Homo-Hetero Pairing Effect Correlation Coefficient:

A Modified Correlation Technique for Social Science Studies

Ntogwa Ng'habi Bundala

Tanzania Police Force, Mara Police Regional,
Criminal Intelligence Bureau
Mob: +255(0)73360418/+255(0)782 33 87 70

E-mail:bundalantogwa@gmail.com

Abstract

The study aimed to examine the pairing relationship between economic growth and psychological human behaviour (psychological well-being) of individuals by application of the homo-hetero pairing effect correlation coefficient technique. The cross-sectional data were used. The data were collected from 211 individuals randomly sampled from two regions in Tanzania. The data were analysed by the homo-hetero pairing effect correlation coefficient and the results were compared to that of the Pearson correlation coefficient and simple regression model. The study found that economic growth and psychological well-being are positively changing in pairs, and the psychological well-being of the individual is the true enabler (optimal independent) of the linear modelling. The study concluded that the improvement of the psychological well-being of individuals significantly improves economic growth and not vice versa. Therefore, the paper recommended that psychological well-being-based initiatives should be established and encouraged in society as found to have a positive impact on economic growth. Moreover, the study specifically, recommends the application of the homo-hetero pairing correlation coefficient in studies of cardiology, neurology, epidemiology, psychology, economics, anthropology, sociology and other fields of the social sciences.

How to Cite:

Bundala, N. N. (2022). Homo-Hetero Pairing Effect Correlation Coefficient: A Modified Correlation Technique for Social Science Studies. *International Journal of Business, Management, and Economics, 3*(3). 213 - 256. DOI: https://doi.org/10.47747/ijbme.v3i3.706

1. Introduction and Literature Review

Social science is the system or procedural rule that scientifically defines society; this is why it is termed a social science. Social science differs from natural science. However, there is no field of study that is more important to human beings than the social sciences (Samuel and Okey, 2015; Maravelakis, 2019; Hernandez and Alpizar-Jara, 2018). The study of social science is more

helpful for understanding what factors or conditions for a better life and what social opportunities are available for human development. Social science studies involve a lot about societies' behaviours and practices, mostly including anthropology (study of the relationship between biological traits and socially acquired characteristics), sociology (study of the relationship among people), economics (study of ways individuals satisfy their unlimited wants and needs-demands relative to limited resources), political science (study of social arrangement to maintain peace and order in society), and psychology(deals with the mind and personality of individual), and other. One of the most debatable issues in social science studies is the relationship between the economic growth and psychological well-being (psychological human behaviour) of the individual that resulted from the Easterlin paradox effect (Easterlin, 1973; 2017). The studies such as Diener and Seligman (2004), Roka (2020), Stevenson and Wolfer (2008; 2013), and Talhelm, Zhang, Oishi, ..., and Kitayama (2014) evidence the positive impact (correlation) between economic growth and psychological well-being, which contradicts Easterlin's (1973; 2017). On the other hand, Stoop, Leibbrandt and Zizzamia (2019) found a negative relationship between economic growth and psychological well-being. In addition, the studies such as Diener and Seligman (2004) and Baro and Sala-i-Martin (2004) suggested that the socio-economic and political measures of GDP have seriously failed to provide a full account of policy decisions at the organizational, corporate, and governmental levels. They concluded that the exclusion of variables such as psychological factors (e.g., personal traits, values, etc.) is not accounted for in the economic principle of demand and supply.

Moreover, the exclusion of the non-economic variables such as motivation, metacognition, and lifestyle leads to economic performance failure in some regions or countries (Diener and Seligman, 2004, Bundala, 2021). One of the potential problems of measuring the relationship between economic growth and psychological well-being is based on the *methodological incompleteness* of the existing correlation technique, particularly the Pearson correlation coefficient. Samuel and Okey (2015) emphasised that Pearson correlation is a *sine qua non* in social science studies; it is an excellent analytical tool when implement correctly and used with its partner regression. Since correlation remains a very important tool in social science study and its techniques are almost inevitable, especially in quantitative studies that involve variables (Samuel and Okey, 2015; Senthilnathan, 2019).

Despite the massive support of the Pearson correlation technique in social science studies, some empirical studies negate its advantages. Some studies condemn the correlation coefficient for its inability to detect the cause-and-effect (Bertoldo, Callegher and Altoe, 2022; Janse, Hoekstra, Jager... and van Diepen, 2021; Coppack, 1990; Grogtay and Thatte, 2017; Rahman and Zhanga, 2016). The coefficient correlation is developed in early 1896 by Karl Pearson, to determine the *strength* which is measured by magnitude ranging from -1 to +1 and the *direction* of the magnitude which is indicated by the negative or positive signs (Janse, et al, 2021; Coppack, 1990; Gogtay and Thatte, 2017; Danacica and Babucea, 2007; Kumar and Chong, 2018). That is, the correlation coefficient is aimed to measure how the two observed variables *x* and *y* are moving together or changing together,i.e, co-vary (Akoglu, 2018; Mukaka, 2012; Emerson, 2015; Bertoldo, Callegher and Altoe, 2022). For example, a cancer disease may be associated or go together with the smoking behaviour of an individual (smoking frequencies), or the Human

immunodeficiency virus (HIV) can be associated with Tuberculosis (TB) disease, but you cannot conclude or say that the TB is caused by HIV, or HIV is caused by TB. But one of the factors or variables can be "an enabler" of another (enabled); not dependent and independent factors or variables. However, sometimes the relationship between enabler and enabled variables may be similar to that of independent and dependent variables, depending on the nature of the data.

The enabler factor or variable is the factor that increases the opportunity for another factor (enabled) to happen. In simple language, enabler factors are inviting agents but may not be loosely defined as cause-and-effect factors. For example, in the correlation between TB and HIV, which one can be an enabler and another is enabled? The question is how to determine the enabler and enabled factors. It is still impossible to determine the enabler and enabled factors/variables by using the available (current) correlation techniques. Substantially, this is the major methodological flaw of the available correlation coefficient techniques. Understanding the strength in terms of the figure (magnitude) and direction in terms of signs is not enough to make a clear decision relating to the paired variables. A lot of methodological questions are left unanswered by the Pearson correlation coefficient, *r*. For instance, if you say *x* and *y* have r = 0.60, it means that they are moving in the same way in the strength of 0.60 (medium), that is, an increase of *x* results in an increase of *y* and it is vice versa for decreasing. Moreover, if *r* = -0.60, it indicates that the variables *x* and *y* are moving in the opposite direction. That is the increase of *x* leads to a decrease of *y*. But, how "they are accurately moving in pairs, i.e., changing up (increases) and down (decreases) together in either direction? What is the pairing effect (error) of the observed variables *x* and *y* as they are moving in pairs (co-related) downwards or upwards? What is an enabler and enabled variables or factors in the linear modelling?

In most practice, when correlation techniques are used in social science, the identification of the enabler and enabled factors/variables was done by using experience or additional knowledge or skills of researchers or decision-makers. For example, if it is found that there is a high correlation between cancer and smoking behaviour, an expert for medical issues (medical officer or practitioner) can identify the *enabler and enabled factors* by using his/her medical skills or experience on the related variables or factors. However, these methods of identification of the enabler and enabled factors have many flaws because they are based on individual expertise, skills, or experience which varies from individual to individual, so the contradiction in interpretation may happen which can cause a dilemma or contra-conclusion problem. Although, the Pearson correlation coefficient gets massive applicability support in social sciences studies (Gokul, Srinivasan, and Swaminathan, 2021; Samuel and Okey, 2015; Senthilnathan, 2019; Onwuegbuzie and Daniel, 1999; Rahman and Zhang, 2016; Li, Chu, Li,..and Jiang, 2022; Sverko, Vrankic, Vlahinic and Rogelj, 2022; Chinnadurai and Bobin, 2021); it requires a complementary technique to answer some of the left methodological questions (Onwuegbuzie and Daniel, 1999; Senthilnathan, 2019; Benesty, Chen and Huang, 2008; Sedgwick, 2012). For this reason, this paper developed a homo-hetero pairing effect correlation coefficient that explained the left methodological flaws by the Pearson correlation coefficient technique. This modern correlation technique can be used in parallel to the Pearson correlation technique. It would be easier and more helpful for decision-makers to determine which variable is enabled and which an enabler is. The identification of enabler and enabled variables is very important and

needy in decision making particularly in social science studies. Therefore, the study aimed to examine the pairing relationship between the economic growth and psychological human behaviour (psychological well-being) of individuals by application of the homo-hetero pairing effect correlation coefficient technique to fill both contradictory evidence and methodological gaps resulting from the studies of Easterlin (1973; 2017), Roka (2020), Stevenson and Wolfers(2008; 2013) and others. Specifically, the study determines the enabler and enabled variables and optimal independent (true enabler) variables in the linear modelling of the economic growth and psychological well-being of the individual. Moreover, the study examines the empirical relevance of the linear modelling (regression model) and methodological coincidence between the homo-hetero pairing correlation coefficient and Pearson correlation coefficient techniques with respect to the linear regression model. The next sections of the paper cover the analytic framework and idea generation, methodology, findings, discussion, and the last sections cover the conclusion and recommendation, and references.

2. Literature Review and Model Development
2.1. Analytic Framework and Idea Generation

Given the pair of the observed variable *x* and *y* in an array of descending order, ask how the relationship of *x* and *y* can be determined? Obviously can be determined in the two directions of changes, that is, hetero pairing change (the change of difference of paired variable), that is, $(x_1 - y_1)$ to $(x_2 - y_2)$. This change determined the "pairing effect or error" of the observed variables *x* and *y*. Moreover, the homo pairing change can be used to determine the relationship of the observed variables. The homo pairing change is the change of difference of the same variable, either *x* or *y*, that is, the change of *x*, $(x_1 - x_2)$ if is moving together with the change of *y*, $(y_1 - y_2)$. The traditional correlation analysis such as Pearson correlation used this *homo-paring change* or effect to determine the correlation of the observed variable *x* and *y*, that is, how the change of *x* $(x_1 - x_2)$) is co-vary with the change of *y*, $(y_1 - y_2)$. Therefore, the covariance of *x* and *y* is weighted to their standard deviation (variability error). In practice this method does detect or predict the "causality effect" of the paired variables *x* and *y*, instead, it shows the positional pairing effect (error) of homo variables (same variables). That is if there is a co-variation of the change of *x* from x_1 to x_2 and the change of *y* from y_1 to y_2. That is, both *x* and *y* are changing or moving in approximately equal "distance" (values) and in the same or opposite direction. This is why the scale of Pearson correlation takes both positive and negative values in the range of -1 to +1.

To measure or examine the correlation of the two observed variables by using only "homo pairing effects have been evidenced to lack its methodological completeness. Some of the relational attributes are left untouched! Arranging the data in the paired data array the Pearson correlation analysis only considers the vertical co-movement or variation of the data (increasing or decreasing upward and downward). That is, the observed variables *x* and *y* change x_1 to x_2 and y_1 to y_2 respectively. Unfortunately, the horizontal co-movement or variation of the paired data, that $x_1 \leftrightarrow y_1$ to $x_2 \leftrightarrow y_2$ was not considered. Overlooking this correlational dimension in which the observed variable *x* and *y* move horizontally (leftwards or rightwards) to the paired data array leads to the methodological incompleteness of the Pearson correlation analysis.

Consequently, this study introduces a new way of measuring the correlation which considers all the change effects of pairing, that are, homo and hetero pairing effects of the paired data *x* and *y*. That is, the homo-hetero pairing effect correlation coefficient, or simply "pairing effect correlation coefficient. This method is very sensitive to data order or arrangement and powerful technique to determine the pairing correlation direction and strength of the paired variables in terms of the *homo-hetero pairing effect*. The method identifies the enabler and enabled factors quantitatively.

To be clear, the Pearson correlation does not consider the "homo-hetero pairing effect" of the paired data array; therefore, its applicability is still limited in social science studies /decisions (Taylor, 1990; Asuero, Sayago and Gonzalex, 2006; Shelef and Schechtman, 2018; Coppack, 1990). Cautionary, the concept of enabler and enabled factors or variables cannot be conceptually considered or confused with dependent and independent variables or factors; they are two different terms and concepts. The enabler factor or variable is the factor that increases or decreases the occurrence opportunity of enabled factors or variables. In other words, the enabler is the factor that increases the "occurrence probability" of the enabled factor/variable; it is the "host" factor or variable. For example, rainfall is an enabler for a farmer to farm. Now, the relationship between rainfall and farmer, intuitively, can be considered positive as when the rainfall falls on the land the farmer would go to the farm. That is, the farmer does not farm because of the rainfall consequences, but he/she is enabled by the falls of the rainfall (he/she would be waiting for rainfall) to fulfil his/her course of action. Another example is a thief that is waiting for the housekeeper to leave out the house and snap/steal a piece of valuable metal. The presence or absence probabilities of the housekeeper at the house determine the will of a thief to steal or not. Therefore, the frequency of leaving out the house of a housekeeper (absence) will be positive related or correlated to the number of theft attempts at the house. In that case, the presence or absence probability of a housekeeper is the enabler factor of the thief or stealing attempts, as he was waiting for that opportunity. Stealing is an enabled event or variable.

On the other hand, the dependent and independent variables or factors are terms that are usually used in regression analysis to determine the causality effect of the independent or explanatory variables or factors on dependent variables. In this concept, the dependent variables or factors are the factors that vary as the independent variables vary and not vice versa. In other words, the dependent variables are the variables or factors that depend on the change or influence of independent variables. Therefore, the independent variables are variables or factors that are not dependent on others, instead are "causal factors" of the dependent variables or factors. Take an example, the two variables advertisement efforts or costs and sales. We expect that once the advertisements efforts are done intensively (increase) would increase the marketing awareness and hence convince or persuade more customers to buy the products or services advertised, that is, more buyers, increase sales. But do the increase or decreases in sales affect or influence the advertisement efforts or costs? In other words, do the increases in advertisement efforts or costs due to the increases in sales? The answer is no! The advertisement efforts or costs are independent of sales; do not depend on sales but sales depend on advertisement efforts or costs. Therefore, the sale is a dependent variable (it depends on the advertisement efforts or costs) and the advertisement efforts or costs are independent variables because do not depend on the sales.

International Journal of Business, Management, and Economics

Vol. 3 No. 3, August 2022

e-ISSN 2746-1351

We can notice the difference between the enabler and the enabled factors by the stated examples. Recall the example of a thief and a housekeeper, the one can say that the probability of presence or absence of the housekeeper in the house and the number or frequency of theft attempts do not qualify to carry the degree or attributes of dependent and dependent variables. The frequency of theft attempts is not due to or not depends on the probability of the presence or absence of the housekeeper in the house or the probability of the presence or absence of the housekeeper at the house is not due to the frequency of theft attempts by theft, but they are associated in a pairing effect order; one can lead another-a paired predecessor and successor factors (enabler and enabled paired variables). For a clear illustration of the pairing effect (error) correlation coefficient, consider the paired observed data *x* and *y* (Table 1). The data are assumed to be linearly related and normally distributed.

Table 1: The Data array of the paired observed variables *x a*nd *y*

x	y
x_1	y_1
x_2	y_2
x_3	y_3
.	.
.	.
.	.
x_n	y_n

Source: Author (2022).

