Scholars Journal of Physics, Mathematics and Statistics

OPEN ACCESS

Abbreviated Key Title: Sch J Phys Math Stat
ISSN 2393-8056 (Print) | ISSN 2393-8064 (Online)
Journal homepage: https://saspublishers.com

Application of Generalizability Theory in the Estimation of Dependability of Critical Thinking Scale for University Students

Sunday Dorathy[1*], Dr. A. Amadioha[1], Dr. G. W. Orluwene[1]

[1]Department of Educational Psychology, Guidance and Counselling, Faculty of Education, University of Port Harcourt, Port Harcourt, Rivers State, Nigeria

DOI: 10.36347/sjpms.2021.v08i09.002 | **Received:** 07.10.2021 | **Accepted:** 14.11.2021 | **Published:** 21.11.2021

***Corresponding author:** Sunday Dorathy

Abstract **Original Research Article**

This study focused on the application of generalizability theory to estimate the dependability of Dorathy's Critical Thinking Scale (DCTS) for university students in the South-South Geo-Political Zone of Nigeria. Three research questions guided the study. A G-study design was used for the study. A stratified proportionate random sampling technique was used to draw a sample size of 1,041 participants from the population. The reliability coefficient realized from the pilot study was 0.86. Research question one was answered using ANOVA by variance component, research question two was answered using relative and absolute errors, while research question three was answered using dependability coefficient. The result showed that the DCTS has a good variance component. For items it was .003 with 1.42%, occasion revealed .007 with 3.30% while the error was 0.202 with 95.3%. The relative and absolute errors were 0.0014 and 0.0111. The dependability coefficient for DCTS was 0.63. Recommendations were made including that stakeholders in the universities should imbibe evaluating students with the use of critical thinking because it is dependable. Also, G-theory should be used to determine the reliability of the instrument as it would result in multiple sources of variations.

Keywords: Dependability, Generalizability Theory, Critical Thinking.

INTRODUCTION

University is a higher educational education where secondary schools, colleges and polytechnics leavers attend to further their educational program for acquisition of higher knowledge, competencies and self-reliance to enable them to handle the issues of life. It is at the university that the degree certificates are awarded at various stages such as the first degree, master's degree and the Doctorate depending on the individual student's level of programs.

From the South Utah University (2016) a university is referred to as an institution of people who require higher learning which is found in running different programmes, they study numerous courses with research facilities that may prompt them to perceive construct in the manner they think it can be executed. Orluwene & Sunday (2018) viewed university as an institution where people are nurtured on how to produce current concepts and eliminate the obsolete ones that may contribute to the reduction of individual and the nation's growth and development.

Kwok (2015) reported that the purpose of the university is to enlarge the mindset and horizon of students to enhance their ability to relate issues and carry out analysis of problems adequately for modification of the society successfully. University students may be independent people who have individual responsibilities to think in different dimensions as to organize and solve the challenges in academic studies. Fermy though (2010) defined thinking as a conscious and active construct and a type of cognitive process that helps students in the production of new knowledge. On the other hand, critical involves analyzing the problems and prospects of anything in life before acting on it.

So what then is critical thinking? Conceptually, critical thinking means being in the state of thinking of objects, entities, content, subjects, challenges where the persons thinking enhances the level of thinking efficiently through analysis, assessment and revising it. In other words, any student who thinks critically can have self-direction, self-discipline, independent monitoring, and correcting one's

self-thinking. In consideration of what it means to think critically Heick (2017) considered it to be the first causation of change of reasoning in the mind of individual towards everything taught and the environment, critical thinking is viewed in a different perspective, educationally; thinking is a method to get information criticized which can be the beginning of clarifying the issues on the ground.

However, to realize adequate thinking, several principles may guide university students to think critically; they have to elicit comprehensive information, comprehend and provide accurate definitions of the words; They also have to ask questions on the procedures that can result in any matter and how it ended, get some responses, evaluate what is being said, ascertain the incident that led to it as it is well as consequences. In other words, for one to think critically facts are inevitable for accurate directions.

