IOSR Journal of Research & Method in Education (IOSR-JRME)
e-ISSN: 2320–7388,p-ISSN: 2320–737X Volume 7, Issue 2 Ver. I (Mar. - Apr. 2017), PP 75-79
www.iosrjournals.org

Estimation of Greatest Lower Bound Reliability of Academic Delay of Gratification Scale

Asst. Prof. Rajib Chakraborty,
(Ghulam Ahmed College of Education, Osmania University, Hyderabad, Telangana, India.

Abstract: *The present study tried to find the best estimation of the reliability of the academic delay of gratification scale, originally prepared by Hefer Bembenutty and Stuart Karabenick, as per the nature of the data obtained on the administration of this scale on the selected sample.479 professional courses students from engineering, law, education and pharmacy of Sultan Ul Uloom Education Society, Hyderabad, were part of the study. The scale was found to be congeneric, with unidimensionality and unequal factor loadings of the items on the factor ADOG through confirmatory factor analysis using SPSS AMOS ver.23.Komolgorov-Smirnov test and Shapiro-Wilk test using SPSS Statistics Ver.23 found the data not to be normal. Under the violation of tau-equivalance, congeneric model and asymmetrical circumstances, R Studio and R softwares were used to find the point estimate of Greatest Lower Bound reliability, as the best estimation of this scale's reliability, which was found to be 0.75. Cronbach's alpha, Omega coefficient, Guttman's Lambda 2, and Composite Reliability, were found to be nearly equal, in and around the magnitude of 0.7 and underestimated the scale's true reliability in these realistic conditions. Educational implications are discussed.*
Keywords: *Academic Delay of Gratification Scale, Cronbach's Alpha, Composite Reliability, Greatest Lower Bound, Guttman's Lambda 2, Omega Coefficient.*

I. Introduction

The Academic Delay of Gratification Scale was prepared by Hefer Bembenutty and Stuart.A. Karabenick to measure the "*willingness to postpone immediately available opportunities to satisfy impulses in favor of academic goals that are temporally remote but ostensibly more valuable*", in the American college students, in 1996. This study reported the 10 items scale's Cronbach's alpha to be 0.77 as the measure of its internal consistency [1].

Though Cronbach's alpha is the most popular estimate of reliability in applied research [2], its limitations are also well known, like the assumptions of tau-equivalence [3] and normality [4]. Violation of the assumption of tau-equivalence leads to under estimation of the true reliability of a scale [5], [6]. This amount of the underestimation of the true reliability can be anywhere between 0.6 to 11.1% depending on the seriousness of the violation [7]. In practice, it is difficult to get data where the equal factor loading of items on a factor of a scale takes place [8]. Different items load to different extent on the factor in reality, which constitutes the congeneric model of measurement.

The omega coefficient proposed by McDonald [9], is cited as the best estimate of reliability when both tau-equivalence exists and does not exists (congeneric measurement) and this corrects the underestimation of reliability by Cronbach's alpha caused due to the violation of tau – equivalence [10]. Raykov's composite reliability for congeneric measures also is an estimator of reliability which does not possess the limitations of Cronbach's alpha underestimation property [11]. When tests are administered only once on the subjects for data collection as part of practical constraints, reliability is then defined as the ratio between true score variance and observed score variance and called the lower bound of reliability after Guttman who proposed six different forms of lower bound reliability from Lambda 1 to 6. However, Guttman proved that among these 6 forms, only lambda 2 is closer to the true reliability and this specific coefficient was found to perform better in a simulation study over the rest of the Guttman lambda reliability estimators for heterogeneous items [12].