Table 1 describes that the variables *x* and *y* can decrease or increase vertically (upward or downwards). The change of *x* from x_1 to x_2 is paired to the change of *y* from y_1 to y_2 and the change difference is co-related between them. Moreover, the change of *x* from x_2 to x_3 is paired or associated with the change of *y* from y_2 to y_3. In general, for each change of *x* from x_i to x_j, there is a corresponding or associated change of *y* from y_i to y_j which correlated with it. We measure the change difference as the proportion of change of *x*, and *y*, that is *x*-homo pairing effect (∂_x) and *y*-homo pairing effect (∂_y), and their *x* and *y* mean homo pairing effect, $\bar{\partial}_x$ and $\bar{\partial}_y$ respectively.

$$\partial_x = \frac{(x_i - x_j) - (y_i - y_j)}{(x_i - x_j)}$$

$$\bar{\partial}_x = \frac{1}{n-1} \sum_{i=1}^{n} \frac{(x_i - x_j) - (y_i - y_j)}{(x_i - x_j)}$$

$$\partial_y = \frac{(y_i - y_j) - (x_i - x_j)}{(y_i - y_j)}$$

$$\bar{\partial}_y = \frac{1}{n-1} \sum_{i=1}^{n} \frac{(y_i - y_j) - (x_i - x_j)}{(y_i - y_j)}$$

Now, the methodological question is whether these pairing effects are moving together, i.e., correlated. The relationship between parameters ∂_x and ∂_y will indicate the pairing effect correlation of the observed variables x and y. Therefore, to get the correlation between the two parameters, we use the common technique of standardised variance or the ratio of signals and their noises. That is, the covariance of ∂_x and ∂_y, is given by

$$Cov\left(\partial_x, \partial_y\right) = \frac{\Sigma(\partial_{xi} - \bar{\partial}_x)(\partial_{yi} - \bar{\partial}_y)}{n-1}$$

Now, the Homo-hetero pairing effect correlation coefficient, Ψ_b is given by,

$$\Psi_b = \frac{Cov\left(\partial_x, \partial_y\right)}{\sigma_{\partial_x}\sigma_{\partial_y}}$$

Since, $\quad \sigma_{\partial_x} = \sqrt{\sum \frac{(\partial_{xi} - \bar{\partial}_x)^2}{n-1}}, \; \sigma_{\partial_y} = \sqrt{\sum \frac{(\partial_{yi} - \bar{\partial}_y)^2}{n-1}}$

Therefore,

$$\Psi_b = \frac{\dfrac{\Sigma(\partial_{xi} - \bar{\partial}_x)(\partial_{yi} - \bar{\partial}_y)}{n-1}}{\left(\sqrt{\Sigma \frac{(\partial_{xi} - \bar{\partial}_x)^2}{n-1}}\right)\left(\sqrt{\Sigma \frac{(\partial_{yi} - \bar{\partial}_y)^2}{n-1}}\right)}$$

$$\Psi_b = \frac{\left(\dfrac{\Sigma(\partial_{xi} - \bar{\partial}_x)(\partial_{yi} - \bar{\partial}_y)}{n-1}\right)\sqrt{(n-1)^2}}{\left(\sqrt{\Sigma(\partial_{xi} - \bar{\partial}_x)^2 \Sigma(\partial_{yi} - \bar{\partial}_y)^2}\right)}$$

$$\Psi_b = \frac{\Sigma(\partial_{xi} - \bar{\partial}_x)(\partial_{yi} - \bar{\partial}_y)}{\sqrt{\Sigma(\partial_{xi} - \bar{\partial}_x)^2 \Sigma(\partial_{yi} - \bar{\partial}_y)^2}}$$

The formula looks like the Pearson correlation coefficient but is quite different, this is the correlation of parameters ∂_{xi} and ∂_{yi} which built on the x-and y-homo change of the observed variables x and y. The substituting the x and y in the *parametric equation* the formula will be

large and more complex to understand. Therefore, this parametric equation is more simplified and easy to use.

2.2. Meaning and Interpretation of Ψ_b

The homo-hetero pairing effect correlation coefficient or simply Bundala pairing effect (error) correlation coefficient is denoted by the Greek letter Psi, (Ψ_b). The subscript small b indicates the name of the author, Bundala. For simplicity, Ψ_b would be termed as Bundala's Psi coefficient or Bundala pairing effect correlation coefficient. This correlation analysis measures the "homo-hetero pairing relationship" of the observed variables x and y. The homo-hetero pairing relationship can be defined symbolically, $(x_i \leftrightarrow x_j) \leftrightarrow (y_i \leftrightarrow y_j)$. Where, $(x_i \leftrightarrow x_j)$ is the x-homo pair and $(x_i - x_j)$ is the x-homo pair effect (error), and $(y_i \leftrightarrow y_j)$ is y-homo pair and $(y_i - y_j)$ is the y-homo pairing effect. Therefore, $(x_i \leftrightarrow x_j) \leftrightarrow (y_i \leftrightarrow y_j)$ is the homo-hetero, i.e., different homo-pairs (x, y). Technically, (Ψ_b) is based on the standardised covariance technique of the x –homo pairing effect and y-homo paring effect.

Logically, the (Ψ_b) it measures how the difference of change (effect) of one variable is correlated with the change (effect) of the other variables. That means it measures how accurately the "correlated variables are paired". The higher Ψ_b indicates the pairing closeness or shortness of pairing distance which is a separation distance (effect) of the paired data. The lower Ψ_b indicates that the two variables are poorly paired. In practices, there is no benchmark value of Ψ_b; it depends on the nature of the studies. In some studies the lower value of Ψ_b is desirable but in some studies is not desirable, and it is vice versa.

2.2.1 Meaning and Interpretation of Ψ_b, when $\Psi_b = 0$ and $r = 1$

Notably, if (Ψ_b) is interpreted along with the Pearson correlation coefficient, r, the r with a higher (Ψ_b) will be considered the optimal or correct one. The values of Ψ_b range from -1 to 1, that is, $-1 \leq \Psi_b \leq 1$. Notably, if the observed variable x and y are perfectly correlated, that is, $r = 1$, then $\Psi_b = 0$. In other words, Ψ_b does not depend on r. The value of Ψ_b can be 0 (perfectly unpaired) while $r = 1$(perfectly paired). Considers a hypothetical example with data of paired variables x and y and parameters of Ψ_b in the hereunder table,

x	y	∂_x	a	∂_y	b	$a.b$	a^2	b^2
1	6	-5	0	0.83	0	0	0	0
2	12	-5	0	0.83	0	0	0	0
3	18	-5	0	0.83	0	0	0	0
4	24	-5	0	0.83	0	0	0	0
5	30	-5	0	0.83	0	0	0	0
6	36	-5	0	0.83	0	0	0	0
7	42	-5	0	0.83	0	0	0	0
8	48	-5	0	0.83	0	0	0	0
9	54	-5	0	0.83	0	0	0	0
10	60							

From the table,

$$\bar{\partial}_x = \frac{1}{n-1} \sum_{i=1}^{n} \frac{(x_i - x_j) - (y_i - y_j)}{(x_i - x_j)} = -5$$

$$\bar{\partial}_y = \frac{1}{n-1} \sum_{i=1}^{n} \frac{(y_i - y_j) - (x_i - x_j)}{(y_i - y_j)} = 0.83$$

$$\sum (\partial_{xi} - \bar{\partial}_x)^2 = \sum a^2 = 0 \text{ and } \sum (\partial_{yi} - \bar{\partial}_y)^2 = \sum b^2 = 0$$

$$\sum (\partial_{xi} - \bar{\partial}_x)(\partial_{yi} - \bar{\partial}_y) = \sum a.b = 0$$

$$\Psi_b = \frac{\Sigma(\partial_{xi} - \bar{\partial}_x)(\partial_{yi} - \bar{\partial}_y)}{\sqrt{\Sigma(\partial_{xi} - \bar{\partial}_x)^2 \Sigma(\partial_{yi} - \bar{\partial}_y)^2}} = \frac{0}{\sqrt{0x0}} = 0$$

The $\Psi_b = 0$, means that the data of the variable x and y are perfectly unpaired but are perfectly correlated (co-vary with zero homo-hetero pairing effects). That is, the x and y-homo pairing effects are not paired, hence not correlated, $\Psi_b = 0$. This can be checked by using x and y-homo pairing effect trends (Figure 1).
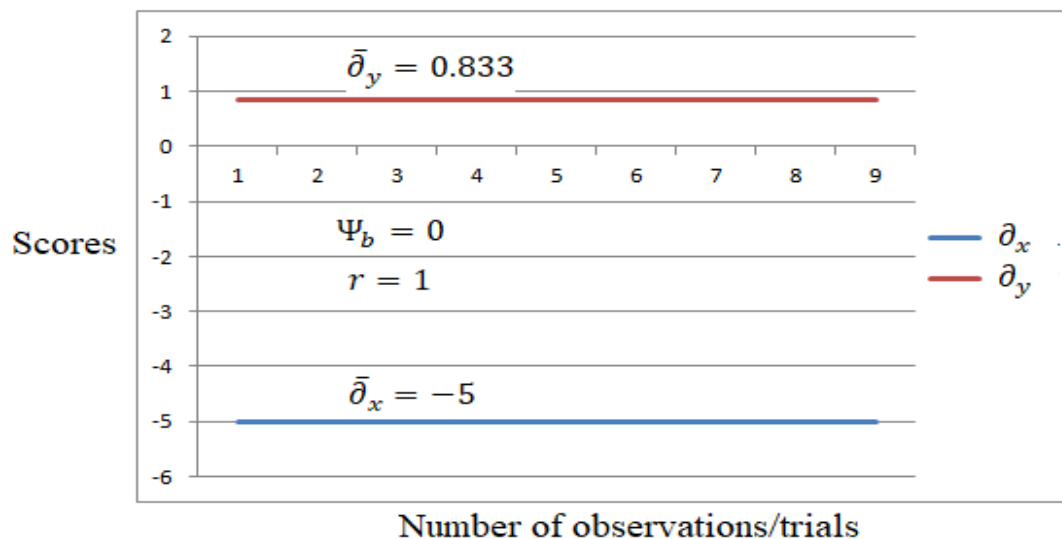
Source: Author (2022)

Figure 1: The x and y- homo pairing effects trends in Ψ_b analysis

Figure 1 shows x and y-homo pairing effects trends. Both x and y homo pairing effects are constant, hence not correlated. That is, the mean x-homo pairing effect(error) is -5 which

indicates the *x* always changing (decreasing) with a constant ratio of 5, and the mean *y*-homo pairing effect(error) is 0.833 which indicates that the values of *y* are always or constantly changing (increasing) with a ratio of 0.833. Since all means of *x* and *y*-homo pairing effects (errors) are constant, there is no overlapping or down and upward co-movement of the observed variables *x* and *y*; therefore, the straight lines are exhibited. That is, why the value of Ψ_b is equal to zero, meaning that there is no homo-hetero pairing effect because the *x* and *y*-homo pairing errors (standard deviation) are zero.

On the other hand, the Pearson correlation coefficient was calculated to re-check the concept of independence of Ψ_b on r.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \ \sum(y_i - \bar{y})^2}}$$

x	y	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
1	6	-4.5	-27	20.25	729	121.5
2	12	-3.5	-21	12.25	441	73.5
3	18	-2.5	-15	6.25	225	37.5
4	24	-1.5	-9	2.25	81	13.5
5	30	-0.5	-3	0.25	9	1.5
6	36	0.5	3	0.25	9	1.5
7	42	1.5	9	2.25	81	13.5
8	48	2.5	15	6.25	225	37.5
9	54	3.5	21	12.25	441	73.5
10	60	4.5	27	20.25	729	121.5
Sum	330	0	0	82.5	2970	495

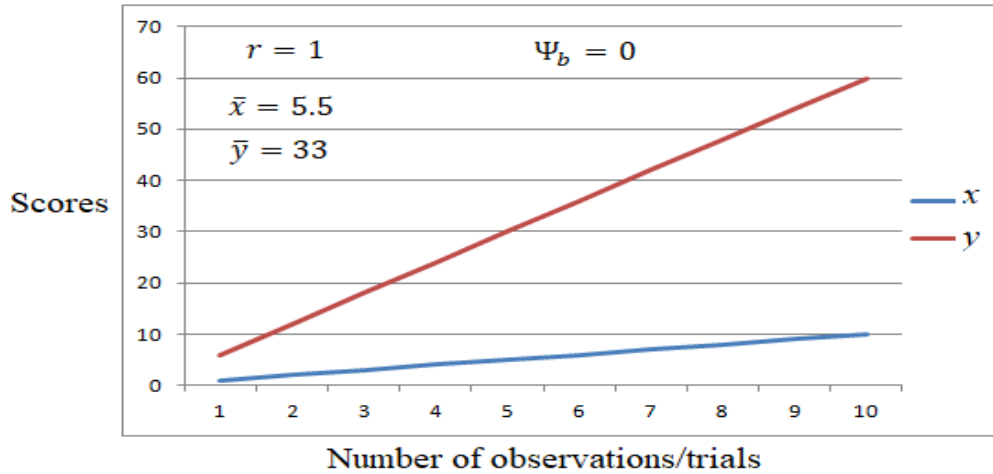Now from the table, the Pearson correlation coefficient can be calculated by using the formula,

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \ \sum(y_i - \bar{y})^2}}$$

$$(x_i - \bar{x})(y_i - \bar{y}) = 2310$$

$$\sum(x_i - \bar{x})^2 = 82.5, \quad \text{and} \ \sum(y_i - \bar{y})^2 = 2970$$

$$r = \frac{495}{\sqrt{82.5 \ x \ 2970}} = 1$$

Therefore, we get $r = 1$ while $\Psi_b = 0$, this means that the data of the variables *x* and *y* are perfectly correlated, $r = 1$ (both are trending, i.e., moving positively upward) with no homo-hetero pairing effects, $\Psi_b = 0$ (Figure 2).

Source: Author (2022)

Figure 2: Data directional trending in the Pearson correlation analysis

Figure 2 shows the graphical presentation of data from the observed variables *x* and *y* as described by the Pearson correlation analysis. The graph shows that the data are perfectly correlated as both are moving in the same direction, with a mean value of *x*, $\bar{x} = 5.5$ and mean value of *y*, $\bar{y} = 33$. The paired data *x* and *y* are perfectly fitted or matched about their respective means values, $\bar{x}$ and $\bar{y}$. Therefore, we conclude that where there is the perfect correlation of paired data *x* and *y*, $r = 1$, then $\Psi_b = 0$.