According to Thomas (2017), critical thinking is vital in informing students on the aptitude to future in a particular aspect of task presented to them by the lecturers; this enables them to be independent at accomplishing responsibility than relying on the efforts of other people's judgments. Again awareness of what is happening politically, philosophically, socially and otherwise are achieved as the students develop opinions critically. In education, critical thinking allows the cognitive capacity of people to transform meanings and disperse ideas that leads to having satisfied feeling in all aspects of life, so that educators will have teaching organization on what they will inculcate into the students.

Emphatically, critical thinking helps increase the students thinking, self-autonomy, dimensional judgments in reasoning and activities. Also, it enhances the students' ability to reason well, particularly those that have the deposition to manifest it; and it motivates students to think logically and be creative. Some of the characteristics of students who think critically may gain include; assumption, deductive, problem solving, inductive, interpretation, definition, analysis that will influence their educational performances positively. In line with that, Lenin (2017) concluded that critical thinking is very important in all fields of life; mostly in the areas that involve human beings, this is because it enables the lecturers to inculcate such level of thinking into the students following several approaches synonymous to the features of critical thinking.

The use or application of critical thinking, academic testing and the consistent researching of the students may aid the elimination of beliefs and opinions that exist in facts rather they can evaluate them with good reconciliation with the facts (Michelle, 2015).

Similarly, critical thinking is accurate in instilling curiosity in the students to ask relevant questions, reject information that is not disseminated, so that they can find the true basis of issues with open-mindedness (Murat 2016).

Critical thinking is presently viewed as one of the highly valued outcomes of education in the entire spectrum of education, mostly in higher professional education, which includes university education. So to achieve a university education that includes graduating competent knowledge and talented individuals, the undergraduates have to solve problems through thinking critical in a variety of ways, which means students need to be good at synthesizing information to be competent in decision making and self-regulatory judgment. Thus, there should be regular assessment of critical thinking among students using a dependable instrument.

A dependable instrument is adequately determined by using generalizability theory instead of classical test theory. This is because with G-theory multiple sources of errors in any given test are disentangled and identified. G-theory is a method of determining the reliability of test scores that disentangled the multiple sources of errors as well as estimate their relative contributions to the measurement error.

The term generalisability theory is a statistical word, it means to expand the outcome of research beyond the area where the study was focused on a large scale to achieve a research work adequately, with there is a need to elicit data from a large population. Another name for it is G theory, where G stands for generalisability. Cronbach, Gleser Nanda and Rajaratnam 1972 explained that the use of ANOVA mean squares for models with different effects helps the reliability estimate for multiple sources of variation such as occasion, items and raters in a testing situation. The testing of such reliability estimate is known as generalizability theory. They were the first to use generalisability theory in estimating reliability with what people can do or work performance assessments; hence it serves as an advanced Cronbach alpha reliability method of internal consistency. Generalizability theory is also called statistical theory which can be used for the evaluation of dependability of measured behaviour. Heitman, Kovaleski & Pugh (2009) opined that G theory is a method of modifying the measurement process to gain reliability in place of classical test theory, which assumes reliability as a single undifferentiated source of measurement errors. To determine and explain the multifaceted nature of measurement error, generalizability theory is used.

However, it is necessary to investigate the dependability of the Dorathy Critical Thinking Scale of university students that possess the characteristics of how university students should think to realize their aims of attending higher educational programs. Hence, generalizability theory explains the degree to which the

result of a study can be accurately be used beyond the scope of the study, so it ensures that research is qualitative if it portrays the actual condition when it is used outside. Briesch, Swaminathan, Welsh & Chafoleas (2013) summarized that generalizability theory produces numerous uses for assessing research for enabling concurrent examination of different ways of variance and give a kind of feedback to relatively and adequately make a decision as well as ascertained the reliability and generalization of the findings holistically.

It is a theory used to overcome the challenges of classical test theory in estimating reliability because it gives the researcher opportunity to determine various sources of variance components and how variables interact. The systematic guide in the effective use of G theory could aid researchers in the opportunity to increase the precision of the construct measured and equally build their confidence in the outcome or findings of the research being executed as it also reveals the dependability and replicates scores Lakes & Hoy (2009). Dependability is the adequacy of generalizing test-takers observed score of the construct of measurement to the mean scores of the test takers considering all conditions allowed by the test users. This implies that the score of the test taker scores are made up of errors from many sources rather than assuming that it is as a result of changes; in other words, test takers score from one testing point cannot be dependent on due to variations.