But, in reality, the researchers also deal with non-normal or skewed data distribution of the sample. In such situations, the performance of alpha, omega and lambda 2 coefficients as estimators of reliability is doubtful and the existence of a new estimator of reliability is desired. One of the most powerful and less known estimators of reliability based on the assumptions of Classical Test theory is the Greatest Lower Bound (GLB) [13]. It produces better results over alpha [14] and together over alpha and omega [15]. A simulation study to find this estimator's functioning in non-normal conditions or asymmetrical distributions, in comparison to the functioning of alpha and omega was carried out [16]. It found that alpha and omega had unacceptable performance under asymmetrical conditions with their bias percentage greater than 13 % and between 1-2 % respectively. But, GLB had better performance even when the skewness value was raised to 0.5 and 0.6.

In this way, GLB is the lowest possible value that a scale's reliability can possess. On calculating this estimator, the scales' reliability, by its very definition, can lie in the interval (GLB,1). Moreover, alpha is affected by the number of items or length of the test and dimensionality, and that is why a series of diagnostics like conducting a confirmatory factor analysis to reveal the dimensionality, inspection of item distribution and computation of alpha, omega and GLB in confidence intervals instead of point estimation to enhance the information value, are recommended to the researchers. Alpha as an estimator of reliability should be abandoned completely [17]. Hence, this study is taken up to estimate the GLB of academic delay of gratification scale along with the estimation of alpha and omega to compare the best estimate with the rest of the underestimates of reliability.

II. Methodology

The tool academic delay of gratification scale consists of 10 items based on academic experiences like "meeting deadlines on assignments, use of the library, interpersonal relations with peers and instructors, and studying course materials". A sample item is like "Go to a favorite concert, play or sporting event and study less for this course even though it may mean getting a lower grade on an exam you will take tomorrow, or Stay home and study to increase your chances of getting a higher grade".

The students respond on a four point Likert scale with responses like "Definitely choose A", "Probably choose A, Probably choose B, Definitely choose B", which are coded from 1 to 4. The total score of a student is obtained by adding all the responses. The total score is divided by 10 to get the mean score as the measure of academic delay of gratification. The higher is the mean score between 1 and 4, the greater is the presence of the characteristic in the subject.Initially, the Academic Delay of Gratification Scale was administered on 479 undergraduate professional courses students from Engineering, Law, Pharmacy and Education of the Sultan Ul Uloom Education Society professional courses colleges in Hyderabad, India. The tests were conducted during classroom sessions by taking prior permission from the Principal of the respective colleges and with the assistance of faculty members. The students took 10 minutes on average to fill the instrument.

Confirmatory factor analysis was carried by using SPSS Amos Ver. 23 on the collected data to confirm the uni-dimensionality of the construct as reported by the authors [18]. Though for a good fit, Chi-square value should be non-significant for 0.05 threshold [19], it does sometimes provide significant value and erroneously indicate a poor fit, as the test is sensitive to sample size [20]. Owing to this reason, degree of freedom Df and probability p values should be reported, along with CMIN/DF which for a good fit has a value less than 3 [21], [22]. RMR and RMSEA goodness of fit indicators for a model must have values less than 0.08 and GFI. IFI, TLI and CFI indicators must have their recommended values above 0.93 for a good fit model [23]. The magnitudes of the factor loadings post CFA were to reveal violation of tau-equivalence and the congeneric nature of the measurement model.Tests of normality like Komolgorov-Smirnov test and Shapiro-Wilk's test using SPSS Statistics Ver.23 of the 10 items was carried out. While composite reliability was calculated using the online calculator [24], R studio and R softwares were used to calculate Cronbach's alpha, Omega coefficient and Greatest Lower Bound in point estimate and confidence interval forms [25],[26]. Lambda 2 value for calculated using SPSS Statistics Ver.23, along with its other forms.

III. Results

Table 1: Goodness of Fit Recommendations and Obtained Statistics from Confirmatory Factor Analysis

Measure	P value	CMIN/DF	RMR	RMSEA	GFI	IFI	TLI	CFI
Benchmark	>0.05	< 3	< 0.08	< 0.08	>0.93	>0.93	>0.93	>0.93
Result	0.001	1.950	0.050	0.045	0.971	0.934	0.913	0.932

Though the obtained p value for Chi-square test is less than the benchmark, it can be expected to be so for a large sample size of n = 479. It does not mean that the result is significant and the model is a poor fit. Expect TLI, all other indicators of goodness of fit, show that the scale is unidimensional in structure.

