument for the four related factors is reliable (table 3).

3.6. Checking reliability of the measuring instrument after modification with test retest
The modified measuring instrument is then reshared back to the same 170 respondents. After the information obtained, then will be tested the correlation of the results of the first test and the results of the second test. From result of test retest, it concluded that the measuring instrument is reliable.

| Table 2. Factors after modification of measuring instruments. |
|-------------------------------------------------------------|
| Factor | Indicator |
|--------|-----------|
| 1      | 8, 16, 26, 36, 39 |
| 2      | 15, 17, 38 |
| 3      | 18, 19, 44 |
| 4      | 7, 28, 32, 43 |

| Table 3. Ordinal alpha reliability of modified measuring instrument. |
|---------------------------------------------------------------------|
| Factor | Name of factor | Alpha ordinal |
|--------|----------------|---------------|
| 1      | Perspective Taking | 0.75          |
| 2      | Personal Distress | 0.75          |
| 3      | Fantasy         | 0.69          |
| 4      | Empathic Concern | 0.65          |
Table 4. Ordinal alpha reliability of measuring instrument in second test.

| Factor | Name of factor      | Alpha ordinal |
|--------|---------------------|---------------|
| 1      | Perspective Taking  | 0.69          |
| 2      | Personal Distress   | 0.78          |
| 3      | Fantasy             | 0.69          |
| 4      | Empathic Concern    | 0.66          |

3.7. Checking ordinal alpha reliability of modified measuring instrument in second test
It is found that all latent variables of the second test result have an ordinal alpha coefficient value greater than 0.6 so that it can be concluded the measuring instrument for all four related factors is reliable (table 4).

3.8. Checking item validity of modified measuring instrument in second test
It is found that all indicators have a significant correlation to the variables it formed, so it can be concluded that the indicators have valid items. So, it can be concluded that the measuring instrument that has been modified based on the results of the second test can be said to be valid items.

4. Conclusion
The modified measuring instrument consists of 15 indicators based on measuring instruments made by Mark H. Davis. Measuring instruments with 15 indicators can be grouped into 4 factors. Factor 1 is perspective taking consisting of 5 indicators, factor 2 is personal distress consisting of 3 indicators, factor 3 is fantasy consisting of 3 indicators and factor 4 is empathic concern consisting of 4 indicators. Measuring instrument with 15 indicators that have been modified proven valid content, valid constructs, valid items and reliable. Measuring instrument with 15 indicators that have been modified suitable to be used for students of Department of Mathematics FMIPA UI in generation of 2012–2016 with active status of academic year 2017/2018.

References
[1] Azwar S 2005 Tes Prestasi: Fungsi dan Pengembangan Pengukuran Prestasi Belajar (Yogyakarta: Pustaka Pelajar)
[2] Flora D B and Curran P J 2004 Psychol. Methods 9 466-91
[3] Nugroho S, Akbar S and Vusvitasari R 2008 Jurnal Gradien 4 372-81
[4] Conover W J 1998 Practical Nonparametric Statistics 3rd edition (Hoboken: John Wiley & Sons, Inc.)
[5] Sugiarto D S 2006 Metode Statistika untuk Bisnis dan Ekonomi (Jakarta: PT. Gramedia Pustaka Utama)
[6] Zulganef 2006 Pemodelan Persamaan Struktural dan Aplikasinya Menggunakan Amos 5 (Bandung: Pustaka)
[7] Gadermann A M, Guhn M and Zumbo B D 2012 Pract. Asses. Res. Evaluation 17 1-13