It also allows the user to find out the number of times dependability is realized by the test takers from the scores relating to its coefficient. It gives them opportunities to evaluate dependable scores that are interpreted in different including total achievement. Generalisability theory can be dependable if the finding factor is generalized from a single of the homogeneity to many in that group example generalization should be made to one item and letter extend to the broader items, and their outcomes or responses will differ.

The word dependability is the measure of a system that reveals a particular theory, the dependability of generalizability theory can be likened to the availability, reliability, maintainability, durability, of this study.

A dependable instrument is a reliable instrument. A reliable instrument produces an authentic result leading to an authentic decision about the students. That means a dependable critical thinking scale for university students will help to provide accurate information about the students, then the students and the lecturers will know the next line of action to involve themselves. For instance, if a student is identified to have good critical thinking ability, he/she will be encouraged to maintain it, because the critical ability is very crucial both in the world or university and that of work. On the other hand, for those that will be identified as having very poor critical thinking ability, the teacher or the lecturer will map out the strategies to use to develop good critical thinking ability; this could be through teaching or assessment strategy as Orluwene and Ajala (2019) found that concept test and concept mapping are very effective, lead effective impact on critical thinking abilities of the undergraduate students of the University of Port Harcourt, Rivers State, Nigeria.

Based on the importance of the critical thinking scale, its dependability must be determined through a very strong method such as; generalizability theory (G.T). generalizability theory is a statistical theory that is used for evaluating reliability dependability of a measurement scale Cronbach, Glesser, Nanda, and Rejaratnam in Webb and Shavelon 2005. It is a method of determining the reliability that identifies the different sources of measurement errors, detangling them as well as determining the estimate of the measurement error, by the use of analysis of variance mean squares. That means generalizability theory is better than classical test theory methods when it comes to determining the reliability or dependability of a measurement device. This is because of the entangled and undifferentiated sources of measuring error in classical test theory in separated the use of generalizability theory.

In generalizability theory, an individual test score is a sample obtained from the universe of admissible observation, which is made up of all possible observations that will help decision-makers to take a decision. This theory believes that any given test has the characteristics, features like forms of a test, as in equivalent form, test item, scorer and test occasion as in test-retest etc that are regarded as facets of measurement. On that basis, a universe of admissible observation is defined using all possible combinations of the levels of the facets such as items, rater and occasion.

For instance, in this level where undergraduates attempted 50 items of CTS scored objectively, (dichotomously) on two occasions, so in G-theory the undergraduates students are the object of measurement and then items and occasions are the facets of the measurement which the universe of admissible observation are all the items and occasions that the decision-makers will likely consider to make any decision.

Furthermore, in G-theory as the students are the object of measurement, each of the students' scores will be decomposed into a component for student, item and occasion. The student component of the score is the systematic variation which reflects variability among the students' ability which item and occasion components of the students scores reflect sources of errors.

To separate their sources of errors, G-theory uses analysis of variance approaches to estimate variance components. They based on the variance component of each facet, the dependability of the test will be determined. That means G-theory determines the dependability of a test from the generalization made from students observed score on a test to the score the students would obtain in the broad universe of admissible observation which represents the true score in classical test theory (Shavelson and Webb, 2005).

Furthermore, it is worthy to note that in classical test theory the undifferentiated error term is only considered to be due to random causes. On the other hand, in Generalizability theory, the error term is partitioned into systematic error and random error, which implies that any observed score is a component of the true score, systematic error and random error. ie X= Ta+Os Es + 6r. The systematic error represents facet variability that can be partitioned further based on the number of facets required in the study design. They are also called systemic variance component that is calculated and applied in estimating the dependability of a test.

In G-theory, the two types of reliabilities are differentiated; relative and absolute reliability dependability (Shavelson and Webb, 1991). Relative dependability is the consistency with which examinees or respondents are ranked based on their performances in the test or scale by the two or more different raters. Then absolute dependability is the consistency with which scores occurred around a particular scale point.