2.2.2 Meaning and Interpretation of Ψ_b, when $\Psi_b \neq 0$ and $r \neq 1$

To get a clear and thorough meaning and interpretation of the concept of Ψ_b and r, considers a more hypothetical example with empirical paired data that explains the conditions of $\Psi_b \neq 0$ and $r \neq 1$. The data and Ψ_b estimates parameters were provided hereunder in the table,

x	y	∂_x	a	∂_y	b	a.b	a^2	b^2
3	6	-0.50	-0.33667	0.33	0.279167	-0.09399	0.113347	0.077934
5	9	-1.00	-0.83667	0.5	0.449167	-0.3758	0.700017	0.201751
4	7	-0.5	-0.33667	0.33	0.279167	-0.09399	0.113347	0.077934
6	10	0.25	0.41333	-0.33	-0.38083	-0.15741	0.170842	0.145034
10	13	0.22	0.38333	-0.29	-0.34083	-0.13065	0.146942	0.116167
19	20	-0.09	0.07333	0.08	0.029167	0.002139	0.005377	0.000851
53	57	0.09	0.25333	-0.11	-0.16083	-0.04074	0.064176	0.025867
32	38	-0.37	-0.20667	0.27	0.219167	-0.0453	0.042712	0.048034
62	79	0.17	0.33333	-0.2	-0.25083	-0.08361	0.111109	0.062917
86	99	0.15	0.31333	-0.17	-0.22083	-0.06919	0.098176	0.048767
11	35	0.11	0.27333	-0.13	-0.18083	-0.04943	0.074709	0.032701
20	43	-0.49	-0.32667	0.33	0.279167	-0.0912	0.106713	0.077934
57	98	Sum				-1.22917	1.747467	0.915892

From the table the Ψ_b can be calculated by using the formula,

$$\Psi_b = \frac{\Sigma(\partial_{xi} - \bar{\partial}_x)(\partial_{yi} - \bar{\partial}_y)}{\sqrt{\Sigma(\partial_{xi} - \bar{\partial}_x)^2 \Sigma(\partial_{yi} - \bar{\partial}_y)^2}}$$

$$\sum(\partial_{xi} - \bar{\partial}_x)^2 = \sum a^2 = 1.747467 \text{ and } \sum(\partial_{yi} - \bar{\partial}_y)^2 = \sum b^2 = 0.915892$$

$$\sum(\partial_{xi} - \bar{\partial}_x)(\partial_{yi} - \bar{\partial}_y) = \sum a.b = -1.22917$$

$$\Psi_b = \frac{-1.22917}{\sqrt{1.747467 \; x \; 0.915892}} = -0.9715$$

Now, we get $\Psi_b = -0.9715$ indicating that the changes (increase and decrease) of *x* and *y* are pairing oppositely. That is, the homo-hetero pairing effect is negative (when a change of *x* is up, the change of *y* is down). This movement cannot be shown or detected in this Pearson correlation analysis; however, it is evidenced when the *x* and *y*-homo pairing effect trending in Bundala pairing correlation analysis is established (Figure 3).
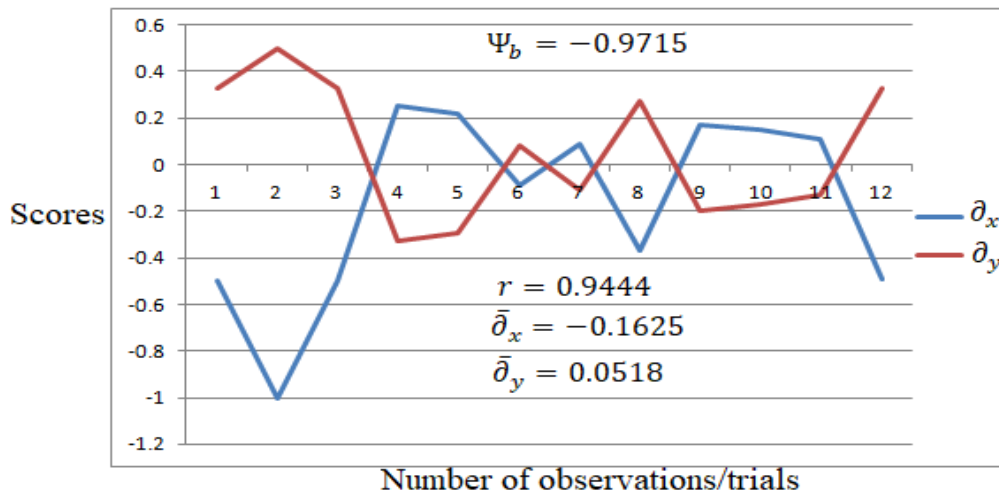
Source: Author (2022)

Figure 3: The *x* and *y*-homo pairing effects trending in Bundala correlation analysis

Figure 3 shows the data trending of the *x* and *y*-homo pairing effects of variables *x* and *y* respectively. It detects the homo effects of the variables *x* and *y* as they move up (increasing and decreasing). The graph shows how the homo pairing effects/changes of *x* and *y* is going up and down as they move together upwards. The graph depicts the mean *x*- homo pairing effects, $\bar{\partial}_x$ is

-0.1625. This means that x changes are decreasing at the ratio or constant values of 0.1625. On the other hand, the mean y- homo pairing effects of y, $\bar{\partial}_y$ is 0.0518. This means, that values of y are increasing positively by a constant ratio $\bar{\partial}_y = 0.0518$. Because, $\bar{\partial}_x \neq \bar{\partial}_y$ then there are unequal up and down movements of x and y-homo pairing changes/effects of variables x and y.

On the other hand, we use the same data we computed the Pearson correlation coefficient, r by using the formula, the condition is $r \neq 1$.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \ \sum(y_i - \bar{y})^2}}$$

The hypothetical empirical data and estimates parameters of r are presented in the table hereunder,

x	y	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
3	6	-27.42	-33.54	751.8564	1124.932	919.6668
5	9	-25.42	-30.54	646.1764	932.6916	776.3268
4	7	-26.42	-32.54	698.0164	1058.852	859.7068
6	10	-24.42	-29.54	596.3364	872.6116	721.3668
10	13	-20.42	-26.54	416.9764	704.3716	541.9468
19	20	-11.42	-19.54	130.4164	381.8116	223.1468
53	57	22.58	17.46	509.8564	304.8516	394.2468
32	38	1.58	-1.54	2.4964	2.3716	-2.4332
62	79	31.58	39.46	997.2964	1557.092	1246.147
86	99	55.58	59.46	3089.136	3535.492	3304.787
11	35	-19.42	-4.54	377.1364	20.6116	88.1668
20	43	-10.42	3.46	108.5764	11.9716	-36.0532
57	98	26.58	58.46	706.4964	3417.572	1553.867
Sum	-	-	-	9030.773	13925.23	10590.89

Now from the table, compute the Pearson correlation coefficient by using the formula,

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \ \sum(y_i - \bar{y})^2}}$$

$$(x_i - \bar{x})(y_i - \bar{y}) = 10590.89$$

$$\sum(x_i - \bar{x})^2 = 9030.773, \quad \text{and} \ \sum(y_i - \bar{y})^2 = 13925.23$$

$$r = \frac{10590.89}{\sqrt{9030.773 \ x \ 13925.23}} = 0.9444$$

From these calculations, $r = 0.9444$ indicates that the paired data are highly correlated in positive ways, and that, the trending changes are done in the same direction (either decreasing or increasing are occurring in the same directions). In other words, the paired data are increasing simultaneously or decreasing simultaneously. But how many degrees of change differences (decreases and increases) of the paired data are co-vary? This question is not answered in r but it answered in Ψ_b. In other words, the question can be asked how the x and y perfectly fit relative to their means values. The paired data trending graph is the best way of interpreting the data (trending) behaviour in r, which indicates either the data x and y are increasing simultaneous or decreasing simultaneously (Figure 4).
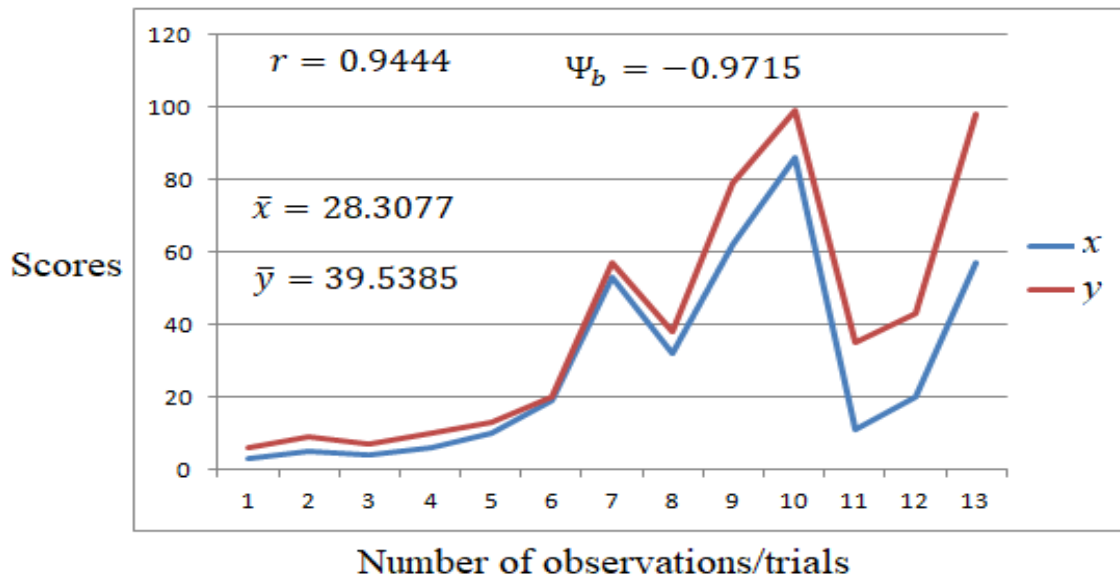
Source: Author (2022)
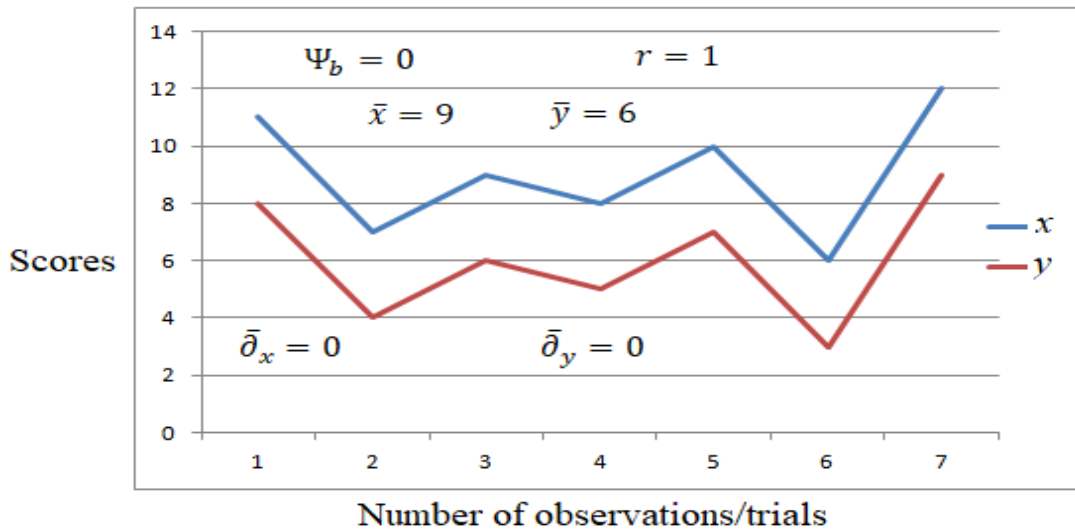Figure 4: The data trending in the Pearson correlation analysis

Figure 4 shows how x and y are trending positively. The data are trending up and down in a positive way, hence a positive $r = 0.9444$. That is, the data are increasing and decreasing simultaneously. The data changes are associated with unequal paired increments which are explained in Ψ_b. The graph shows the data moving in the same direction, with a mean value of x, $\bar{x} = 28.3077$ and mean value of y, $\bar{y} = 39.5385$. The paired data x and y are highly fitted to their respective means values, $\bar{x}$ and $\bar{y}$.

2.2.3 Meaning and Interpretation of Ψ_b, when $\bar{\partial}_x = \bar{\partial}_y = 0$

When $\bar{\partial}_x = \bar{\partial}_y = 0$, we get a single straight line separator which combines both x and y-homo pairing effects of variables x and y. When the both $\bar{\partial}_x$ and $\bar{\partial}_y$ are equal to zero, then, $\Psi_b = 0$ and $r = 1$ (Figure 6). Consider the hypothetical empirical data and the x and y-homo pairing effects of the variable x and y in the table below,

x	y	∂_x	∂_y
11	8	0	0
7	4	0	0
9	6	0	0
8	5	0	0
10	7	0	0
6	3	0	0
12	9		

Computing the Ψ_b and r from the above data we get, $\Psi_b = 0$ and $r = 1$. Since the $\partial_x = 0$ and $\partial_y = 0$, the x and y-homo pairing effect trend cannot be established, instead of the data trending of the Pearson correlation analysis (Figure 5).

Source: Author (2022)

Figure 5: The data trending in the Pearson correlation analysis

Figure 5 evidences the co-movement of the paired data is equal in magnitude and direction. The x and y-homo pairing effects are equal, that is, $-\partial_x = -\partial_y$ and $\partial_x = \partial_y$. Therefore, the mean x and y-homo pairing is zero ($\bar{\partial}_x = \bar{\partial}_y = 0$), hence $\Psi_b = 0$. The graph shows the data moving in the same direction, with a mean value of x, $\bar{x} = 9$ and mean value of y, $\bar{y} = 6$. The paired data x and y are perfectly fitted to their respective means values, $\bar{x}$ and $\bar{y}$.

2.2.4 Meaning and Interpretation of Ψ_b, when $\Psi_b = 1$ and $r \neq 1$

When there is a perfect homo-hetero pairing effects of the two variables, x and y, $\Psi_b = 1$. This does not grant $r = 1$. Consider the hypothesis empirical data and their x and y-homo pairing effects in the table below,

x	y	∂_x	∂_y
2	7	2	2
4	5	2	2
7	2	-4	0.8
9	12	2	2
12	9	2	2
10	11		

From the table, we compute the homo-hetero pairing effect coefficient, Ψ_b and the Pearson correlation coefficient, r. The $\Psi_b = 1$ and the $r = 0.5140$. The $\bar{\partial}_x$ is 0.8 and $\bar{\partial}_y$ is 1.76 (Figure 6).
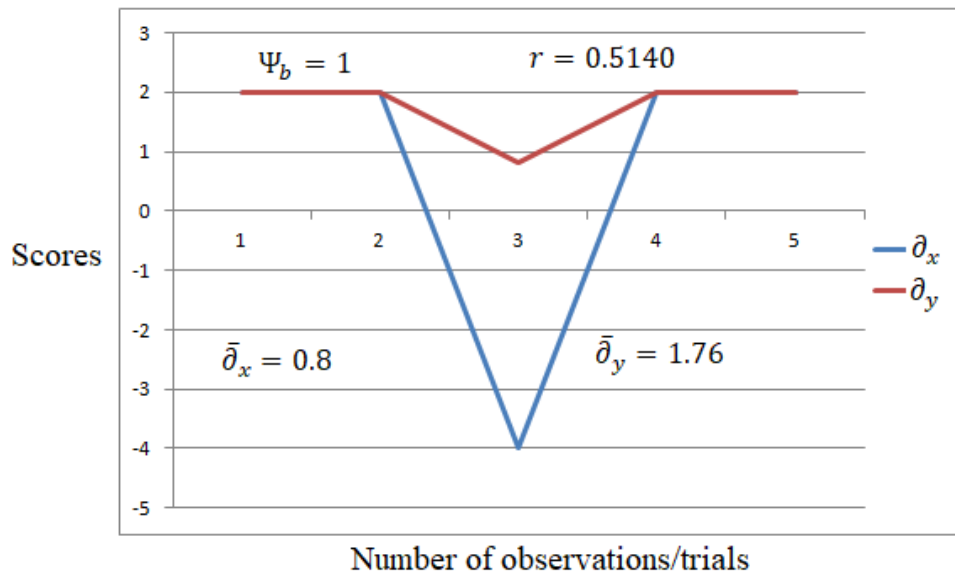
Source: Author (2022)

Figure 6: The x and y-homo pairing effects trending in Bundala correlation analysis

Figure 6 shows x and y-homo pairing effects trending of variables x and y respectively. The figure shows the perfect homo-hetero pairing effect coefficient, $\Psi_b = 1$, and $r = 0.5140$. The graph depicts the mean x-homo pairing effects, $\bar{\partial}_x$ is 0.8. This means that x changes are decreasing at the ratio or constant values of 0.8. On the other hand, the mean y- homo pairing effects of y, $\bar{\partial}_y$ is 1.76. This means, that values of y are increasing positively by a constant ratio $\bar{\partial}_y = 1.76$. This evidences that the perfect homo-hetero pairing effect correlation of the two variables, x and y, does not guarantee for perfect Pearson correlation of the variables.

