# Examining Dimensionality and Validity of the Academic Integrity Survey Instrument

Zhidong Zhang[1] & Luis Garcia[1]

[1] The College of Education and P-16 Integration, The University of Texas-Rio Grande Valley, Texas, USA

Correspondence: Zhidong Zhang, The College of Education and P-16 Integration, The University of Texas-Rio Grande Valley, Texas, USA. E-mail: zhidong.zhang@utrgv.edu

**Abstract**

Dimensional reduction is one of the methods to ensure the quality of a questionnaire. This study examined two methods to reduce the dimension of the questionnaire: multidimensional scaling (MDS) and exploratory factor analysis (EFA). The questionnaire, Awareness of Academic Dishonesty consists of 30 questions. Participants included 110 college students. Multidimensional scaling analysis reduced the multidimensions to essentially two dimensions. The exploratory factor analysis reduced the multidimensions to three dimensions. MDS allowed the researchers to evaluate the questionnaire items by looking at the similarities of these data points. EFA provided an alternative thought about the construct of the questionnaire.

**Keywords:** dimensionality, multidimension scaling (MDS), reliability, validity, questionnaire, and academic integrity

## 1. Introduction

The quality evaluation of the questionnaire as a survey instrument was a concern by many researchers in a variety of fields. Developing and validating a questionnaire in a survey research study are challenging tasks. Evaluators of the survey study instrument assessed this differently based on their different research interests and educational backgrounds (Baker, 2013; Baker, & Keeter, 2022; Biemer, 2010; Bremer, 2013). The quality of the survey instrument, the questionnaire is a comprehensive concept and even to this day, the researchers have not reached an identical view on how to evaluate the survey quality with the questionnaire. However, some aspects of the quality evaluation were highlighted on reliability, validity, instrumental dimensions, and a critical appraisal.

This study is to explore two aspects of the survey instrument quality: validity and instrumental dimensions. The purposes are to allow the beginning investigators to know how to assess the questionnaire quality and rationale of what are effective questions/ items of the questionnaire objectively.

## 2. Perspectives

### 2.1 Reliability of a Questionnaire

Reliability of the questionnaire was described as the degree to which the item scores can be replicated in a different investigation occasion. Reliability is one of the important characteristics of questionnaire quality (Kember & Leung, 2008; Sangoseni, Hellman, & Hill, 2013; Wong, Ong, & Kuek, 2010). "Lack of reliability may arise from divergence between observers or instruments of measurement such as a questionnaire or instability of the attribute being measured which will invariably affect the validity of such questionnaire" (Bolarinwa, 2015, p. 195).

Questionnaire reliability is usually examined in three aspects: Equivalence, stability and internal consistency (homogeneity). Equivalence reliability is to examine the correlation of scores between different versions of the same instrument. It can also be done between instruments that measure the same or similar constructs. Stability is can be measured in an alternate form (Nevill, Lane, Kilgour, Bowes, & Whyte, 2004), which refers to the amount of agreement between two different questionnaires on a research construct that are administered at nearly the same point in time. "It is measured through a parallel form procedure in which one administers alternative forms of the same measure to either the same group or different group of respondents" (Bolarinwa, 2015, p. 196). Internal consistency of the questionnaire refers to how reliable the instrument measures what the researchers believe it will measure. The most common way to measure internal consistency is by using a statistic known as *Cronbach's Alpha*, which calculates the pairwise correlations between items in a survey (Zach, 2022).

*2.2 Validity of a Questionnaire*

Validity is the extent to which an instrument such as a questionnaire measures what is supposed to measure. There are several different ways to examine the validity (Agarwal, 2012). Face and content validity are subjective opinions of the researchers. Face validity is often seen as the weakest form of validity. Therefore, the researchers usually seek other forms to validate the instruments. The other two concepts of validity are often used in questionnaire designs, criterion validity, and current validity. Criterion validity is the extent to which the measures derived from the survey relate to other external criteria. concurrent validity is measured at the same time as the survey, either with questions embedded within the survey, or measures obtained from other sources (Deniz & Alsaffar, 2013).

Another validity is construct validity, which is considered a higher-level concept because the "constructs are higher level concepts which are not directly observable or measurable (nature) while variables (sometimes used interchangeably with indicators or measures) seek to measure the underlying construct (nature exposed to our method of reasoning)" (Agarwal, 2012, p.2). Thus, the construct is latent because it is not directly observed. The hypothesis is that an individual's responses to each of the survey questions are influenced by the underlying latent construct, which can be measured through responses to questions related to the construct (Heisenberg, 1958; Morrison, n.d.).

