

ANALYSIS OF THE RELATIONSHIP BETWEEN ECONOMIC & SOCIAL FACTORS AND THE GLOBAL HAPPINESS INDEX USING A CANONICAL CORRELATION ANALYSIS APPROACH

Julia Widiyanti⁽¹⁾, Cantika Dhiya⁽¹⁾, Ganesya Intantalia⁽¹⁾, Maulana Syah Putra Ramadhani⁽¹⁾, Elly Pusporani^(1*), Elly Ana⁽¹⁾

¹Departement of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Indonesia

*Corresponding Author. E-mail: elly.pusporani@fst.unair.ac.id

ARTICLE INFO

Article History:

Received: 27-Jun. 2025

Revised: 04-Apr. 2026

Accepted: 06-Apr. 2026

Keywords:

Canonical correlation;
Economic factors; Social
factors; World
Happiness

ABSTRACT

This study aims to analyze the multivariate relationship between socioeconomic factors and global happiness indicators. Data taken from the 2019 World Happiness Report, which covers 125 countries, was analyzed using canonical correlation methods. The socio-economic dimension is represented by four variables, namely social support, freedom to make life choices, log GDP per capita, and healthy life expectancy. Meanwhile, the second group of variables representing happiness indicators includes ladder score, positive affect, and negative affect. The first canonical function shows a very strong and statistically significant relationship between the two sets of variables. The resulting canonical correlation is very high, namely $\rho_1 = 0.9243$ with a p-value < 0.001 , and this function is able to explain 85.4% of the overlapping variance. The main contributing variables are social support with a loading of 0.983 and log GDP per capita of 0.914 in the socio-economic set, as well as ladder score with a loading of 0.997 in the happiness set. Meanwhile, the second and third canonical functions generated in this analysis show relatively limited contributions. The results of this analysis confirm that social support and economic prosperity play a fundamental role in a country's happiness. The policy implications of these findings emphasize the need for integrated interventions that simultaneously strengthen social capital and promote sustainable economic growth.

This is an open access article under the CC-BY-NC-SA license



How to Cite:

Widiyanti, J., Dhiya, C., Intantalia, G., Ramadhani, M. S. P., Pusporani, E., & Ana, E. (2026). Analysis of the relationship between economic & social factors and the global happiness index using a canonical correlation analysis approach. *Journal of Mathematics Education and Science*, 9(1), 80-93. <https://doi.org/10.32665/james.v9i1.5076>

INTRODUCTION

Happiness is a key indicator of social life and the ultimate goal of every human being worldwide (Fukushima et al., 2021). Anas (2022) define happiness as a unique and significant phenomenon whose effects span the cognitive, affective, and behavioral domains of a person's life. Happiness is defined as a positive mental state, reflected in high levels of life satisfaction, strong positive emotions, and low levels of negative emotions (Carr, 2003). Globally, happiness is closely linked to various fundamental aspects of life, including economic conditions, employment, social

dynamics, and the fulfillment of basic needs (Cimpoeru, 2023). The strong impact of these aspects indicates that societal happiness and national development have a mutually reinforcing relationship. This means that an increase in societal happiness often goes hand in hand with inclusive and multidimensional progress in a country. Therefore, to formulate policies that can improve and enhance the quality of development in a country, the government must evaluate the performance or outcomes of previously implemented policies through various happiness index measurements (Indah & Magriasti, 2023).

This phenomenon marks a paradigm shift from a conventional economic approach centered on Gross Domestic Product (GDP) toward a more holistic framework that incorporates considerations of quality of life and life satisfaction as the ultimate goals of development (OECD, 2013). This shift considers quality of life and life satisfaction as more comprehensive indicators of progress (Pillai & Hu, 2024). In a number of developing countries, happiness indices have not only become a measure of subjective well-being but have also been adopted as a benchmark in formulating public policies aimed at achieving national development goals (Agustina & Wulandari, 2023). One of the primary indicators used is the ladder score, which represents individuals' subjective assessments of their lives as a whole. This indicator is multidimensional because it reflects not only economic achievements but also encompasses various non-economic factors, thereby providing a more complete and comprehensive picture of a nation's level of well-being (Helliwell et al., 2023).

Previous research on happiness indices has been conducted using various approaches. For example, Angela analyzed the effects of economic growth and per capita GDP on the happiness index across 33 provinces in Indonesia using the Chi-square association test (Angela, 2018). Meanwhile, Kumalasari and Yasa (2020) used regression analysis on data from 156 countries to identify the significant influence of macroeconomic variables on happiness levels. However, these methods generally focus on causal relationships by explicitly defining dependent and independent variables or analyzing one-to-one relationships between variables.

Based on this review, no previous study has examined the multidimensional relationship between socioeconomic factors and overall happiness simultaneously using a canonical correlation approach. This method was chosen because it aligns with the exploratory nature of this study. Canonical correlation was specifically developed to test relationships among sets of multivariate variables without requiring researchers to make causal assumptions at the outset. Based on this review, this study aims to address the existing research gap by investigating the multivariate and simultaneous relationships between socioeconomic determinants and measures of overall well-being. Based on the theoretical framework and previous empirical findings, it is proposed that there is a strong positive and statistically significant relationship between the set of socioeconomic variables and the set of happiness variables.