2.3. Relationship Between r and Ψ_b

The Pearson correlation coefficient, r measures the "degree (strength) and direction of the co-variation of the x and y about their means $\bar{x}$ and $\bar{y}$ values respectively. The direction measure in r is indicating the "trend direction" of the paired data. Therefore, the paired data can be trending in a negative direction (negative r), or trending in a positive direction (positive r). On the other, the Bundala's Psi coefficient, Ψ_b measures the strength (pairing distance between the changes/effects of x and y) and direction of the co-variation of the x-homo pairing effects, $\partial_{x,}$ and y-homo pairing effect, $\partial_{y,}$ about their means homo pairing effects, $\bar{\partial}_{x,}$ and $\bar{\partial}_{y,}$ respectively. The direction measured in Ψ_b indicate the pairing direction or relationship of the paired data x and y. That is, the opposite direction of the pairing effects or changes of x and y indicated by negative Ψ_b, and the change of x and y occurs in the same direction (either increasing or decreasing) together, the value of Ψ_b is positive.

Both r and Ψ_b can be used to measure the correlation behaviour or information of the observed data, but Ψ_b explains how the observed data are "paired", that is the degree and direction of the pairing effect. Advantageously, Ψ_b can be used to check the robustness and perfectness of the r accuracy of co-movement of x and y as indicated by the dispersion (error) of the increase or decrease of the observed data in the pair. Noting that the co-movement of the paired variable x and y, increase or decrease may not always equal, i.e., $(x_i - x_j) = (y_i - y_j)$, it sometimes defers in size or magnitude, although are moving in the same direction, i.e., $(x_i - x_j) \neq (y_i - y_j)$. The variances or deviation of the magnitude change (decrease and increases) are detected by Ψ_b. In other words, if the Ψ_b is related to the r, the value of Ψ_b will indicate or measures the variance or variation of the changes (perfectness) of the observed variable x and y, as the Ψ_b measures the correlation of the difference or error pairing in r.

Technically, Ψ_b can be used to indicate or measure the pairing error in r. In this case, the r can be positive indicating that the data are either increasing or decreasing in the same direction but does not describes or explains the level or degree of the decrease of x and y or increase x and y are different or co-vary. In this case, it is impossible to identify the enabler and enabled factor or variable in r. In practice, r may be positive but Ψ_b is negative, which indicates that with either increase or decrease of paired variable x and y, the homo pairing change effect of observed variables x and y changes oppositely (∂_x going up and ∂_y going down). On the other hand, the r can be positive with a positive Ψ_b, indicates that either increase or decrease of paired variables x and $y,$ the homo pairing effect (error) changes in the same direction (∂_x and ∂_y upwards ∂_x and ∂_y downwards). And, the r can be negative and Ψ_b is positive, indicates that the data are oppositely changes (decrease or increases) while the magnitude or size of the changes is positively co-vary or related; they are going (moving) up (increase) and down (decrease) together. Also, the

negative r can be associated with negative Ψ_b indicates that the data are oppositely changing and the homo pairing effect also is moving oppositely (∂_x is going up and ∂_y is going down).

Generally, we can conclude that there is no either magnitude (strength) or directional relationship between Pearson correlation coefficient, r, and Bundala pairing correlation coefficient, Ψ_b. The r can be higher and Ψ_b is higher, or r is higher but Ψ_b is low and it is true and vice versa. Moreover, the r can be negative while Ψ_b positive or negative. This is because the Ψ_b does depends on r.

2.4. Enabler and Enabled Variable or Factor

Consider two correlated data, x and y in the paired data array. The methodological question is which between the two variables x and y, one is the enabler of the other variable (enabled variable)? This question traditionally was talked about by using experience or skills on the attributes of the variables in the study. This study introduces a new scientific way of identifying the enabler and enabled variables. The basic principle underpinning this concept is that the enabler should be totally or wholly absorbed in the enabled variable (Kenny, 1979; Onwuegbuzie and Daniel, 1999). In non-technical language, the enabler variable is a sub-part or element of the enabled variable. If x is the enabler variable of y, then, the following "pair-rule" would be adhered to, that $x \leftrightarrow y, x < y; \ x \to y$. This means, that the value of x is less than the value of y and is totally or wholly absorbed in the value of y, i.e., $y > x$. On the other hand, if y is an enabler of x, then, the paired relationship will adhere to $x \leftrightarrow y, x > y; \ x \leftarrow y$. That is, if the value of y is less than x than y, then y is totally and wholly absorbed in the value of x, *i.e.,* $y < x$; this is the mathematical reasoning. In logical reasoning, take an example of the pairing relationship between the risk factors of the infection of TB and HIV. Using the "pair-rule", that is,

$$\text{TB} \leftrightarrow \text{HIV}, \text{TB} > HIV; \ TB \leftarrow \text{HIV}$$

This means, in clinical language or reasoning, HIV is one of the clinical symptoms or signs of TB. Therefore, HIV is a sub-part or element of the symptoms or signs (risk factors) of TB. Now, alternatively,

$$\text{TB} \leftrightarrow \text{HIV}, \text{TB} < HIV; \ TB \to \text{HIV}$$

This means that if the risk factor of TB is less than the risk factors of HIV, then the risk factors of TB are part of the risk factors (symptoms and signs) of HIV. In other words, the risk factors of TB are the enabler factors or variables of HIV, and therefore HIV is the enabled variable (factor). For paired data (homo-hetero paired data) the enabler and enabled are identified with the same principle, but the mean x and y- homo pairing effect are applied. That is when x is an enabler of y the pairing rule is given by, $\bar{\partial}_x \leftrightarrow \bar{\partial}_y, \bar{\partial}_x > \bar{\partial}_y; \ x \leftarrow y$ and when x is an enabler of y the pairing rule is given by $\bar{\partial}_x \leftrightarrow \bar{\partial}_y, \bar{\partial}_x < \bar{\partial}_y; \ x \to y$. For more clarity, take examples from Figures 1 and 5. In Figure 1, the data have $r = 1$ and $\Psi_b = 0$. The mean homo paring effect for x and y are -5 and 0.833 respectively. Using the pair-rule we can identify the enabler and enabled variable, that is, $\bar{\partial}_x = -5$ and $\bar{\partial}_y = 0.833$, then,

$$\bar{\partial}_x \leftrightarrow \bar{\partial}_y, -5 > 0.833; \ x \leftarrow y: \text{Not true}$$

$$\bar{\partial}_x \leftrightarrow \bar{\partial}_y, -5 < 0.833; \ x \to y\text{:True}$$

Now from this example, we identify that x is an enabler factor or variable and y is the enabled factor or variable. That is, the change of the x will influence or enable y to change. The nature of change can be determined by the mean homo-hetero deviation, denoted as ϖ_b; this is the enabling effect value, or simply the enabler value. It is calculated by using the formula,

$$\bar{\partial}_y - \bar{\partial}_x = \varpi_b; \text{If} \ \ x \text{ is an enabler factor/variable}$$
$$\bar{\partial}_x - \bar{\partial}_y = \varpi_b; \text{If y is an enabler factor/variable}$$

Now, using the mean homo-hetero deviation equations, enabling effect value is obtained from the formula, $\bar{\partial}_y - \bar{\partial}_x = \varpi_b$, since x is an enabler factor; $\bar{\partial}_x = -5$ and $\bar{\partial}_y = 0.833$, we get, $0.833 - (-5) = 5.833$. This implies that the increases in effects of x are positively influenced by the increases in effects of y. The 5.833 indicates the trending-off effect value over the y-homo pairing effect. This extra mean value of y is totally and wholly absorbed in the value of x, to affect it (enabling it to change). This is why the Pearson correlation is positive. This value is always positive and indicates the positive or enabling effect of the variable. Furthermore, take another example in Figure 5 where $r = 1$, $\Psi_b = 0$, $\bar{\partial}_x = 0$ and $\bar{\partial}_y = 0$. From the pair-rule, since all the values of homo pairing changes in x and y are zero, indicating that the data has no homo-hetero pairing effect (x and y are not paired). The change of x is not paired with the change of y. That is why we say the paired data x and y have no homo-hetero pairing relationship because $\Psi_b = 0$. A question to ask is, which enabler and enabled variable? When $\Psi_b = 0$, there is a relationship between the enabler and enabled variable in the linear modelling. The mean homo-hetero deviation value is zero, which explains that there is no trending-off change value of either the x or y-homo pairing effect on y or x respectively. No variable its effect exceeds the other. The enabling effect of each variable is zero. These kinds of data are either all dependent or independent variables or all are enablers or enabled variables.

Let's learn a more example in Figure 3, in which $r = 0.9444$, $\Psi_b = -0.9715$, $\bar{\partial}_x = -0.1625$ and $\bar{\partial}_y = 0.0518$. In this case, the paired data of the observed variables x and y are positively correlated since it has a positive value of $r = 0.9444$. On the other hand, the paired data has a negative highly pairing relationship since it has negative values of $\Psi_b = -0.9715$. This means the positive correlation of x and y is associated with opposite changes in the homo pairing effects of x and y. The data are paired in a negatively way. That is, there is an enabler and enabled variable. Using the pair-rule, that is,

$$\bar{\partial}_x \leftrightarrow \bar{\partial}_y, \bar{\partial}_x > \bar{\partial}_y; \ x \leftarrow y\text{: If } y \text{ in an enabler variable of } x$$
$$\bar{\partial}_x \leftrightarrow \bar{\partial}_y, = -0.1625 > 0.0518; \ x \leftarrow y\text{: Not true}$$

$$\bar{\partial}_x \leftrightarrow \bar{\partial}_y, \bar{\partial}_x < \bar{\partial}_y; \ x \to y\text{: If } x \text{ is an enabler variable of } y$$
$$\bar{\partial}_x \leftrightarrow \bar{\partial}_y, -0.1625 < 0.0518; \ x \to y\text{: True}$$

Therefore, the x is an enabler variable of the variable y. Now use the mean homo-hetero deviation to calculate the enabler value or enabling effect value.

Since x is an enabler variable of variable y, then we used, $\bar{\partial}_y - \bar{\partial}_x = \varpi_b$

$$0.0518 - (-0.1625) = 0.2143$$

We notice that the increase of the change of the x variable positively influences the change of variable y. The trending-off change value of 0.2143 indicates the enabling effect value of the variable x on y. The higher the trending-off change values the stronger the enabling effect. Hence, we conclude that x and y are positively correlated with a highly negative homo-hetero pairing effect correlation.

2.5. Statistical Significance Test

It is a formal procedure of comparing the observed data with a claimed truth (hypotheses) (Moore, Notz and Flinger, 2013). Therefore, the statistical significance test aims to draw statistical or empirical evidence (inference) that the claimed hypotheses are representing the population of the study sample. The common statistical significance is the student t-test which is suitable or appropriate for small sample size, in practice, not more than 30 observations because as the sample becomes large the sample becomes normally distributed according to the central limit theorem and therefore is approximated to z-tests, and becomes relevant for a normal distribution (Moore, et al. 2013). The assumption of Ψ_b is that the observed variables x and y are linearly pairing or matched variables, therefore the appropriate statistical significance test is the *homo-hetero paired t-test*, modified from the normally paired t-test (Moore, et al. 2013). That is,

$$t_{\Psi_b} = \frac{\sum(\partial_{xi} - \partial_{yi})\sqrt{n-1}}{\sqrt{n\sum(\partial_{xi} - \partial_{yi})^2 - (\sum(\partial_{xi} - \partial_{yi}))^2}} \sim t(n-1)\text{df}$$

Where $t(n-1)$df is the degree of freedom, given by $n-1$.

2.6. Hypotheses Test

The hypothesis is the tentative solution or claimed truth of the problem on hand (Paiva, 2020). The hypotheses are of two types, null hypothesis (H_0) which is a negation statement about the claimed truth or expected facts. By default, the null hypothesis is represented with a negative statement about the fact in the issue (Massey and Miller, n.d). The second hypothesis is an alternative hypothesis or research hypothesis put down or described by a researcher concept. It is denoted by (H_1)or(H_a). It is a positive statement that explains the attentive solution to the research problem (Massey and Miller, n.d; Paiva, 2020). The alternative hypothesis H_1 can be formulated in three forms; it can follow the upper-tailed, lower-tailed and two-tailed tests (Paiva, 2020; Massey and Miller, n.d). For example, a researcher may want to know if $H_0: t_{\Psi_b} = t_{\alpha/2}$ for the null hypothesis and set an alternative $H_1: t_{\Psi_b} \neq t_{\alpha/2}$ (two-tailed test). The alternative

hypothesis $H_1: t_{\Psi_b} \neq t_{\alpha/2}$ can be further decomposed into two tests, upper (right)-tailed test ($H_1: t_{\Psi_b} > t_{\alpha}$) and lower (left)-tailed test ($H_1: t_{\Psi_b} < t_{\alpha}$). Where t_{Ψ_b}, $t_{\alpha/2}$ and t_{α} are computed statistic t-values and critical statistics t-value in a two (left and right)-tailed test and one-tailed test respectively.

2.7. Bundala's Ratios of Linear Correlation (Modelling)

Bundala's ratios or values are ratios that describe the conditions for optimality and empirical validity of the linear correlation techniques or models. The ratios are based on the relationship between true enabler (independent) and true enabled (dependent) variables in the linear modelling. The ratios are called by the name of the author to make them distinctive from other common ratios. There are two ratios developed by the author, which are the Gamma ratio (Γ_b) and Zeta ratio(Z_b). They are explained in detail in the next sections.

2.7.1. Gamma Ratio (Value)

The Bundala's gamma value or ratio (Γ_b) is the ratio used to identify or classify the enabler variables into either dependent or optimal independent variables. In other words, it is the test of independence of the variables for linear modelling (regression). In linear modelling, the enabler variables can be either dependent or independent variables.