Another concept associated with construct validity is dimension, which is related to the construct. In a survey study, the questionnaire scales consist of a series of questionnaire items, which are organized in different groups, which may be the external representation of the construct. Therefore, assessing questionnaire dimensionality is one aspect of validating the internal structure of the questionnaire scale (Rios, & Wells, 2014).

*2.3 Dimensionality of Questionnaire Scales*

When they initiate engaging in a survey research project, the researchers may not consider how to organize these questionnaire items in different groups and the meaning of grouping these items. Thus, examining the dimensions ensures the survey quality and the construct validity (ER Services (n.d.).

## 3. Modes of Inquiry

The issues of dimensionality are related to several aspects of the questionnaire quality such as reliability and validity. The dimensionality is also related to convergence. When working with multiple constructs in a survey study, it is important to satisfy convergent and discriminant validities in order to satisfy construct validity. If researchers can demonstrate that they have evidence for both convergent and discriminant validity, then the researchers demonstrate that they have evidence for construct validity (Trochim, Donnelly, & Arora, 2006).

One method to strengthen the convergent and discriminant validity is the reduction of the dimensions of the multidimensional scales (Sarveniazi, 2014; Weng, & Young, 2017; Zhang & Takane, 2010).

This study is to examine how to strengthen the convergent validity by reducing the dimensions of the multivariate constructs. The authors compared two reduction methods, multidimensional scaling and factor analysis, to evaluate different findings and results.

## 4. Data Resources and Evidence

In this study, there were 110 participants participated in the survey questionnaire. The focus of the survey had questions about the frequency of academic dishonesty.

There are 23 items in the questionnaire scales. These items can be classified into 4 dimensions: Cheating, Plagiarism, Obtaining Unfair Advantage and Falsification of records and Official Documents. These items were associated with 9-point Likert scales where 1 stood for "strongly disagree" and 9 for "strongly agree". One hundred and 8 college students responded to the questionnaire.

*4.1 Multidimensional Scaling for Dimensionality Reduction*

Since real-world data is usually highly unstructured, the extraction of features of the data for analysis challenged researchers. Therefore, dimension reduction is the one of effective strategies to better represent the structure and construct of a set of data. Wijaya (2020) summarized several reasons why we have to reduce the dimensions of a survey instrument:

> A higher number of features increase variance in data, which could cause overfitting — Especially where the number of observations is less than the amount of the features present. The density and distance between data become less meaningful, which means the distance between data is equidistant or equally similar/different. This affects clustering and outlier detection as critical information from the data is undervalued. Combinatorial Explosion or a large number of values would lead to a computationally

intractable problem where the process just takes too long to finish. (p. 2)

There are 30 questions in the questionnaire of Awareness of Academic Dishonesty. As shown in Figure 1, the questionnaire items fall into two groups roughly Group One consists of A24, A1, A18, A8, A6, A29, A15, A2, and A25. Group Two consists of A7, A10, A20, A30, A5, A9, A14, A19, A4, A17, A11, A13, A16, A12, A3, A22, A23, A26, and A28. "The data for Multidimensional Scaling (MDS) analyses is usually coined as proximities, which indicate the overall similarity of the elements in the data" (Imperial, 2019). MDS look for a special configuration of the elements so that the distance between the elements matches their proximities as closely as possible. The outputs showed that there are two "piles" of data points. Each pile of data points indicates there is a similarity among these data points.
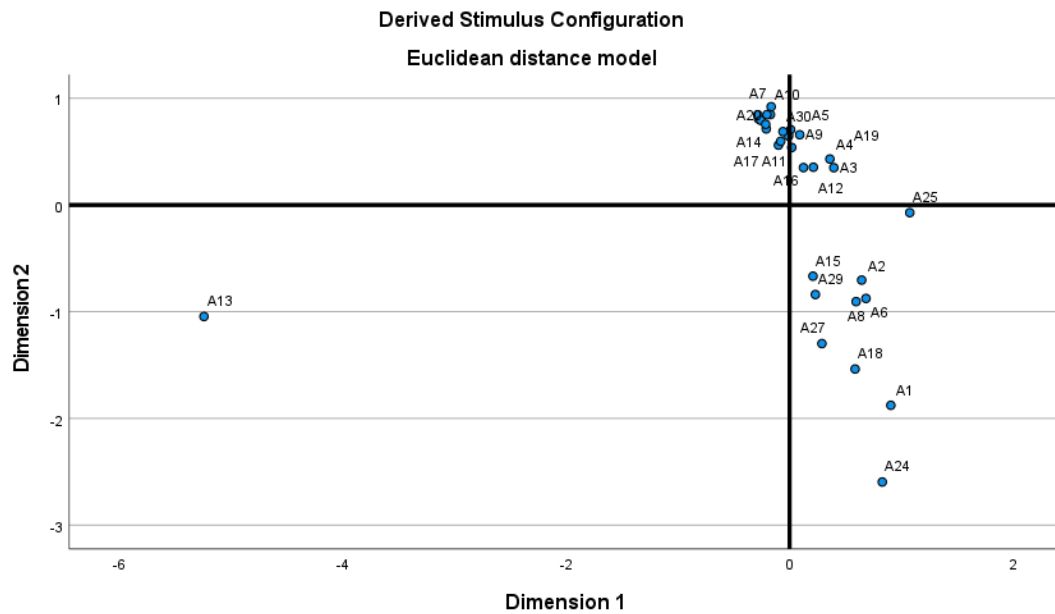