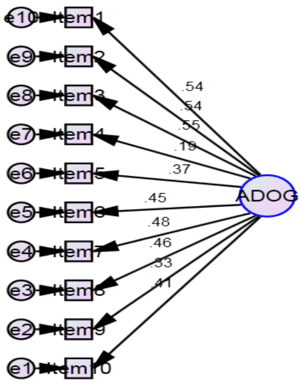

Figure 1: Path diagram and the Factor Loadings of 10 Items of ADOGS from Confirmatory Factor Analysis

Since the factor loadings of the ten items of the scale vary from 0.19 to 0.55, the present measurement model is congeneric and tau-equivalance assumption is violated.

Table 2: Tests of Normality

	Kolmogorov-Smirnov*			Shapiro – Wilk*		
	Statistic	df	Sig.	Statistic	df	Sig.
Item 1	.201	479	.000	.831	479	.000
Item 2	.210	479	.000	.856	479	.000
Item 3	.206	479	.000	.844	479	.000
Item 4	.408	479	.000	.635	479	.000
Item 5	.196	479	.000	.860	479	.000
Item 6	.186	479	.000	.864	479	.000
Item 7	.183	479	.000	.868	479	.000
Item 8	.257	479	.000	.807	479	.000
Item 9	.207	479	.000	.829	479	.000
Item 10	.230	479	.000	.830	479	.000

Lilliefors Significance Correction
* Both the tests are sensitive to sample size and report significance even for slight deviation from normality.

The Kolmogorov-Smirnov test and the Shapiro – Wilk test work based on the null hypothesis that the submitted data is that of a normal distribution. Obtaining a significance result for all the items, post test implies the null hypothesis is to be rejected and the alternate hypothesis of the items in the measurement model being non-normal is to be accepted.

Table 3: Items' Asymmetry Statistics

Item No.	**N**	**Skewness**			
		Statistics	**Std. Error**	**Statistics / Std. Error**	**Acceptable Range**
Item 1	479	.017	.112	0.151	-1.96 to 1.96 for smaller samples and -2.58 to 2.58 for larger sample.
Item 2	479	.183	.112	1.633	
Item 3	479	-.379	.112	-3.38	
Item 4	479	-1.410	.112	-12.58	
Item 5	479	.196	.112	1.75	
Item 6	479	-.087	.112	-0.776	
Item 7	479	-.120	.112	-1.071	
Item 8	479	-.561	.112	-5.00	
Item 9	479	.287	.112	2.562	
Item 10	479	-.300	.112	-2.678	

Item 3, Item 4, Item 8 and Item 10 have their skewness / std. error ratio beyond the rule of thumb acceptable range of -2.58 to 2.58 for a large sample size like N=479. This indicates the asymmetry or departure from normality in the distribution of these items.

Table 4: Reliability Estimates

S.No.	**Type of Reliability**	**Source of Calculation**	**Point Estimator**	**Confidence Interval**
1.	Cronbach's Alpha	R Studio / R	0.7	(0.66,0.74)
2.	Omega	R Studio / R	0.7	(0.66,0.74)
3.	Guttman's Lambda 2	SPSS Statistics Ver. 23	0.703	-
4.	Composite Reliability	AMOS / Online Calculator	0.699	-
5.	GLB	R Studio / R	0.75	-

Cronbach's alpha, Omega coefficient, Guttman's Lambda 2 and Composite reliability together underestimated the true reliability of the scale. These indicators of reliability are not valid as the items are heterogeneous and their distribution is skewed or asymmetric. Only greatest lower bound is the indicator of reliability under the violation of tau-equivalance, existence of congeneric model and non-normal distribution of items. R Studio / R software found this statistic to be 0.75 for ADOGS. It implies that this scale has an acceptable reliability value when administered in tact on urban professional courses students of India.