Let *x* and *y* be two paired variables, that if *x* is an enabler variable of *y*, then, $x \to y$ and its regression model or function is $y = a_1 + b_1 x$, and when *y* is an enabler of *x*, that $x \leftarrow y$, the regression equation becomes $x = a_2 + b_2 y$. From the concept of Bundala's Gamma ratio, Γ_b

$$\Gamma_b = \frac{\text{Regression Coef. when enabled variable } (y) \text{ is independent}}{\text{Regression Coef. when enabler variable } (x) \text{ is independent}}$$

$$\Gamma_b = \frac{\text{Coef. } y}{\text{Coef. } x} = \frac{b_2}{b_1}$$

The coefficient of the linear equation $y = a_1 + b_1 x$ and $x = a_2 + b_2 y$ can be written in expanded form, that is,

$$b_1 = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$b_2 = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum y^2) - (\sum y)^2}$$

Therefore,

$$\Gamma_b = \frac{\text{Coef. } y}{\text{Coef. } x} = \frac{b_2}{b_1} = \frac{\left(\frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum y^2) - (\sum y)^2}\right)}{\left(\frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}\right)}$$

This can be simplified to,

$$\Gamma_b = \frac{n(\sum x^2) - (\sum x)^2}{n(\sum y^2) - (\sum y)^2}, \text{if } x \text{ is an enabler of } y$$

$$\Gamma_b = \frac{n(\sum y^2) - (\sum y)^2}{n(\sum x^2) - (\sum x)^2}, \text{if } y \text{ is an enabler of } x$$

Alternatively, the Γ_b can be computed using the concept of the total sum of the squares errors (TSSE) and total mean squares errors (TMSE) of the linear modelling (regression). That is,

$$\Gamma_b = \frac{\text{TSSE or TMSE of regression model when enabled variable } (y) \text{is independent}}{\text{TSSE or TMSE of regression model when enabler variable } (x) \text{is independent}}$$

$$\Gamma_b = \frac{TSSE_x}{TSSE_y} = \frac{TMSE_x}{TMSE_y} = \frac{\sum_{i-1}^{n}(x_i - \bar{x})^2}{\sum_{i-1}^{n}(y_i - \bar{y})^2}; \text{ if } x \text{ is an enabler variable}$$

$$\Gamma_b = \frac{TSSE_y}{TSSE_x} = \frac{TMSE_y}{TMSE_x} = \frac{\sum_{i-1}^{n}(y_i - \bar{y})^2}{\sum_{i-1}^{n}(x_i - \bar{x})^2}; \text{ if } y \text{ is an enabler variable}$$

That is, if Γ_b implies x is an enabler of y, then Γ_b^{-1} implies that y is an enabler of x.

Specifically, Γ_b is used to evaluate the validity of the enabler variable to qualify for "an optimal independent variable". The optimal or true independent is the variable that maximises the output (dependent) with fewer predictive errors or noise. Therefore, it is possible by using Bundala's Gamma value or ratio to identify the dependent and independent variables in the correlation analysis or linear modelling. The optimal range (true independent variable) of Bundala's gamma value Γ_b is from 0 to 1, that is $0 < \Gamma_b < \pm 1$. If $\Gamma_b = 0$ or $\Gamma_b = \pm 1$ indicates the data are not paired; hence there is no effect of enabler and enabled variables, and if the $\Gamma_b > \pm 1$ the variable x is not a true enabler or not a truly independent variable, it is an enabled (dependent) variable. Therefore, the Bundala Gamma value can be the most applicable statistical test in regression modelling/linear modelling to identify the optimal dependent and independent variables if they are wrongly regressed. It is an appropriate test to identify scientifically, the optimal independent and dependent variables in linear modelling.

2.7.2. Zeta Ratio (Value)

Bundala's Zeta ratio of the linear correlation (Z_b) describes the empirical relevance of the linear models, such regressions model. The ratio is taken as the percentage of the value of a unit contribution or impact of the true enabler (optimal independent) variable on the true enabled (dependent) variable when the true enabler variable takes the value of 0 and 1, that is, x (0,1). This ratio can be derived from the general linear equation,

$$f(x) = \sum_{i=1}^{n} \beta_i x_i + \beta_0, for\ x(0,1)$$

$$f(1) = \sum_{i=1}^{n} \beta_i + \beta_0 \text{ and } f(0) = \beta_0$$

Therefore,

$$Z_b = \frac{f(1) - f(0)}{f(1)}$$

$$Z_b = \frac{\sum_{i=1}^{n} \beta_i + \beta_0 - \beta_0}{\sum_{i=1}^{n} \beta_i + \beta_0} = \frac{\sum_{i=1}^{n} \beta_i}{\sum_{i=1}^{n} \beta_i + \beta_0}$$

If the linear equation is reduced to simple regression with one dependent and one independent variable, the Zeta value or ratio is calculated by using the formula,

$$Z_b = \frac{\beta_i}{\beta_i + \beta_0}$$

Where β_i and β_0 are beta coefficient and *y*-intercept or regression constant values respectively; β_i and β_0 are calculated as

$$\beta_i = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{n(\sum x_i^2) - (\sum x_i)^2}$$

$$\beta_0 = \frac{1}{n}\sum_{i=1}^{n} y_i - \beta_i \frac{1}{n}\sum_{i=1}^{n} x_i$$

The value of Z_b can take positive or negative values but never becomes zero. Positive values indicate the true enabler (optimal independent) variables in the model have a positive empirical impact on the true enabled (dependent) variable. Moreover, if the value is negative implies that the true enabler variable has a negative empirical impact on the dependent variable. The constant value β_0 is the value where the true enabler variables do not exist or are valued at zero scores. The interpretation of the Zeta ratio follows the following rule, when $\beta_0 = \beta_i$ or $\beta_0 < \sum_{i=1}^{n} \beta_i$(for multiple regression), the Z_b value is 0.5, therefore, the recommended Z_b value is at least 0.5 which indicates that $\beta_0 < \beta_i$ or $\beta_0 < \sum_{i=1}^{n} \beta_i$(for multiple regression). In the optimal linear equation, i.e., $\beta_0 = 0$ the Zeta value is equal to one, $Z_b = 1$. This implies that all the impacts or enabling effects in the enabled variable are due to the enabler variable. The value of constant or natural value is zero. On the other hand, when the constant value is negative, that is $\beta_0 < 0$, the Zeta value or ratio becomes more than 1, that is $Z_b > 1$. Moreover, the value of Z_b may become less than a critical value of 0.5, if $\beta_0 > \beta_i$ or $\beta_0 < \sum_{i=1}^{n} \beta_i$ (for multiple regression). If the Zeta value or ratio is greater than one, the model has more noise and is influenced by an unexplained value (constant value) or zero value of the independent variable(true enabler). In most practice, the optimal value of Zeta values should fall under the optimality range of $0.5 < Z_b \leq 1.5$ (Bundala, 2021).

3. Methodology

The study aimed to determine the pairing relationship between the economic growth and psychological human behaviour (psychological well-being) of the individual by application of the homo-hetero pairing effect correlation coefficient, or simply Bundala pairing correlation coefficient, Ψ_b technique. The study used the cross-sectional survey research design. The data were collected from 211 individuals randomly sampled from two regions in Tanzania, in 2020. Tabachnick and Fidell's (2019) approach was used to determine the sample size of the study. They suggest the sample size of $N > 50 + 8m$ for multivariate data analysis, and $N > 104 + m$ for testing individual predictors, where N is the sample size, m is the number of independent variables. The minimal sample size is $N > 104 + 1 = 105$ for a single independent variable. Therefore, the sample size of 211 individuals is reasonable for this study.

The psychometric scale technique was used to process the data as suggested by Bundala (2022), which is suitable for cross-sectional studies. The data is primarily analysed by using Ψ_b, and the results are compared with that of the *r* and regression model. This is to cross-examine the conclusion of the Ψ_b with *r* and regression model. The study variables are average gross domestic product (AGDP) per capita (measured in monthly annualised personal income) and psychological well-being. The annualised personal income is the summarization of the monthly monetised income from farming, business activities, salary, and other activities of the individual which are computed on the monthly consumption basis of the individual. Therefore, average GDP per capita (AGDP) = monthly income of an individual, *i* x 12 (number of months in a year). The monthly income is assumed to be constant over a year due to the nature of economic activities being similar and has almost constant monthly returns. On the other hand, the psychological human behaviour of an individual (psychological well-being) is measured by a psychometric scale developed by Bundala (2021). This psychometric scale is directly converting the 5-Point scale to an index number that ranges from 0 to 1. The general formula of the psychometric scale is,

$$Hube = \frac{\sum_{i=1}^{n} \sum_{i=1}^{m} Q_i \; x \; LS_i}{\sum_{i=1}^{n} Q_i \; x \; LS_{max}}$$

Where *n* is the number of variables measured in psychological factor, *m* is the number of questions that measure or is a proxy for variables, Q_i the question posed for variable *i*, LS_i the score of individual questions on the Likert Scale, and LS_{max} is the maximum score of individual questions on the Likert scale. The psychological human behaviour (Hube) which measures the psychological well-being of the individual is composed of three observed variables namely lifestyle, metacognition and motivation of the individual. The indicators of lifestyle are hedonistic, adventuristic and individualistic behaviour characteristics/attributes of the individual. Metacognition which measures the psychological awareness of the individual is indicated by three indicators namely, knowledge-based awareness, regulation awareness-based awareness and experiences-based awareness. The motivation of the individual is measured by two indicators

namely geographical motives of the individual to work and government policy and regulations support-based motive on work. The 5-Point Likert Scale was used to scale the observed variables lifestyle, metacognition and motivation.

The tests of linearity assumptions were done. The normality test was done by using the Jarque-Bera normality test and was found to be 7.04 and Chi (2) is 0.0296 which indicates that the data are normally distributed (Thadewald and Herbert, 2004). On the other hand, the multicollinearity test was done by using variance inflation factors (VIF) and was found to have a mean VIF of 4.38 which indicates the absence of a collinearity problem because it is below 5 (Murray, Nguyen, Lee...and David, 2012). Moreover, the reliability test was done for cross-sectional data by using Cronbach alpha which is found to be 0.6617 which is reasonably better as suggested by Tavakol and Dennick (2011).

4. **Result and Discussion**

The study examined the pairing relationship between economic growth and psychological well-being. The Bundala pairing correlation coefficient technique is used and its results are compared with other linear (bivariate) analytic techniques which are Pearson correlation and simple regression analysis. The data were analysed with the aid of the Microsoft (MS) Excel program.

4.1. Relationship Between Economic Growth and Psychological Well-Being

The objective is to examine the pairing relationship between economic growth and the psychological well-being (psychological human behaviour) of the individual by the application of t_{Ψ_b}. The null hypothesis is stated as "there is no significant pairing relationship between economic growth and the psychological well-being ($H_0: t_{\Psi_b} = t_{\alpha/2}$). The alternative hypothesis stated that "there is a significant pairing relationship between the economic growth and psychological well-being ($H_1: t_{\Psi_b} \neq t_{\alpha/2}$). The descriptive statistics are displayed (Table 2).

Table 2: Descriptive statistics for cross-section survey data in 2020 in Tanzania

Variable	Obs	Mean	Std. Dev.	Min	Max
Hube	211	.7250284	.1592616	.201	1
lnAGDP	211	.3978714	.5228518	-1.261602	1.607837
d_Hube	210	-.0635013	2.353792	-5.124332	4.94572
d_lnAGDP	210	.7557632	1.084683	-3.328894	4.3129

Source: Author (2022).

Table 2 shows the descriptive statistics for average gross domestic product transformed in natural logarithm (lnAGDP). The average of lnAGDP is 0.3979 TZS million (natural logarithm) and ranges from -1.2616 to 1.6078 TZS millions in natural logarithm. The other variable is psychological well-being measured in Hube (human behaviour index composed of three sub-

indicators, lifestyle, motivation and metacognition). Hube has an average of 0.7250; the score is based on the psychometric scale developed by Bundala (2021).

The study applied Ψ_b to examine the pairing relationship between economic growth and psychological well-being. The results were compared to the *r* and linear regression model. The parameters of Ψ_b and *r* is computed with the aid of the MS Excel program and it is summarised (Table 3).

Table 3: Parameters of Bundala's Psi and Pearson's rho coefficients

Ψ_b parameters		*r* parameters	
Hube observations (x)	211	Hube observations (x)	211
lnAGDP observations (y)	211	lnAGDP observations (y)	211
$\bar{\partial}_x$	-0.0635	$\bar{x}$	0.725028
$\sum(\partial_x - \bar{\partial}_x)^2$	1157.93	$\sum(x - \bar{x})^2$	5.326494
$\bar{\partial}_y$	0.755763	$\bar{y}$	0.397871
$\sum(\partial_y - \bar{\partial}_y)^2$	245.8965	$\sum(y - \bar{y})^2$	57.40855
$\sum(\partial_x - \bar{\partial}_x).(\partial_y - \bar{\partial}_y)$	155.4533	$\sum(x - \bar{x})(y - \bar{y})$	7.550915
$\sum(dx_i - dy_i)$	-172.046		
$\sum(dx_i - dy_i)^2$	1233.871		
$\left(\sum(dx_i - dy_i)\right)^2$	29599.67		

Source: Author (2022).

From Table 3, the Ψ_b is computed by using the formula,

$$\Psi_b = \frac{\Sigma(\partial_{xi} - \bar{\partial}_x)(\partial_{yi} - \bar{\partial}_y)}{\sqrt{\Sigma(\partial_{xi} - \bar{\partial}_x)^2 \Sigma(\partial_{yi} - \bar{\partial}_y)^2}}$$

$$\Psi_b = \frac{155.4533}{\sqrt{1157.93 \; x \; 245.8965}} = 0.2913$$

The statistical significance test can be computed by using the formula,

$$t_{\Psi_b} = \frac{\Sigma(\partial_{xi} - \partial_{yi})\sqrt{n-1}}{\sqrt{n\Sigma(\partial_{xi} - \partial_{yi})^2 - (\Sigma(\partial_{xi} - \partial_{yi}))^2}} \sim t(n-1)\text{df}$$

$$t_{\Psi_b} = \frac{-172.046 \sqrt{211-1}}{\sqrt{211x \ 1233.871 - 29599.67}} = -5.190$$

From the MS Excel program, we calculate the p-value, with a t-value of $t_{\Psi_b} = 5.190$, that is, $\left(t_{\Psi_b} < t_{\alpha/2} \right)$, degree of freedom, $df = 211 - 1 = 210$, t- paired test, two-sided tail test, which is equal to 0.000.

Decision rules*: Rule one,* is to compare either the *p-value* and *alpha value* (significance level) and *rule two is* to compare the computed statistics t-value and its critical values at the given significance level and degree of freedom. That is if the p-value < alpha value; we reject the null hypothesis and on the other hand, if the computed statistic t-value is greater than the critical value, $t_{\alpha/2}$ then we reject the null hypothesis, that is $t_{\Psi_b} > t_{\alpha/2}$. For this study the significant level, $\alpha = 0.05$, and degree of freedom, df $= n - 1$, that is, 211-1 =2010.

Decision on rule one: Compare the p-value and alpha value, the p-value is 0.000, and the alpha value is 0.05, which indicates that the p-value is less than the alpha value therefore we have a zero (0.000) probability to accept or support the null hypothesis. Therefore, we reject it and accept the alternative hypothesis that there is a pairing relationship between economic growth and human psychological well-being.