Figure 1. An Euclidean Distance Model of the Awareness of Academic Dishonesty

*4.2 Dimensionality Reduction With Factor Analysis*

Exploratory factor analysis (FA) is a dimensionality reduction technique that attempts to group intercorrelated variables together and to produce interpretable outputs (Henrique, 2021). Exploratory factor analysis assumes that there are several latent variables in a model. These variables are unobserved but they consist of a construct and also these variables explain a significant proportion of the variation common among the manifest variables. This study is to examine the latent variables of the questionnaire. This dimensional reduction is a different practice from a latent construct perspective. Again, there are 30 questions in the questionnaire on Awareness of Academic Dishonesty.

As shown in Figure 2, there are three dimensions in this study. Items A7, A16, and A22 consist of dimension 1. Items A1, A13, and A24 consist of dimension 2, and the other 30 items are in dimension 3. Exploratory factor analysis reported three dimensions. The result is slightly different from the one of MDS.

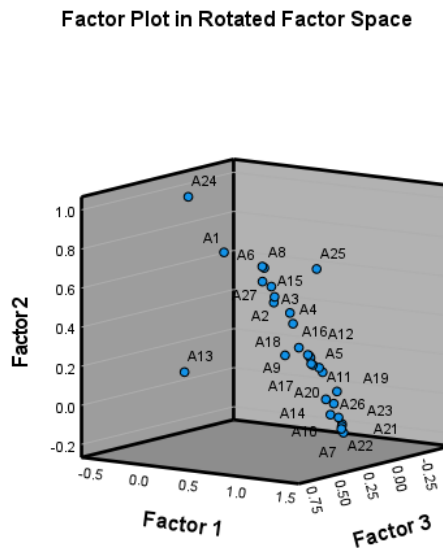**Factor Plot in Rotated Factor Space**



Figure 2. A Factor Plot of the Questionnaire of Awareness of Academic Dishonesty

## 5. Conclusions

Two dimensional reduction methods were introduced for the analysis of the questionnaire on Awareness of Academic Dishonesty. Multidimensional Scaling reduced the dimensions of the questionnaire to two dimensions. Exploratory factor analysis reduced the dimensions of the questionnaire to three dimensions.

The assumptions and philosophies are different between these two-dimensional reduction methods. MDS is a non-linear dimensionality reduction technique that tries to preserve the distances between instances while reducing the dimensionality of non-linear data. Thus, the new researchers examined the questionnaire dimensions and visualized the distributions of all of these data points. They know the questionnaire and the data better via evaluating the data point pattern in the reduced data space.

This analysis aids beginning researchers to guarantee the quality of the survey. This analysis provides a good foundation for the next step of analyzing latent variables and constructing them with reduced dimensions. Factor analysis is used to develop a new set of uncorrelated variables, called latent variables, with the hope that these new variables better represent the construct of the data being analyzed. The factor analysis model assumes there is a smaller set of uncorrelated latent variables driving the value of the variables that are actually being measured. This is an advanced step to know the construct of a set of measuring structures.

## 6. Scholarly Significance of the Study

This study examined the dimensionality of the questionnaire on Awareness of Academic Dishonesty. Multidimensional Scaling (MDS) and Exploratory Factor Analysis (EFA) were applied to the data. MDS recognized two dimensions of the 30 questionnaire items. The factor analysis reported a 3-dimension construct. This study recommends an effective strategy. When beginning researchers develop their questionnaire as an instrument of the survey study, they can use MDS to examine the patterns and distribution of the data points and variables. Typically, the factor analysis provides an advanced tool for inspecting the construct, so as to achieve the purpose of ensuring the quality of the questionnaire.

## 7. Limitations

Multidimensional scaling and factor analysis are two different methods to reduce the dimensionality of the scale such as questionnaires. However, the theories and assumptions are different. Beginning researchers should carefully think about the dimensions and constructs and then make a choice between them.