To determine the magnitude of the absolute dependability and the G-Coefficient, one must first calculate the relative and absolute variances that will be required in the study. G theory can provide trustworthy findings within the time frame; it is dependable for researchers to conceptualize their statistical framework to have access to the information related to the without restrictions of any kind, such as the partition of scores in numerous sources of variations for identification and estimation of the sources of measurement error, in other to decide to enhance the process in the measurement.

G theory is also flexible in the maintenance of sample tasks where the errors occur frequently for the researcher to make improvements on the consistency of the measurement through the increment of the number of items or tasks to ascertain if the scores are adequately useable. However the duration of adoption of G theory is limited to the sample size during the study to enable the researcher to handle the information effectively in different dimensions but after the study, but after the study, the findings are accurately generalized beyond the sample for any period time practical for long duration. In other words, in depending on generalizability theory there is no fear of time limit in its use, unlike the classical test theory (CTT) it does not have the tendencies of becoming obsolete in short while. To unveil the safety of G theory, several researches can be carried out with it just like an assessment in various angles, applying it in psychological and sociological studies including development or construction of an instrument to have variability (Foote and Huang 2011, Heitman, Kolaeskl and Pugh 2009).

Specifically for this study, G theory can be dependable in the sense that the scores that are realized are interpreted by recognizing and finding out the degree of measurement. G theory has to do with the accuracy of the use of G to obtain from an individual observe score of a measure such as in the students critical thinking to the average score that the students would have gotten if all the conditions are equally considered. If a single score got from a one-time test that was administered once, it may not be regarded as been dependable and cannot be taken as an average score of all the acceptable occasions, this is because the students' scores may not be the same in all the occasions of the test format changes or period of administration of the test that may affect the test taker as the essential aspect sources of inconsistency also known as errors in measurement unlike classical test (CTT) that estimate single sources of error at a period.

Most importantly, G theory allows the researcher to estimate multiple sources of errors in the measures of a study estimating separately in one analysis, this can be achieved by adopting Spearmen Brown Prophecy formulate to establish to have the length of the test. Again, it aids the decision-makers in ascertaining how many occasions as regards to this research, the number of occasions can be determined in test format of the sub construct which is about seven and the domains of administration the same for all the students based on the sapling technique considered to get dependable scores and this may produce a summary of coefficient that reflects the extent of dependability.

Moreover, G theory helps the researcher to decide and investigate how dependable the scores are for various types of interpretations for relative standing of a test taker and absolute level of performance for instance Ngozi scored higher than 85% of her peers for relative standing while the absolute or the real performance of Ngozi can be realized without the consideration of her peers' performance.

The dependability of G theory is very vital in various approaches whether simple of differences terms of achievement or psychological test, observing behaviour, survey researchers and so, one of which is when the measurement is based on one source of variation also known as a facet, this aids the researcher to assess of measuring error from only one source of study from a particular area that can be generalized from an area of study to the larger group for decision

making. In consonance with G theory, since the multiple-choice items of critical thinking scale construct are in three option responses and the researcher has to score them based on correct and incorrect choices of the options made relating to the items of the construct this can be generalized if the student is thinking critically and not on the item, it is essential to develop on the item that has equal difficulty with individual scores that may be from one item to another, the student performance from any sample items may be generalized. In case the item difficulty is not the same, a student score will depend on a particular sample on the test, although it is proper to generalize from the sample to the population.

In addition, the one sample dependability can come in different sources of variability by estimating differences among the participants who are thinking critically and the sources that vary as a result of differences in the difficulty of the construct items which may be difficult to generalize items sample to item universe appropriately. The next has to do with students educational experimental that accompanies the students for the test such as students that are given the test on military aspect those who were brought up in such background has better opportunity to perform higher than those who are from civilian environment. The fourth source of variable gives the researcher the freedom to handle the variability that may occur out of randomness in the application of G theory in the development and validation of the critical thinking scale.

To be précised, G theory is dependable for this study because it allows the estimation of the variation by expressing the magnitude of variability as regards the variance component, so the important sources of variation are assessed which can reduce variations that are not needed. G theory is dependable in two facet universe or population where the research is based on the test that is administered different days, the population of this study is mostly considered in two facets of the items of the construct and the occasion or venue in the definition.