Decision on rule two: The computed statistics t-value, t_{Ψ_b} is -5.190 and the critical value is, $t_{0.025}$ -1.972 (from the t-table), at the degree of freedom of 210 (two-sided tails). The rule requires $t_{\Psi_b} \geq t_{\alpha/2}$ rejecting the null hypothesis. Therefore, we compare $t_{\Psi_b} = -5.190 > t_{0.025} = -1.972$. Then from this rule, we conclude that the calculated t-statistic value falls under the rejection area of the null hypothesis; hence we reject the null hypothesis and accept the alternative hypothesis. Rule one and rule two can be summarised by using the t-distribution curve (Figure 7).
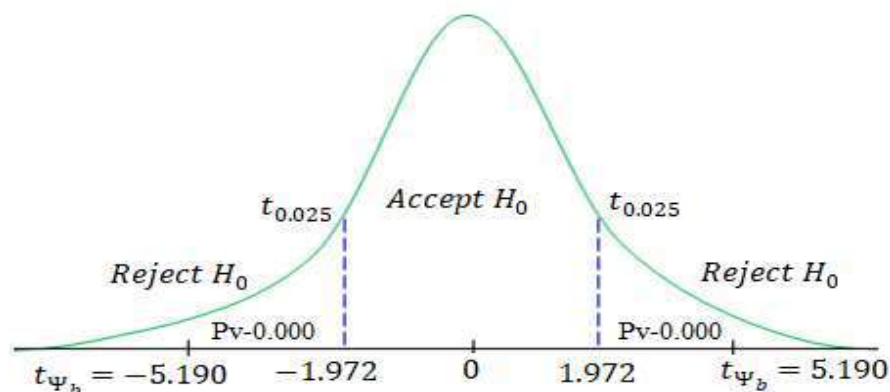
Source: Author (2022)

Figure 7: The t-distribution curve for decision rule with Bundala's Psi coefficient

Figure 7 shows the t-distribution curve. The figure describes the two decision rules. Rule one describes that the p-value at $t_{0.025} = \pm 1.972$ is 0.000 which is less than the alpha value (significance level) of 0.5/2 =0.025. Moreover, the second rule is shown by the rejection and acceptance boundary value, the critical value at $t_{0.025} = \pm 1.972$. The computed statistic t-value, $t_{\Psi_b} = -5.190$ falls under the null hypothesis rejection area with a more than 95 percent of confidence level. In general, we can combine or summarise the decision rule one and two in a single decisional statement that the probability that the computed t-statistic value, $t_{\Psi_b} = \pm 5.190$ is less that the statistic critical value, $t_{0.025} = \pm 1.972$ is zero. That is $P(t_{\Psi_b} < t_{\alpha/2}) = p - value = 0.000$. Therefore, the rejection of the null hypothesis is clearly illustrated by using the t-distribution curve.

4.2 Enabler and Enabled Variables /Values

The identification of enabler and enabled variables is very important in the correctional analysis as it helps the decision-makers to establish a work plan (decide accordingly). The enabler and enabled variables can be identified by using the pair-rule conditions as explained,

$$\bar{\partial}_x \leftrightarrow \bar{\partial}_y, \bar{\partial}_x > \bar{\partial}_y; \ x \leftarrow y: \text{If } y \text{ is an enabler variable of } x$$
$$\bar{\partial}_x \leftrightarrow \bar{\partial}_y, -0.0635 > 0.7558; \ x \leftarrow y: \text{Not true}$$

$$\bar{\partial}_x \leftrightarrow \bar{\partial}_y, \bar{\partial}_x < \bar{\partial}_y; \ x \rightarrow y: \text{If } x \text{ is an enabler variable of } y$$
$$\bar{\partial}_x \leftrightarrow \bar{\partial}_y, -0.0635 < 0.7558; \ x \rightarrow y: \text{True}$$

Thus, x (Hube) is an enabler variable and y(lnAGDP) is the enabled variable, the value of the enabler effect is computed by using the mean homo-hetero deviation, ϖ_b that is,

$$\bar{\partial}_y - \bar{\partial}_x = \varpi_b; \ \text{Since } x \text{ is an enabler variable of } y$$
$$0.7558 - (-0.0635) = 0.8391$$

The positive enabling effect value or simply enabler value of 0.8391 indicates the relative increment change of the mean y- homo pairing effect (economic growth) on the mean x-homo pairing effect (human psychological well-being)- the relative trending-off change value. It is the value that indicates the relative change variation of the paired data. Therefore, we have evidence that an increase in human psychological well-being scores increase also economic growth. Therefore, we conclude that there is a positive pairing effect between economic growth and the human psychological well-being of an individual. The homo-hetero pairing effect correlation coefficient (Ψ_b) is 0.2913. This value can be interpreted as "the pairing effect is moderate". This can be well demonstrated by using the x and y-homo pairing effect trending/variation graph (Figure 8).
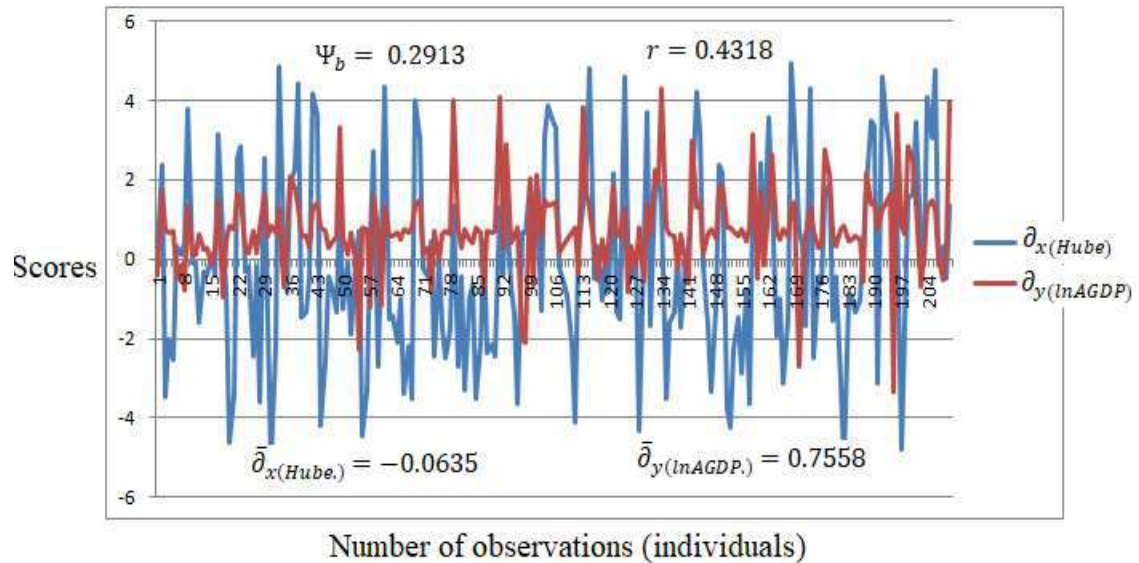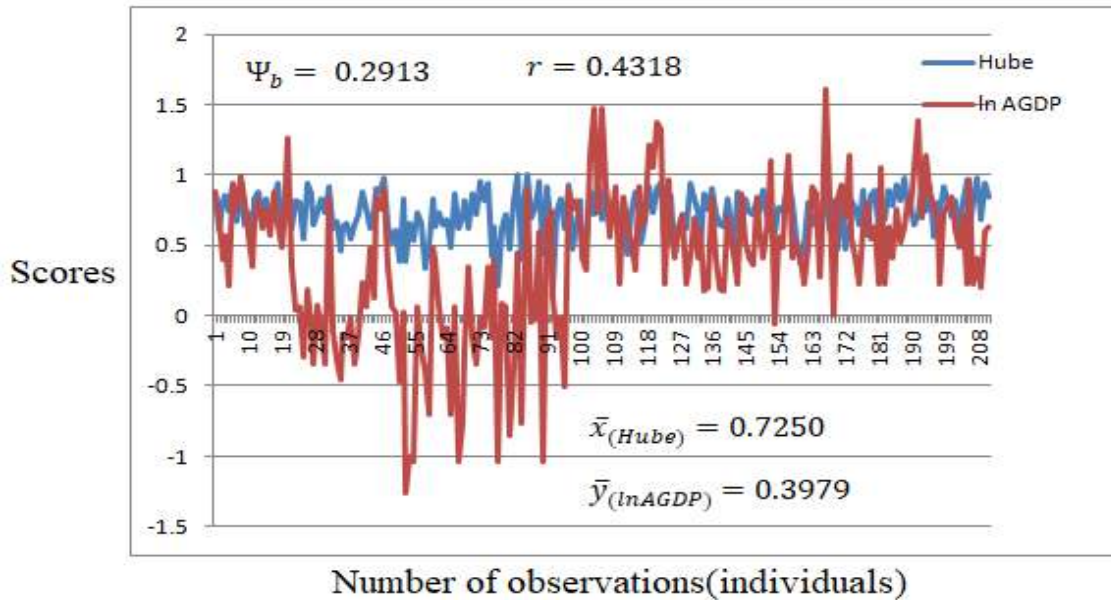
Source: Author (2022)

Figure 8: The *x* and *y*- homo pairing effects trending of the paired data

Figure 8 shows the *x* and *y* homo pairing effects trending. The figure indicates that there is a positive *relative trending-off changes value* of lnAGDP, which is averaged at 0.8391. The mean *x*-and *y*-homo pairing effect is -0.0635 and 0.7558 respectively. The data are paired with $\Psi_b = 0.2913$ and correlated with $r = 0.4318$. The value of Ψ_b indicates the pairing effect of the homo-hetero pairing in the data correlation/relationship between them. This matching or pairing effect cannot be shown or detected in the Pearson correlation analysis data trends (Figure 9).

The directional trending of the paired data in the Pearson correlation analysis shows that there are two classes of the population, the sub-populations that have a low economic growth (annualised monthly personal income) which is indicated with a negative value of lnAGDP and that class or sub-population that has a high economic growth which is represented by the positive values of lnAGDP (Figure 9). On the other hand, the psychological well-being of individuals is trending positively with less variance. That is, there are no significant classes based on the psychological well-being status among the individual in the population studied (Figure 9).

Source: Author (2022)

Figure 9: The data directional trending in Pearson correlation data analysis

Figure 9 shows the directional trending of the paired data in the Pearson correlation analysis. The data have r =0.4318, which means that they are moderately positively correlated. On the other hand, the pairing relationship of the data is shown by using Ψ_b = 0.2913, which indicates that data are moderately positively homo-hetero paired. The positive value of Ψ_b indicate that the data are paired or pairing in the same direction, not indicate the trending direction as done in the r. The mean value of the x and y variables are 0.7250 and 0.3979 respectively.

4.3 Bundala Ratios of Linear Correlation

Bundala ratios of the linear correlation are used in the correlation analysis to identify the true enabler (optimal independent) and truly enabled (dependent) for linear modelling. There are two types of Bundala ratios. One is the Gamma ratio which measures the "independence degree" of the enabler variable relative to the enabled variable. The second ratio is the Zeta ratio which is used to determine the "empirical relevance" of the enabler in the linear modelling (Bundala, 2021). These ratios are calculated to determine or identify the optimal independent and dependent variables among the enabler and enabled variables.

4.3.1 Gamma Ratio

Bundala's Gamma ratio is used to evaluate the "degree of the independence" of the enabler variable to the enabled (dependent) variable. In other words, it is used to evaluate the optimality of the independent variable in linear modelling (regression). It is given by the formula,

$$\Gamma_b = \frac{\text{Coef. } y}{\text{Coef. } x} = \frac{b_2}{b_1} = \frac{n(\sum x^2) - (\sum x)^2}{n(\sum y^2) - (\sum y)^2}$$

The coefficient y (InAGDP) and x (Hube) can be simply obtained by running two regression models, that is, $y(\ln AGDP) = a_1 + b_1 x(Hube)$ (Table 4) and $x(\text{Hube}) = a_2 + b_2 y(\ln AGDP)$ (Table 5).

$$\Gamma_b = \frac{\text{Coef. } y}{\text{Coef. } x} = \frac{b_2}{b_1}$$

$$y(\ln AGDP) = -0.6299394 + 1.417614x(Hube)$$

$$x(Hube) = 0.6726966 + 0.1315295y(\ln AGDP)$$

Therefore,

$$\Gamma_b = \frac{\text{Coef. } y}{\text{Coef. } x} = \frac{0.1315295}{1.417614} = 0.0928$$

Alternative,

$$\Gamma_b = \frac{TSSE_x}{TSSE_y} = \frac{TMSE_x}{TMSE_y} = \frac{\sum_{i-1}^{n}(x_i - \bar{x})^2}{\sum_{i-1}^{n}(y_i - \bar{y})^2}$$

$$\Gamma_b = \frac{TSSE_x}{TSSE_y} = \frac{TMSE_x}{TMSE_y} = \frac{5.3265}{57.40855} = 0.0928$$

We interpret that Γ_b is within the optimal range of $0 < \Gamma_b < \pm 1$, therefore, the Hube is a true enabler (optimal independent) variable that maximises the impact of economic growth, the enabled (dependent) variable.

To be clear with the optimal independent variable or true enabler variable, let's consider the linear equations when both enabler and enabled variables are involved and the independent and dependent variables are reversed. In the first equation, y (lnAGDP) is the dependent variable, and x (Hube) is the independent variable (Table 4).

Table 4: $y(lnAGDP) = -0.6299394 + 1.417614x(Hube)$

Source	SS	df	MS			
				Number of obs	=	211
				F(1, 209)	=	47.90
Model	10.7042859	1	10.7042859	Prob > F	=	0.0000
Residual	46.7042613	209	.223465365	R-squared	=	0.1865
				Adj R-squared	=	0.1826
Total	57.4085472	210	.273374034	Root MSE	=	.47272

lnAGDP	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
Hube	1.417614	.2048257	6.92	0.000	1.013825	1.821404
_cons	-.6299394	.1520284	-4.14	0.000	-.9296452	-.3302337

Source: Author (2022)

Table 4 shows the linear regression model of economic growth (lnAGDP) and human psychological well-being or simply psychological well-being (Hube). Economic growth is the dependent variable and psychological well-being is the independent variable. The unit impact of the Hube on economic growth (coefficient) is positive 1.41764, with a range of 1.013825 to 1.821404 at 95 percent of the confidence interval. The t-value is positive 6.92, with a p-value of 0.000, and the model is determined at an R-squared value of 0.1865. This implies that, although the data are poorly fitted to the model as indicated by the small value of R-squared, human psychological well-being has a positive impact on economic growth and is significant at a 95 percent of confidence level, since its p-value of 0.000 is less than the critical value of significance level, 0.05. Therefore, we reject the null hypothesis as we have zero probability of acceptance of the null hypothesis (p-value). We accept the alternative hypothesis that there is a pairing relationship between economic growth (dependent) and human psychological well-being (independent).

On the other hand, the linear regression was run when the variables are interchanged, the psychological well-being becomes the dependent variable and the economic growth becomes the independent variable. The aim is to evaluate or examine how the unit impact of each variable will be affected by the misallocation or use of a non-optimal enabler in linear modelling. In other words, it tests the unit independence impact of each variable in the linear modelling, and then we can able to identify which variable has a high unit independence impact on another variable, hence it is an optimal independent or true enabler of its pair variable, enabled (dependent) variable. The linear equation of Hube against economic growth is established (Table 5).

Table 5: $x(Hube) = 0.6726966 + 0.1315295y(lnAGDP)$

Source	SS	df	MS		Number of obs	=	211
					F(1, 209)	=	47.90
Model	.993167656	1	.993167656		Prob > F	=	0.0000
Residual	4.33332611	209	.020733618		R-squared	=	0.1865
					Adj R-squared	=	0.1826
Total	5.32649377	210	.025364256		Root MSE	=	.14399

Hube	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
lnAGDP	.1315295	.0190042	6.92	0.000	.094065	.1689939
_cons	.6726966	.0124674	53.96	0.000	.6481187	.6972746

Source: Author (2022)

Table 5 shows the regression model for the relationship between the Hube (dependent) and (lnAGDP) as an independent variable. The model is determined at an R-squared value of 0.1865, with a positive coefficient of 0.1315295, a t-value of 6.92, a p-value of 0.000, and a range of 0.094065 to 0.1689939 at 95 percent of the confidence interval. The model shows that there is a positive relationship between human psychological well-being (dependent variable) and economic growth (independent variable). Since its p-value is 0.000 which is less than the critical value of significance level, 0.05, therefore we reject the null hypothesis at a 95 percent of confidence level and accept the alternative hypothesis that there is a significant pairing relationship between economic growth and human psychological well-being (remember in the correlation analysis the variable can be interchanged without the effect of the strength or magnitude of the correlation coefficient).