## References

Agarwal, N. K. (2012, January). *Verifying survey items for construct validity: A two-stage sorting procedure for*

*questionnaire design in information behavior research.* https://doi.org/10.1002/meet.2011.14504801166

Baker, R., & Keeter, W. S. (2022). Evaluating survey quality in today's complex environment. *American Association for Public Opinion Research*. Retrieved July 14 2022, from https://www.aapor.org/Education-Resources/Reports/ Evaluating-Survey-Quality.aspx

Baker, R., Brick, M. J., Bates, N. A., Battaglia, M., Couper, M. C., Dever, J. A., … Tourangeau, R. (2013). *Report of the AAPOR task force on non-probability sampling*. Retrieved July 13, 2022, from https://www.aapor.rg/aapor_main/media/mainsitefiles/ nps_tf_report_final_7_revised_fnl_6_22_13.pdf

Biemer, P. P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly, 74*(5), 817-848.

Bolarinwa, O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal, 22*, 195-201.

Bremer, J. (2013). Research quality: The interaction of sampling and weighting in producing a representative sample online: An excerpt from the ARF's foundations of quality 2 initiative. *Journal of Advertising Research, 53*(4), 363-71.

Deniz, M. S., & Alsaffar, A. A. (2013). Assessing the validity and reliability of a questionnaire on dietary fibre-related knowledge in a Turkish student population. *Journal of Health Population and Nutrition, 31*(4), 497-503

ER Services. (n.d.). *Research methods in psychology*. Retrieved, July 16, 2022, from https://courses.lumenlearning.com/suny-hccc-research-methods/chapter/chapter-6-measurement-of-construc ts/

Heisenberg, W. (1958). The Copenhagen interpretation of quantum theory (Chapter 3). In G. Allen, & Unwin (Eds), *Physics and philosophy*. Retrieved May 31, 2011, from http://www.marxists.org/reference/subject/philosophy/works/ge/heisenb3.htm

Henrique, A. (2021. Nov. 15) *Dimensionality reduction with factor analysis on student performance data.* Geek Culture. Retrieved July 18, 2022, from https://medium.com/geekculture/dimensionalityreduction-with-factor-analysis-on-student-performance-data -fd4ca7082f63

Imperial, J. (2019). The multidimensional scaling (MDS) algorithm for dimensionality reduction. *Data Driven Investor.* Retrieved July 18, 2022, from https://medium.datadriveninvestor.com/the-multidimensional-scaling-mds-algorithm-for-dimensionality-red uction-9211f7fa5345

Kember, D., & Leung, D. Y. (2008). Establishing the validity and reliability of course evaluation questionnaires. *Assessment & Evaluation in Higher Education, 33*, 341-53.

Morrison, J. (n.d.). *Assessing questionnaire validity.* Select Statistical Service. Retrieved July 15, 2022, from https://select-statistics.co.uk/blog/assessing-questionnaire-validity

Nevill. A. M., Lane, A. M., Kilgour, L. J., Bowes, N., & Whyte, G. P. (2010). Stability of psychometric questionnaires. *Journal of Sports Sciences, 19*(4), 273-278.

Rios, J., & Wells, C. (2014). *Validity evidence based on internal structure, 26*(1), 108-116.

Sangoseni, O., Hellman, M., & Hill. C. (2013). Development and validation of a questionnaire to assess the effect of online learning on behaviors, attitude and clinical practices of physical therapists in United States regarding of evidence-based practice. *Internet journal of allied health sciences and practice, 11*, 1-12.

Sarveniazi, A. (2014). An actual survey of dimensionality reduction. *American Journal of Computational mathematics, 4*, 55-72.

Trochim, W. M., Donnelly, J. P., & Arora, K. (2006). *Research methods: Essential knowledge*. Cengage Learning.

Weng, J., & Young, D. S. (2017). Some dimension reduction strategies for the analysis of survey data. *Journal of Big Data, 4*(43), 1-19.

Wijaya, C. Y. (2020). *Five must-know dimensionality reduction techniques via prince*. Towards Data Sciences. Retrieve, July 18, from https://towardsdatascience.com/5-must-know-dimensionality-reduction-techniques-via-prince-e6ffb27e55d1

Wong, K. L., Ong, S. F., & Kuek, T. Y. (2012). Constructing a survey questionnaire to collect data on service quality of business academics. *European Journal of Social Sciences, 29*, 209-221.

Zach, O. (n.d). *A simple explanation of internal consistency.* Statology. Retrieved, July 15, 2022, from https://www.statology.org/internal-consistency/

Zhang, Z., & Takane, Y. (2010). Statistics: Multidimensional scaling. In E. Baker, B. McGaw, & P. Peterson (Eds.), *International Encyclopedia of Education* (3rd ed.). Oxford, UK: Elsevier.