Statement of Problem

Dependability is a very essential aspect of generalizability theory that helps a researcher to ascertain if the study carried out under the theory (G theory) would be relied on in terms of availability, reliability, maintainability, durability etc to confidently conclude that the findings can be generalized beyond the area of the study. Therefore, the problem of this study is to investigate whether a study such as the critical thinking scale when investigated under generalizability theory can produce findings that can be dependable for educationists, students, lecturers, researchers, analysts, and individual countries to work within the society.

Aim and Objectives of the Study

This study aimed to determine the dependability of generalizability theory in the critical thinking scale for universities in Nigeria using generalizability theory. Specifically; the following objectives were considered;

1. Ascertain the magnitude of the facets of the variance components
2. Determine the relative error estimate and absolute error estimate in critical thinking scale (CTS)
3. Establish the index dependability of the critical thinking scale (CTS)

Research Questions

1. What are the magnitudes of the facets/variance components?
2. What are the relative error estimate and absolute error of the CTS?
3. What is the dependability of the critical thinking scale (CTS)?

METHODOLOGY

This study involved the broad blueprint to execute the study and the methods of analyzing data which was aided by answering the stated research questions. Specifically, the study entailed the description of the research design, study area, population for the study, sample and sampling techniques, nature/source of data, method of data collection, validity/reliability of instrument and method of data analysis.

Research Design

This study was a G study designed specifically to separate and estimate the different facets of measurement errors that are feasible. The study involves item and occasion as facets where all the undergraduates were measured equally on all the levels of facets. That is all the 300 undergraduate students who responded to the 50 items on 2 occasions. Thus, the study is a two facets universe study.

Population of Study

The population of this study was 34665 which consist of all the level 300 students in all federal-owned universities in the south-south in Nigeria because they need to think critically to perform very well, reflect on their knowledge and information given to them as it relates to assignments, research and test they have to do, it helps students to be open-minded and also be informed to increase their ability to evaluate and judge issues with evidences which can aid a productive essay and efficiently write papers that can be generally accepted without bias. Critical thinking ability opportunities to control directed, self-discipline, self-monitoring, self-corrective thinking construct, as the students think rationally they communicate effectively when analyzing any concept and can always provide a solution to problems with adequate involvement in making it real for real-life situations. This is because

critical thinking skill is important for would-be graduating graduates and employees/hired of labour). The level 300 students will soon become graduates. The number of universities in the south-south includes; University of Benin (UNIBEN), Benin City, Edo State, Federal University Otuoke Bayelsa, University of Port-Harcourt, University of Uyo, University of Calabar (UNICAL), Calabar, Cross River State for federal owned universities. **Source;** Nigeria Bulletin, List of Federal Universities in Nigeria Cimaobi, K (2014) and Lists of Approved State Universities in Nigeria – Daily school news.

Sample and Sampling Technique

There are five Federal Government-owned universities in the south-south. The sampling technique for this study was stratified random sampling also known as proportionate random sampling to select a total of 1,041 300 level students of about (3%) out of a population of 34,665 third-year students from five universities, as the sample for the study. School location and students' levels of study are the strata for sample selection.

Stratified random sampling is a method of sampling technique that requires the researcher to divide the population of the study into sub-group and the strata are gotten from the participants' common criteria's. Stratified random sampling is used because the researcher tends to dictate a particular sub-group within a population and have a representative sample. Stratified is divided into strata where the sample is usually drawn and this may be very useful if the variable to stratify is easy to handle to select participants from more than one group (Adeleke, 2010). The instrument was administered to the students using an accidental sampling technique whereby the researcher and research assistant administered the instrument on any available three hundred students they met in the lecture halls as at the time of administration of the instrument.

Validity of the Instrument

The instrument was validated by experts in educational psychology, experts in measurement and evaluation and experts in English language to ascertain the grammatical usage and tenses. While the content validity of determining the extent the items of critical thinking was done by judgment panel experts to rate how closely they relate to the domains.