IV. Conclusion

Cronbach's alpha is reported as the estimate of reliability in most of the studies, without paying heed to the nature of the construct and the items comprising it. Under realistic circumstances, the assumption of tau-equivalence, gets violated, the measurement model is congeneric and the obtained data of the items has skewness to varying extent in it. In such scenarios, the greatest lower bound reliability of the instrument can be considered as the only estimator of its true reliability to report in studies. Academic delay of gratification is one of the vital constructs to be promoted in the students to make them self regulated learners[27]. To achieve the same, the trait is to be reliably assessed in students from different cultures. This calls for the estimation of the reliability of the ADOGS tool using estimators other than Cronbach's alpha. Though the authors of the tool reported the tools Cronbach's alpha to be lying between 0.70 to 0.84 [28], [29], these estimates cannot be trusted any more for the apparent limitations of Cronbach's alpha as an estimator of reliability. For instance, alpha is related to both internal consistency and homogeneity[30]. While former relates to the degree of interrelatedness among the items [31], [32], the later refers to unidimensionality [32],[33],[34]. Alpha is a function of internal consistency or interrelatedness of items which is a necessary but not a sufficient condition for homogeneity. Moreover, it has been shown to be unrelated to internal consistency of a test in certain studies [2]. It is also dependent on number of items, sample of the test and dimensionality of the construct. Ignoring of these limitations and violation of assumptions of essential tau-equivalence should stop and the practice of mentioning Cronbach's alpha as the tool's reliability estimator should be abandoned and a more comprehensive assessment of the scale's reliability through greatest lower bound estimation should be adopted [17].

The present study established the unidimensionality of the academic delay of gratification construct through confirmatory factory analysis and then estimated the greatest lower bound reliability of the ADOGS tool to be 0.75, which is an acceptable value of this psychometric property [35]. It also showed how other estimators like composite reliability, Omega coefficient and Guttman's lambda 2 also underestimate reliability when the items have asymmetric or skewed distribution. The estimation of the comprehensive GLB reliability of ADOGS tool paves the way for its administration on Indian students with a fair degree of confidence. However, India is a culturally diverse nation and the replication of this study in other states and comparison of the results would be a very fruitful enterprise to take up by other educational researchers. The study should also be replicated on rural students, students from different socio-economic background and ethnicity with larger sample size.

References

[1] Bembenutty, H., & Karabenick, S. A., Academic Delay of Gratification Scale. A New Measurement for Delay of Gratification, Paper presented at the *Annual Meeting of the Eastern Psychological Association*, Philadelphia, PA, March 1996.

[2] Sijtsma, K., On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107,2009.

[3] Cronbach, L. J., Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), 297-334, 1951.

[4] Green, S. B., & Yang, Y., Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika*, *74*(1), 155-167,2009.

[5] Raykov, T.Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, *32*(4), 329-353,1997.

[6] Graham, J. M., Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and psychological measurement*, *66*(6), 930-944,2006.

[7] Green, S. B., & Yang, Y., Commentary on coefficient alpha: A cautionary tale. *Psychometrika*, *74*(1), 121-135,2009.

[8] Teo, T., & Fan, X., Coefficient alpha and beyond: Issues and alternatives for educational research. *The Asia-Pacific Education Researcher*, *22*(2), 209-213, 2013.

[9] McDonald, R. P., *Test theory: A unified treatment* (Psychology Press, 2013).

[10] Dunn, T. J., Baguley, T., & Brunsden, V., From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399-412, 2014.

[11] Raykov, T., Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, *21*(2), 173-184,1997.

[12] Tang, W., Cui. Y., A Simulation Study for Comparing Three Lower Bounds to Reliability, Paper presented on April 17 2012 at the *AERA division D: Measurement and Research Methodology,* Section:1, Educational Measurement, Psychometrics and Assessment..