From both regression models comes the same findings, we reject the null hypothesis at a 95 percent of confidence level. Also, we noticed a change in unit independence impact of the variable when the variable is interchanged, the enabler variable (Hube) to be enabled variable and the enabled variable (lnAGDP) to be the enabler variable. This difference can be measured or noticed by the *resource input-output (RIO) ratio*. The RIO ratio is the ratio that measures the relative unit impact of independent to dependent variables. For example, in the linear regression equation present in Table 4, the RIO is 1: 1.417614 or 0.7054, which means that a unit impact of the independent variable on the dependent variable is 1.417614 or about 70.54 percent of the total output (dependent variable) is due or explained by a unit impact of the independent variable. The remaining 28.46 percent (100-70.54) is explained by the *y*-intercept or the regression constant value. This ratio is very important in economic decisions as can be used for resource allocation, this is why the ratio is called the resource input-output (RIO) ratio. Simply, it is the ratio of input resources to output resources. In other words, RIO is the percentage that defines or explains the independence degree of the independent variable in the linear model. Moreover, in the linear regression equation in Table 5, when the Hube becomes or is treated as

the dependent variable and regressed to lnAGDP, the unit impact of the lnAGDP on Hube is 0.1315295 (Table 5). This means that the RIO ratio is changed to 0.1315295:1 which is equal to 0.1315 or 13.15 percent of the total output (dependent variable) is explained by a unit impact of the independent variable. About 86.85 percent is explained by the constant value of the regression model. We can notice that RIO ratios differ significantly. The RIO ratio for the optimal independent variable or true enabler (Hube) is higher than that of the non-optimal independent variable (lnAGDP). Therefore, the RIO ratio is the best statistical tool that can also be used to identify the optimal independent or test the optimality of the independence of the independent variable in linear modelling.

Therefore, from the concept of RIO ratio, we notice the relevance of relating (use of pair-rule) the value of β_0 and β_i . That is, the linear equation with a higher y-intercept, β_0 than beta coefficient, β_i has a less RIO ratio, that is, $\beta_0 > \beta_i$ and it is vice versa for $\beta_0 < \beta_i$. This concept originated from the enabler and enabled variables identification principle that the impact or effect of the "enabler coefficient " on the linear model should be totally and wholly absorbed in the effect of the "enabled coefficient" in the linear model. The enabler coefficient value should be greater than the enabled coefficient value. That is, $\beta_0 < \beta_i$ if β_0 enabler coefficient of is β_i and $\beta_0 > \beta_i$ if β_i is the enabler coefficient of β_0. The empirical relevance of the linear model is obtained when the true enabler variable or coefficient is used in the regression model; therefore the general rule is regression constant (y-intercept) to be less than the beta coefficient. To achieve the optimality of the independent variable the rule should be adhered to $\beta_0 < \beta_i$. We conclude that, for empirical relevance of the linear model, the beta coefficient should be greater than the constant value (y-intercept). We can apply this rule and get the same conclusion or results as the RIO ratio conclusion. Therefore, from the regression model in Table 4,

$$y(lnAGDP) = -0.6299394 + 1.417614x(Hube)$$

$$\beta_0 = -0.6299394 \text{, and } \beta_i = 1.417614$$

Therefore, $\beta_0 < \beta_i$ is true since$-0.6299394 < 1.417614$, then we conclude that the variable (Hube) is a true enabler (optimal independent) and lnAGDP is a dependent variable or truly enabled variable. Moreover, when the Hube (independent) variable is treated as the dependent (enabled) variable, the linear model becomes differently (Table 5). We apply the same concept of enabler coefficient and enabled coefficient; that is,

$$x(Hube) = 0.6726966 + 0.1315295y(lnAGDP)$$
$$\beta_0 = 0.6726966 \text{ and } \beta_i = 0.1315295$$

Therefore, $\beta_0 < \beta_i$ is not true since $0.6726966 > 0.1315295$, then we conclude that this model is empirically non-optimal, the enabler variable (lnAGDP) is not a true enabler; hence is not an optimal independent variable. If the values of RIO ratios are adjusted by their constant value

(commonly y-intercept), the value at a unit value of the independent variable, *x* (Hube) =1 the Bundala's Zeta ratio is used to identify the empirical optimality of the linear model (explained later).

4.3.2 Zeta Ratio

Another important Bundala linear correlation ratio is the Zeta ratio. The Zeta ratio describes the ability of the true enabler or identifies the true enabler by examining the "relevance of unit impact of the enabler variable (optimal independent variable) on the dependent variable in the empirical model. The rule described by Bundala (2021), the optimality of the Zeta ratio is at least 0.5. Therefore, by computing the Zeta ratios for regression equations in Table 3 and Table 4, and comparing them, a researcher or decision-maker can able to identify which linear model is relevant for empirical use. The higher the Zeta ratio, the higher the empirical relevance of the linear model is. The Zeta ratio is calculated by using the formula,

$$Z_b = \frac{\beta_i}{\beta_i + \beta_0}$$

Where β_i and β_0 are beta coefficients and y-intercept or regression constant values respectively. Consider the equation in Table 4,

$$y(lnAGDP) = -0.6299394 + 1.417614x(Hube)$$

$$\beta_0 = -0.6299394, \text{ and } \beta_i = 1.417614$$

$$Z_b = \frac{\beta_i}{\beta_i + \beta_0} = \frac{1.417614}{1.417614 - 0.6299394}$$

$$Z_b = 1.800$$

Then, compare the Zeta ratios from the equation in Table 4 to that of the equation from Table 5, that is,

$$x(Hube) = 0.6726966 + 0.1315295y(lnAGDP)$$

$$\beta_0 = 0.6726966, \text{ and } \beta_i = 0.1315295$$

$$Z_b = \frac{\beta_i}{\beta_i + \beta_0} = \frac{0.1315295}{0.1315295 + 0.6726966}$$

$$Z_b = 0.1635$$

Therefore, we compare the Zeta ratio of the equation in Table 4 which is 1.800 and that of the equation in Table 5, which is 0.1635, the Zeta ratio of the equation in Table 4 is greater than that of the equation in Table 5. This means, that the true enabler variable is Hube in the equation in Table 4. Moreover, in the equation in Table 4, its Zeta ratio exceeds the recommended upper optimality value of 1.5; therefore, the equation was violated with higher noises (standard deviation) in comparison to that of the equation in Table 5. Its standard deviation error is 0.2048257; hence the model has a less predictive accuracy compared with that in an equation in

Table 5. The equation in Table 4 has a higher sum of squares (SS) than that of the equation in Table 5 which is 57.4085472 and 5.32649377 respectively. On the other hand, the equation in Table 5 has the lowest value Zeta ratio which is 0.1635 indicating that the "assumed enabler" is not optimal; its effect on the output (Hube) is about 16.35 percent of the total effect of its unit value. This equation has fewer noises (standard deviation error) of 0.0190042 and higher predictive power than that in Table 4. Therefore, we concluded, that the true enabler variable which maximises the impact of the enabled (dependent) variable is human psychological well-being (Hube) and the dependent variable is economic growth (lnAGDP) exhibited in the equation in Table 4.

4.4 Pearson Correlation Coefficient, r

To enhance the conclusion of the finding, the conclusion or finding of Ψ_b is compared with the conclusion or finding of the Pearson correlation analysis. Therefore, the Pearson correlation coefficient, r is computed by using the formula,

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \ \sum(y_i - \bar{y})^2}}$$

$$r = \frac{7.550915}{\sqrt{5.326494 \ x \ 57.40855}} = 0.4318$$

Calculate the statistic significance test for r, use the formula

$$t_r = \frac{r.\sqrt{n-2}}{\sqrt{1-r^2}}$$

$$t_r = \frac{0.4318\sqrt{211-2}}{\sqrt{1-(0.4318)^2}} = 6.92$$

From the MS Excel program, we calculate the p-value, with a t-value of $t_r = 6.92$, that is, $(t_r < t_{\alpha/2})$, degree of freedom, $df = 211 - 2 = 209$, t- paired test, two-sided tail test, which is equal to 0.000.

Decision rules*: Rule one* is to compare either the *p-value* and *alpha value* (significance level) and *rule two is* to compare the computed statistics t-value and its critical values at the given significance level and degree of freedom. That is if the p-value < alpha value; we reject the null hypothesis and on the other hand, if the computed statistic t-value is greater than the critical value, $t_{\alpha/2}$ then we reject the null hypothesis, that is $t_r > t_{\alpha/2}$. For this study the significant level, α = 0.05, and degree of freedom, df = n − 2, that is, 211-2 =209.

Decision on rule one: Compare the p-value and alpha value, the p-value is 0.000, and the alpha value is 0.05, which indicates that the p-value is less than the alpha value, therefore we have a

zero (0.000) probability to accept or support the null hypothesis. That is, we reject it and accept the alternative hypothesis that there is a positive correlation between economic growth and human psychological well-being.

Decision on rule two: The computed statistic t-value, t_r is 6.92 and the critical value is $t_{0.025}$ =1.972 (from the t-table), at the degree of freedom of 209 (two-sided tails). The decision rule requires $t_r \geq t_{0.025}$ rejecting the null hypothesis. Therefore, we compare $t_r = 6.92 > t_{\Psi_{0.025}} = 1.972$. Then from this rule, we conclude that the calculated t-statistic value falls under the rejection area of the null hypothesis; hence we reject the null hypothesis and accept the alternative hypothesis. Rule one and rule two can be summarised by using the t-distribution curve (Figure 10).
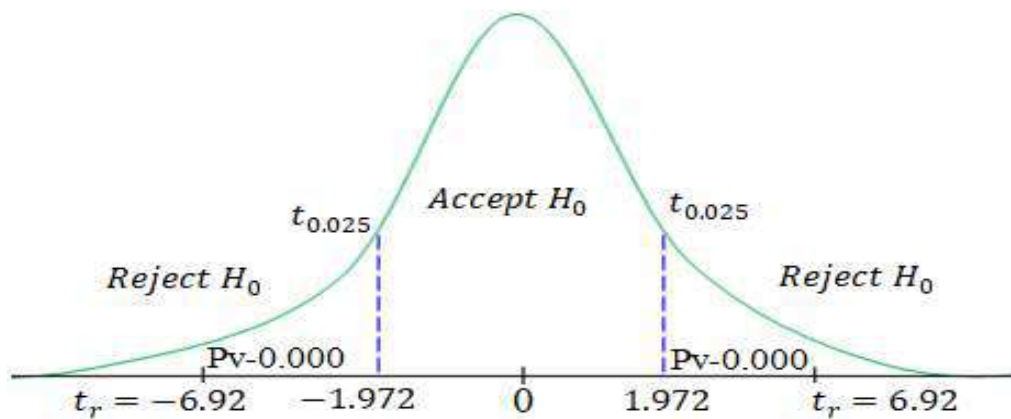
Source: Author (2022).

Figure 10: The t-distribution curve for decision rule with r

Figure 10 shows the t-distribution curve. The figure describes the two decision rules. Rule one describes that the p-value at $t_{0.025} = \pm 1.972$ is 0.000 which is less than the alpha value (significance level) of 0.5/2 =0.025. Moreover, the second rule is shown by the rejection and acceptance boundary value, the critical value at $t_{0.025} = \pm 1.972$. The computed statistic t-value, $t_r = 6.92$ falls under the null hypothesis rejection area with a more than 95 percent of confidence level. In general, we can combine or summarise the decision rule one and two in a single decisional statement that the probability that the computed t-statistic value, $t_r = \pm 6.92$ is less that the statistic critical value, $t_{0.025} = \pm 1.972$ is zero. That is $P(t_r < t_{\alpha/2}) = p - value = 0.000$. Therefore, the rejection of the null hypothesis is clearly illustrated by using the t-distribution curve.

4.5 The Methodological Cross-Examination of Ψ_b and r

The term cross-examination is commonly used in legal or law applications. It means the detailing of the witness (fact) provided to clear the doubt rose during the direct examination of the fact in issue. In this study, the term is used with the same meaning but for different purposes and logic.

Cross-examination aims to test the testimony (methodological completeness) provided by the two competing correlation techniques, the Pearson correlation coefficient and its competing partner Bundala pairing correlation coefficient before the Judge linear regression. It looks like a comical fact but it is a logical fact.

If we start the trial, the testimony or witnessed evidence by the Pearson correlation coefficient about the relationship between economic growth and human psychological well-being is that they are positively correlated at 0.4318. It is statistically supported with a rule that its empirical probability of t-value of 6.92, i.e., the p-value is 0.000 which indicates the probability to support the null hypothesis that it is true is zero (error type I). This is the empirical evidence provided by the Pearson correlation coefficient before the Judge linear modelling (regression model). On the other hand, the Bundala pairing correlation coefficient provided its testimony before the Judge that an increased change difference in the human psychological well-being level of the individual leads to an increased change difference (positively) in economic growth level. It is determined by the relative trending-off changes value (*x*-enabler effect value) of 0.8391. Moreover, the change difference of variables are paired as homo-hetero pairing effect coefficient, $\Psi_b = 0.2913$. Hence, it concludes that economic growth is positively influenced by the human psychological well-being of the individual. This evidence is supported by the same rule that its empirical probability of t-value -5.190, i.e., the p-value is zero, indicates the probability to accept the null hypothesis is zero.

Based on the direct empirical evidence provided by both correlation techniques before the Judge linear regression, some criteria were set to cross-examine the testimony (empirical evidence) provided by each of the two correlation techniques in the trial session (finding presentation). The criteria are based on the question that "who can be my best close partner". A Judge prepares a judging criteria matrix (Table 6).

Table 6: Methodological judging criteria matrix for Ψ_b and *r* coefficients

Judging criteria	*r*	Regression (Judge)	Ψ_b	Ruling
Do you quantify the strength and direction of the relationship between the two variables	Yes	Yes	Yes	Out
Do you predict how *x*(enabler variable) enables *y* (an enabled variable)	No	Yes	Yes	Yes
Both *x* and *y* are assumed to be random variables	Yes	No	No (one variable is an enabler of other	Yes
Do *x* and *y* are interchangeable, i.e., identical results are obtained when *x* and *y* are interchanged	Yes	No	Yes	Fair
Produces a statistical model	No (single statistic	Yes	Yes (Gamma and Zeta ratios)	Yes
Do you detect the homo paring effect of the *x* and *y* trending	No	No	Yes	Extra
Do you identify the optimal independent and dependent variable from the two related variables	No	Yes	Yes	Yes

Source: Author (2022).