Reliability of the Instrument

The reliability of this instrument DCTS was achieved by first, administration of the instrument to one hundred (100) students based on the recommendation of. Tay, & Jebb (2017) on the recommended one hundred to two hundred (100-200) sample sizes of the participants similar to the main group that was studied. The trial testing of the instrument for the study was conducted to test the adequacy of the testing instrument which was executed by calculating KR20 reliability with the use of statistical package for social sciences (SPSS) and the reliability coefficient realized was 0.86, and this proves that the instrument is reliable for utilization for this study.

Method of Data Collection

Before administering the instrument to the undergraduates in the selected universities, the researcher obtained permission from the head or management of the selected universities by presenting a letter of access that was given to the researcher to the head of department to the various heads. Thereafter the researcher and the assistance employed an accidental approach and administered copies of the instrument to available three hundred level undergraduates in the lecture hall as at the time of administration. After the students' responses, the copies of the instrument were reviewed, scored and collected ready for data analysis.

Method of Data Analysis

The statistical tools that were used to answer the research questions were analysis of variance via variance component. Research question one was answered using analysis of variance (ANOVA) by variance component method. Research question two was answered using relative and absolute errors. Research question three was answered using the dependability coefficient.

Table 1: Summary of analysis of variance showing the estimate of variance component and percentage of score variation for DCTS

Source of Variation	sum of squares	df	mean square	variance component	% variance component
Corrected Model	636.216	50	12.724		
Intercept	35466.442	1	35466.442		
ITEMS	298.302	49	6.088	.003	1.42
OCCASION	352.613	1	352.613	.007	3.30
Error	20969.536	104046	.202	.202	95.3
Total	73495.000	104097		0.212	
Corrected Total	21605.752	104096			

From the result in table 1, it can be seen that ANOVA table has a sum of squares for the item 298.302 with df 49, and the mean square was 6.088, accompanied with the variance component of .003 indicating 1.42% of the total variance for critical thinking scale for university students. Then for occasion, the sum of square is 352,613, df of 1, and a mean square of 352.613. This is associated with a variance component of 007, which accounted for 3.30% total variance. Finally, on the same table it is shown that for error, sum of square of 20969.536 at df of 104096 was obtained alongside a mean square of 0.202. This is associated with a variance component value of 0.202, which accounted for 95.3% of the total variance for DCTS.

Research Question 2: What is the relative error estimate and absolute error estimate?

Table 2: The research question 2 scores were answered using relative error estimate and absolute error, the result are revealed in table 2

Relative Error Variance	Universe Score	Absolute Error Variance
0.0114	0.007	0.0111

The table 2 reveals that the values estimated for relative error variance and absolute error variance were 0.0114 and 0.0111 accordingly, while the universe score showed a value of 0.007.

Research Question 3: What is the dependability of the critical thinking scale (CTS) items?

Table 3: The research question scores were answered with dependability coefficient, the result yielded are written in table 3

Universe Score	Absolute Error Variance	Index of dependability
0.007	0.0111	0.63

From table 3 the result shows that the value for index of dependability of generalizability theory for the instrument (CTS) was 0.63. It can also be seen that it had same value of 0.0111 for absolute error variance and the universe score was 0.007 respectively.

DISCUSSION OF FINDINGS

In consideration of the facets magnitude discussion, one of the functions of generalizability theory is to ascertain the reliability of the instrument if various sources of variations adds or contributes to measurement error. Generalizability theory eliminates different sources of error in a measurement, so the result realized from this study showed that multiple sources of contributed to the measurement error in the critical thinking scale CTS, the source of measurement error in this study are items and occasion.

The estimated value for the relative error and absolute error estimate of the CTS items had very minor coefficient meaning that so much error were not found in developing the items, the same consciousness was in place concerning the absolute errors which also had insignificant coefficient. Decision can be relative or absolute, relative decision is like a norm referenced test while absolute decision is in form of standard or degree of achievement; therefore whether the errors were to check norms in terms of critical thinking or the degree at which the students can think critically, the results has proven that the errors made were not significant or large. The estimation of standard error of measurement in generalizability and decision studies depends on the relative or total error from the measurement to create awareness for psychometricians concerning the significant interval of a person's score. The variance categories may be determined in the standard errors, if the standard errors are high it means that the variance categories are fluctuating when the participant size is not large. But if the total value is higher than the standard error, the variations of the categories may not be considered to be zero.