[13] Woodhouse, B., & Jackson, P. H. (1977). Lower bounds for the reliability of the total score on a test composed of non-homogeneous items: II: A search procedure to locate the greatest lower bound. *Psychometrika*, *42*(4), 579-591.

[14] Lila, M., Oliver, A., Catalá-Miñana, A., Galiana, L., & Gracia, E., The intimate partner violence responsibility attribution scale (IPVRAS). *The European Journal of Psychology Applied to Legal Context*, *6*(1), 29-36, 2014.

[15] Wilcox, S., Schoffman, D. E., Dowda, M., & Sharpe, P. A. (2014). Psychometric properties of the 8-item english arthritis self-efficacy scale in a diverse sample. *Arthritis*.

[16] Trizano-Hermosilla, I., & Alvarado, J. M. , Best Alternatives to Cronbach's Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements. *Frontiers in psychology*, *7,* 2016.

[17] Peters, Gjalt-Jorn Y. "The alpha and the omega of scale reliability and validity: why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality." *European Health Psychologist* 16.2,: 56-69, 2014.

[18] Bembenutty,. H., Karabenick, S. A.,"Academic Delay of Gratification", *Learning and Individual Difference,* Volume 10, Number 4, 1998, pages 329-346, ISSN: 1041-6080, 1998.

[19] Barrett, P., "Structural Equation Modelling: Adjudging Model Fit," *Personality and Individual Differences*, 42 (5), 815-24, 2007.

[20] Gerbing, D. W., & Anderson, J. C.," The effects of sampling error and model characteristics on parameter estimation for maximum likelihood confirmatory factor analysis", *Multivariate Behavioral Research*, 20, 255−271, 1985.

[21] Kline, R.B., *"Principles and Practice of Structural Equation Modelling*" (2nd Edition ed.). New York: The Guilford Press, 2004.

[22] Hayduk, L., Cummings, G.G., Boadu, K., Pazderka-Robinson, H., and Boulianne, S., "Testing! Testing! One, Two Three – Testing the theory in structural equation models!," *Personality and Individual Differences*, 42 (2), 841-50, 2007.

[23] Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., Ouwerkerk, J. W., & Spears, R., "Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification", *Journal of Personality and Social Psychology*, 95(1), 144-165, 2008.

[24] Colwell, Scott R., The composite reliability calculator. Technical Report, DOI: 10.13140/RG.2.1.4298.088, 2016.

[25] RStudio Team, RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, URL http://www.rstudio.com/, 2016.

[26] R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/, 2017.
[27] Bembenutty, H., Self-regulation of learning and academic delay of gratification: Gender and ethnic differences among college students. *Journal of advanced academics*, *18*(4), 586-616, 2007.
[28] Bembenutty, H., Sustaining motivation and academic goals: The role of academic delay of gratification. Learning and Individual Differences, 11, 233–257, 1999.
[29] Bembenutty, H., & Karabenick, S. A., Inherent association between academic delay of gratification, future time perspective, and self-regulated learning Educational Psychology Review, 16, 35–57, 2004.
[30] Cortina, J. M. , What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*, *78*(1), 98, 1993.
[31] Crano, W. D., & Brewer, M. B., *Principles of research in social psychology*. McGraw-Hill, 1973.
[32] Green, S. B., Lissitz, R. W., & Mulaik, S. A. , Limitations of coefficient alpha as an index of test unidimensionality1. *Educational and Psychological Measurement*, *37*(4), 827-838, 1977.
[33] Gulliksen, H., The reliability of speeded tests., Psychometrika, 15, 259-269, 1950.
[34] Lord, F. M., Novick, M. R., & Birnbaum, A., *Statistical theories of mental test scores*, Reading, MA: Addison – Wesley, 1968.
[35] George,D., Mallery,.P., SPSS for Windows Step by step: A Simple Guide and Reference, 11.0 update (4^{th} ed.), Boston: Allyn & Bacon, 2003.