Table 6 shows the methodological judging criteria matrix for the Bundala pairing correlation and Pearson correlation coefficients before the Judge linear regression. In the common language "the methodological completeness of the two correlation techniques Ψ_b and r were evaluated by using the attributes of the linear regression model as the standard linear technique, henceforth, a Judge. The ruling was done by three outcomes namely "out" which means all the judging criteria are evidenced (possessed) in each technique, either they have which is indicated by Yes or they have not which is indicated by No. The second judging ruling is "Yes" which indicates the Bundala correlation coefficient judging criteria are paired with that of the Judge and are contra to the Pearson correlation. A fair ruling is assigned when the outcome of the Bundala and Pearson correlation is contra to the Judge. The Judge is stand-alone to its criteria. The extra ruling is done when only one technique, particular Ψ_b has unique criteria that the Judge and r have not. In this trial, the Bundala pairing correlation coefficient has the extra ruling of having the ability to detect the homo-pairing effect of the trending paired data. The Judge rules out –that the Bundala correlation coefficient has hit the goal. That is, before the Judge linear regression the Bundala pairing correlation is a close partner because they share several decisional criteria.

5 Discussion

Principally, the paper aimed to examine the pairing relationship between economic growth and psychological well-being. The study applied the homo-hetero pairing effect correlation coefficient technique. This method is used to fill both the contradictory evidence gap on the relationship between the economic growth and psychological well-being of individuals (Roka, 2020; Easterlin, 1973; 2017; Stevenson and Wolfers, 2008; 2013) and the methodological gap (incompleteness) in the Pearson correlation coefficient technique that is raised or highlighted in several studies (Kennny, 1979; Rahman and Zhang, 2016; Janse, et al. 2021; Bertoldo, Callegher and Altoe, 2022). In other words, the homo-hetero pairing effect correlation coefficient or Bundala's Psi coefficient, Ψ_b (Bundala pairing correlation coefficient) for brevity is a new technique for correlation analysis in social science studies. The traditional correlation technique fails to identify the dependent and independent variables in linear modelling (Rahman and Zhang, Kenny, 1979; Emerson, 2015; Samuel and Okey, 2015; Maravelakis, 2019; Mudelsee, 2003). With the application of the Bundala pairing correlation coefficient, the study identified the enabler and enabled variable in the pairing relationship between the economic growth and psychological well-being of the individual, which further was classified into optimal independent and dependent variables by using the Bundala's ratios namely Gamma ratio(Γ_b) and Zeta ratio (Z_b). These ratios help to identify which variable in the pair is a true enabler of the other.

The study uses Ψ_b and found that the economic growth and psychological well-being of the individual are pairing related positively. That is, the increase of paired changes of the variable y is due to the increase of the paired changes of the variable x. Moreover, the study found that the psychological well-being of the individual is a true enabler (optimal independent) variable and has a positive impact on economic growth. This study confirms Roka (2020), Stevenson and Wolfer (2008; 2013), Talhelm et al. (2014), Diener and Seligman (2004), and Baro and Sala-i-Martin (2004). Moreover, this finding contradicts Stoop et al.(2019), and Easterlin (1973; 2017).

On the other hand, the Bundala pairing correlation coefficient has shown a significant role in linear regression. It was used to test the effectiveness of the independence of the variable in linear modelling. The study found that psychological well-being has a higher degree of independence than economic growth. Hence, psychological well-being is an optimal independent variable (true enabler) in linear modelling. The identification of the true enabler variable is most important in correlation studies because it helps to separate the paired variables into dependent and independent variables, which was not done in the Pearson correlation coefficient technique (Akoglu, 2018, Kumar and Chong, 2018; Senthilnathan, 2019). In other words, the Bundala pairing correlation coefficient has solved the problem of the Pearson correlation coefficient failing to define/classifies the paired variables into independent or dependent variables. The Bundala ratios either the Gamma ratio or Zeta ratio are the best techniques or methods that separate or identify which variables are either dependent or independent. The adverse effect of failing to identify the optimal independent or true enabler is that the use of non-optimal independent will have minimum impact on the output or dependent variable in the linear modelling. Consequently, may mislead the decision-makers.

Considering a finding of this study, in the linear correlation analysis of economic growth and psychological well-being, without the knowledge of which variable is independent and which other is dependent, the two possible regression models can be established, one can regress economic growth as the dependent variable and the psychological well-being as an independent variable. On the other hand, psychological well-being can be regressed as a dependent variable and economic growth as the independent variable. The two results are not similar in terms of unit independence impact (coefficient)-RIO ratio, statistical noises (standard deviation) and predictive accuracy (sum of the squares), which can confuse the decision-makers. Notably, this methodological error cannot be detected in the Pearson correlation coefficient instead by using the Bundala pairing correlation coefficient. The regression model that has a true enabler or optimal independent (psychological well-being) has a higher coefficient value (unit impact) or resource input-output (RIO) ratio than that has a non-optimal independent variable (economic growth). Therefore, the uses of this technique help to solve the methodological flaws that are exhibited in the traditional correlation techniques as claimed by several researchers (Kumar and Chong, 2018; Akoglu, 2018; Mukaka, 2012; Emerson, 2015; Bertoldo, Callegher and Altoe, 2022; Coppack, 1990; Shelef and Schechtman, 2018).

6 Conclusion and Recommendations

This paper aimed to examine the pairing relationship between economic growth and psychological well-being by the application of the homo-hetero pairing effect correlation technique or simply Bundala pairing correlation coefficient. The application of this technique is due to claimed methodological flaws of the existing correlation techniques, particularly, the Pearson correlation coefficient and the contradictory evidence gap on the relationship between economic growth and the psychological well-being of the individuals. The study found that the changes in economic growth are increases in pair with the increases in the change in the psychological well-being of the individual. In other words, the change in economic growth is positively paired with the change in the psychological well-being of the individual. Moreover, the psychological well-being of the individual is the optimal independent variable because have

found to have a positive empirical relevance in the regression model. Therefore, the study concluded that as far as the economic growth and psychological well-being are positively changing in pair, the improvement of the psychological well-being of the individual significantly improves the economic growth and not vice versa. Therefore, the paper recommended that psychological well-being-based initiatives should be established and encouraged in society as found to have a positive impact on economic growth.

The comparison of the Bundala pairing correlation coefficient and Pearson correlation coefficient techniques evidenced that the Bundala pairing correlation coefficient is a more robust and effective correlation technique. Hence, the Bundala pairing correlation coefficient is strongly recommended for contemporary social science studies. In addition, the specific recommendations for the application of the Bundala pairing correlation coefficient include the use in medical studies such as cardiology to examine risk factors or symptoms and signs of the heart diseases such as the homo-hetero pairing effect of systolic pressure and diastolic pressure can be detected by using the Ψ_b. Moreover, Ψ_b can be applied in other branches of medicine such as in neurology and epidemiology; it can be used to study the brain coordination disorders and control and monitoring of communicable and non-communicable diseases. This is because the Bundala pairing correlation coefficient identifies or detects the pairing effect or associate disease and its enabler risk factors or symptoms and signs. Moreover, the paper recommends the application of the Bundala pairing correlation coefficient in other branches of social science such as anthropology, sociology, economics, psychology and others.

References

Akoglu, H. (2018). User's Guide to Correlation Coefficient. Review Article. *Turkish Journal of Emergency Medicine, 18*(2018), 91-93.

Asuero, A.G., Sayago, A. and Gonzalex, A.G. (2006). The Correlation Coefficient: An Overview. *Critical Review in Analytical Chemistry*, *36*,41-59, doi: 101080/10408340500526766

Barro, R.J. and Sara-i-Martin, X. (2004). *Economic Growth*. New York, McGraw-Hill.

Benesty, J., Chen, J and Huang, Y.A.(2008). On the Importance of Pearson Correlation Coefficient in Noise Reduction. *IEEE Transactions on Audio, Speech and Language Processing, 16*(4), 757-765.

Bertoldo, G., Callegher, C.Z., and Altoe, G. (2022). Designing Studies and Evaluating Research Results: Type M and Type S Errors for Pearson Correlation Coefficient. *Meta-psychology*, 6, MP.2020.2573. doi: 10.15626/MP.2020.2573

Bundala, N. N. (2022). The Hidden Demographics Barriers of the Economic Growth: A Psychometric Approach. *International Journal of Business, Management and Economics, 3*(1). 24 - 51. doi:10.47747/ijbme.v3i1.471

Bundala, N.N (2021). *Structural Modelling of Psychological Limiting Factors for Economic Growth in Tanzania*. A PhD Thesis, Submitted for PhD award at the Open University of Tanzania, faculty of Arts and Social Science, Dar Es Salaam, Tanzania.

Chinnadurai, V. and Bobin, A. (2021). Picture Fuzzy Hypersoft TOPSIS Method Based on correlation Coefficient. *Journal of Hyperstructures*, *10*(2), 86-107.

Coppack, S.W. (1990). Limitations of the Pearson Product Moment correlation. *Clinical Science, 79*(3), 287.

Danacica, D.E and Babucea, A.G. (2007). Methodological Aspects in Using Pearson Coefficient in Analysing Social and Economic Phenomena. *European Research Studies, 11*(3-4), 89-97.

Diener, E. and Seligman, M.E.P. (2004). Beyond Money towards an Economy of Well-being. *Journal of Psychological Science in the Public Interest*, *5*(1), 1-31.

Easterlin, R. A. (2017). Paradox Lost? *Review of Behavioural Economics, 4*(4), 311-339. doi: 10.1561/105.00000068

Easterlin, R.A. (1973). Does Money Buy Happiness? *The Public Interest*, *30* (winter), 3-10.

Emerson, R.W.(2015). Causation and the Pearson Correlation Coefficient. *Journal of Visual Impairment and Blindness, 109*(3), 242-244.

Gogtay, N.J. and Thatte, U.M. (2017). Principles of Correlation Analysis Statistics for Researchers. *Journal of the Association of Physicians of India*, 65(march), 78-81.

Gokul,T., Srinivasan, M.R. and Swaminathan, V.D. (2021). Measuring Interchangeability In School Lunch intervention Data Using Concordance Correlation Coefficient. *Statistics and Application, 19*(2), 241-253.

Hernandez, O. and Alpizar-Jara, R. (2018). *Practical Use of Correlation Coefficient in the Social Sciences.* ICOTS10, Contribute Paper in M.A. Sorto, A.White and L.Guyot Eds), Looking Back, Looking Forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July 2018), Kyoto, Japan.

Janse, R.J., Hoekstra, T., Jager,K.J., Zoceali, C., Tripepi, G., Dekker, F.W. and van Diepen, M. (2021). Conducting Correlation Analysis: Important Limitations and Pitfalls. *Clinical Kidney Journal*, *14* (10), 2332-2337.

Kenny, D.A. (1979). Correlation and Causality. A Book Review. New York: John Wiley and Sons.

Kumar, S. and Chong, I. (2018). Correlation Analysis to Identify the Effective Data in Machine Learning: Prediction of Depression Disorder in Emotion Strategy. *International Journal of Environmental Research and Public Health*, 15, 2909. doi:103390/ijerph15122907

Li, M., Zhu, X., Li, W., Tang,X., Yu,X, and Jiang,Y. (2022). Retrieval of Nitrogen Content in Apple Canopy Based on Unmanned Aerial Vehicle Hyperspectral Images Using a Modified Correlation Coefficient Method. *Sustainability*,2022,14,1992. doi: 10.3390/su 14041992

Maravelakis, P.(2019).The use of statistics in Social sciences. *Journal of Humanities and Applied social science, 1*(2), 87-97. doi: 10.1108/JHASS-08-2019-0038.

Massey, A.and Miller, S.J, (n.d).*Tests of Hypotheses Using Statistics*. Mathematics Department, Brown University, Providence, RI 02912.

Moore, D. S., Notz, W. I, and Flinger, M. A. (2013). *The Basic Practice of Statistics (6thed.).* New York, NY: W. H. Freeman and Company.

Mudelsee, M.(2003). Estimating Pearson Correlation Coefficient with Bootstrap Confidence Interval From Serially Dependent Time Series. *Mathematical Geology*, *35*(6), 651-665.

Mukaka, M. M. (2012). Statistics Corner: A Guide to Appropriate Use of Correlation Coefficient in Medical Research. *Malawi Medical Journal*, *24*(3), 69-71.

Murray, L., Nguyen, H, Lee, Yu., Remmenga, M.D. and Smith, D.W. (2012). Variance Inflation Factors In Regression Models With Dummy Variables. *Conference on Applied Statistics in Agriculture.* doi. 10.4148/2475-7772.1034

Onwuegbuzie, A.J. and Daniel, L.G.(1999). *Uses and Misuses of the Correlation Coefficient.* Paper Presented at the Annual Meeting of the Mid-South Education Research Association, Point Clear, AL, November 17-19, 1999.

Paiva, A.(2020). *Hypothesis Tests*. ECE 3530, Spring, 2020.

Rahman, M.and Zhang, Q. (2016). Comparison Among Pearson Correlation Coefficients Tests. *Far East Journal of Mathematical Science (FJMS), 99*(2), 237-255. doi:10.17654/MS0 99020237

Roka, D.(2020). Nexus of Economic Growth on Happiness and Inequality Re-examine the Paradox. *Asian Journal of Economics, Business and Accounting, 15*(1), 35-49.

Samuel, M. and Okey, L.E. (2015). The Relevance and Significance of Correlation in Social Science Research. *International Journal of Sociology and Anthropology Research,2* (3), 22-28.

Sedgwick, P.(2012). Pearson's Correlation Coefficient. *BMJ*, 2012, 345, e4483. doi: 10.1136/ bmj.e4483

Senthilnathan, S.(2019). The usefulness of Correlation Analysis. *Article in SSRN Electronic Journal*. doi: 102139/ssm.3416918.

Shelef, A. and Schechtman, E. (2018). Correlation and time Interval over which the Variables are measured; A non-parametric approach. *PLoS ONE, 13*(1), e0206929. doi:10.1371 /journal.pone.02006929

Stevenson, B. and Wolfers, J.(2008). Economic Growth and Subjective Well-being Reassessing the Easterlin Paradox. *Brookings Papers on Economic Activity, 1, 1-87.*

Stevenson, B., Wolfers, J. (2013). Subjective Well-Being and Income: Is There Any Evidence of Satiation? *American Economic Review: Papers & Proceedings, 103*(3): 598–604. doi:10.1257/aer.103.3.598

Stoop, N., Leibbrandt, M. and Zizzamia, R. (2019). *Exploring Psychological Well-being and Poverty Dynamics in South-Africa Evidence from NIDS Waves 1-5*. Cape Town SALDRU, UCT,(SALDRU Working Paper Number 240 Version 1/NIDS Discussion Paper 2019/7).

Sverko, Z.,Vrankic, M.,Vlahinic, S.and Rogelj, P.(2022). Complex Pearson Correlation Coefficient for EEG Connectivity Analysis. *Sensors ,2022*, 22, 1477, doi:10.3390/3220 4177.

Tabachnick, B.G. and Fidell, L.S. (2019). Using *Multivariate Statistics,* 7th Ed. Pearson Education, Inc. USA.

Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X. and Kitayama, S. (2014). Large-Scale Psychological Difference within China Explained by Rice versus Wheat Agriculture. *Sciences, 344(6184)*, 603-608.doi: 101126/science.1246850

Tavakol, M. and Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education. 2011*(2), 53-55, doi:10.5116 /ijme.4dfb .8dfd

Taylor, R. (1990). Interpretation of the Correlation Coefficient: A Basic Review. *Journal of Diagnostic Medical Sonograph, 6*(1), 35-39.

Thadewald, T. and Büning, H. (2004). *Jarque-Bera Test and Its Competitors for Testing Normality: A Power Comparison, Diskussionsbeiträge*, No.2004/9, Freie Univer-sität Berlin, Fachbereich Wirtschaftswissenschaft, Berlin.

Copyrights