Tavakol and Brennan, (2013) Used critical thinking as a case study, that thinking can be the entity of measurement while the errors in the measurement may be as a result of unreliability of factors like inventory and those rating and some problem that emanate from the sample drawn from those taking inventory, rater and how they interacted as the inventory helps in calculating the variations of the component (Min, Shavelson, Yin and Wiley 2015)

While the dependability of the Dorathy's CTS findings of research question of the instrument developed has an acceptable index with a very minimal absolute error; this implies that the instrument is moderate for university lecturers, students, examining bodies to depend on. In that regard, the dependability of the instrument is adequate in terms of availability, reliability, maintainability, durability, as applied in this study. Based on that, G theory here has the ability of providing trust worthy findings within the time frame. the adequacy of this study as seen in the result has made it necessary to say that generalizing test takers observed score of construct of measurement got putting all conditions under consideration allowed by those who will used the test desirable.

The theory is an adequate approach of estimating measurement reliability from many factors (facet) in a modified pattern than the classical test theory of reliability and gives room to find out the dependability of measurement and to establish different sources of variations such as interaction effects (Heitman, Koocaleski and Pugh, 2009). Also, Preuss (2013) investigated clinical assessment protocols using generalizability theory to realize information that have reliability in terms of clinical achievement, he reviewed two benefits of the theory because it allows many measurement sources and generalization of findings of study areas to various facets as well as re-estimation of reliability. The study was directed on how generalizability can be adopted in the determination of data and construction of assessment protocols as it relates to patients, measures were taken on the parameters of gait, stance time- the period of limb contact with the ground and a support time when limb was alone in contact down.

RECOMMENDATIONS

1. The stake holders in the universities should imbibe evaluating students with the use of critical thinking because it is dependable.
2. G-theory should be used to determine the reliability of instrument in it will review multiple sources of variations.

REFERENCES

- Adeleke, J. O. (2010). The Basic of Research and Evaluation Tools.Somerest Ventures, Guinness Round About by NEPA Gate. Oba Akran Avenue End. Ogba-Ikeja.
- Akol, M., & Brennan, R. L. (2013) Michael Education Assessment: a brief over view of concept in generalizability theory and Structural Equation Modeling. *Psychological Methods*, 23(1), 1-26.
- Briesch, A. M., Swaminathan, H., Welsh, H., & Chafoleas, S. M. (2013). Generalizability Theory: A practical guide to study design, implementation, and interpretation. *Journal of School Psychology*, 52(1).
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). The dependability of behavioral measurements: Theory of generalizability for scores and profiles. Retrieved August 4th, 2021.
- Heitman, R. J., Kovaleski, J. E., & Pugh, S (2009). Application of generalizability theory in estimating the Reliability of the Ankle Complex Laxity Measurement. *Journal of Athletic Training*, 44(1).
- Lakes, K. D., & Hoyt, W. (2009). Application of Generalizability Theory to Clinical Child and Adolescent Psychology, *J Clin Child Adolesc Psychol*, 38(1).
- Lenin, I. (2019). Critical Thinking and its Importance in Education. https.//www.researchgate.net. Retrieved July 30th, 2021.
- Michelle, A. C. (2015). Importance of Critical Thinking. https.//www.researchgate.net. Accessed July 30th, 2021.
- Min, L., Shavelson. R, Yin, Y., & Wiley, E. (2015). Generalizability Theory. Encyclopedia of Clinical Psychology accessed 12th September, 2018.
- Murat .K (2016) The Significance of Critical Thinking Ability in Terms of Education. International Journal of Humanities and Social Sciences, Vol. 6, NO.7
- Preuss, R. A. (2013). Using generalizabily theory to Develop clinical assessment protocols. *Psychological Therapy*, 93(4).
- Tay., & Jebb, A. (2017). Scale Development in S. Regeberg (Ed), the SAGE Encyclopedia of Industrial and Organizational Psychology 2nd edition, thousand oaks CA sage Available at: https://www.researchgate.net/publication/305462834
- Thomas, G. (2017). Importance of Critical Thinking, Ability in Students. Retrieved 31st July 2018. https://www.careerindia.com/tips/importance-critical-thinking-ability-students