In this study, a canonical correlation analysis was conducted on two sets of variables: the happiness set (X), comprising the variables ladder score, positive affect, and negative affect; and the socioeconomic set (Y), represented by social support, freedom to make life choices, log GDP per capita, and healthy life expectancy (Helliwell et al., 2019). This study is expected to provide a comprehensive understanding of the nature of the relationship between happiness indicators and the various socioeconomic aspects that determine them. Furthermore, this study is expected to offer a comprehensive understanding of the relationships among these dimensions and serve as a foundation for developing more effective public policies aimed at continuously improving the happiness and quality of life of the community.

METHOD

Types and Sources of Data

This study is a descriptive quantitative study using a multivariate method, specifically canonical correlation analysis. The data source used in this study is secondary data obtained from the 2019 World Happiness Report. This dataset was published by the Sustainable Development Solutions Network (SDSN) and is openly accessible via the Kaggle website. The data covers various economic, social, and happiness indicators from countries around the world.

Population and Sample

The population in this study includes all countries listed in the 2019 World Happiness Report. From this population, a sample of 143 countries was selected that had complete data for all variables used in the analysis. The inclusion criteria for the sample were the availability of complete data for the seven variables analyzed, including both happiness-related variables and economic and social factors.

Research Variables

This study uses two main sets of variables:

1. Happiness Variables (X)
 - a. Ladder Score (X_1): individuals' subjective assessment of their overall life.
 - b. Positive affect (X_2): the average frequency of positive emotions experienced by respondents in a day.
 - c. Negative affect (X_3): the average frequency of negative emotions experienced by respondents in a day.
2. Set of Economic and Social Factors (Y)
 - a. Social Support (Y_1): the extent to which an individual has others they can rely on in difficult situations.
 - b. Freedom to Make Life Choices (Y_2): the degree of freedom people have in making their own life decisions.
 - c. Log GDP per capita (Y_3): the logarithm of Gross Domestic Product per capita, representing the level of economic well-being.
 - d. Healthy Life Expectancy (Y_4): indicating the level of public health.

Research Procedures

The data analysis procedure in this study was conducted through the following systematic steps:

1. Data Collection and Preliminary Processing
Data were drawn from the publicly available 2019 World Happiness Report dataset. Subsequently, data cleaning was performed to ensure that only countries with complete data for all seven variables used were included in the analysis.
2. Grouping Variables into Two Sets
 - a. Set of happiness variables (U_i): ladder score (X_1), positive affect (X_2), dan negative affect (X_3).
 - b. Set of socioeconomic variables (V_i): log GDP per capita (Y_1), social support (Y_2), healthy life expectancy (Y_3), dan freedom to make life choices (Y_4).
3. Assumption Tests for Canonical Correlation Analysis

Before performing a canonical correlation analysis, assumption tests are conducted to ensure the suitability of the data. To meet the requirements of the analysis, the following assumption tests are performed:

a. Detection of Multicollinearity

Multicollinearity is assessed by calculating the Variance Inflation Factor (VIF) for each set of variables. Mathematically, VIF is formulated as:

$$VIF_j = \frac{1}{1 - R_j^2} \quad (1)$$

where R_j^2 is the coefficient of determination from the regression results of the j-th variable on all other variables in the same set. A VIF value > 10 indicates the presence of serious multicollinearity that requires attention (Salmeron-Gomez et al., 2025).

b. Linearity Test

The linearity of the relationship between sets of variables was tested using the Ramsey RESET test. The testing procedure involved regressing each variable in one set against all variables in the other set, then adding quadratic and cubic terms of the predictor variables to the model (Bae et al., 2020). The F-value is obtained through the following equation:

$$F = \frac{(R_{new}^2 - R_{old}^2)/q}{(1 - R_{new}^2)/(n - k - q - 1)} \quad (2)$$

Here, R_{old}^2 is the coefficient of determination of the linear mode, R_{new}^2 is that of the model including quadratic and cubic terms, q is the number of additional predictors, n is the sample size, and k is the number of initial predictors. A relationship is considered linear when the p-value of the F-test statistic is > 0.05.

c. Normality Test

Data normality was tested univariately for each variable using the one-sample Kolmogorov-Smirnov test. This test works by comparing the observed data distribution to the theoretical normal distribution. The Kolmogorov-Smirnov test statistic is defined as:

$$D = \max |F_n(x) - F_0(x)| \quad (3)$$

where $F_n(x)$ is the empirical cumulative distribution function and $F_0(x)$ is the normal cumulative distribution function with parameters estimated from the data (Puritz et al., 2023). The test decision is based on the p-value; if p-value > 0.05, then the data are normally distributed.

4. Canonical Correlation Analysis

Canonical correlation analysis was used to examine the relationship between two sets of variables, happiness variables (X_1, X_2, X_3) and socioeconomic variables (Y_1, Y_2, Y_3, Y_4), after conducting assumption tests to ensure data suitability. The data analysis procedure in this study was carried out through the following stages.

a. Significance Test for Pairs of Canonical Variates

The significance test uses Wilks' Lambda (Λ) with the following formula (Sherry & Henson, 2005).

$$\Lambda_k = \sum_{i=1}^n (1 - \rho_i^2) \quad (4)$$

where ρ_i is the i-th canonical correlation, n is the number of possible canonical correlations (min(p, q)), p and q are the number of variables in the first and second sets, respectively. The hypotheses used in this test are as follows.

H_0 : The canonical correlations are not significant.
 H_1 : At least one canonical correlation is significant.
with the test statistic

$$F \approx -\left(n - \frac{p + q + 3}{2}\right) \cdot \ln \Lambda_k \quad (5)$$

The critical region for rejecting H_0 is when $F_{uji} > F_{(\alpha; v_1; v_2)}$, where $v_1 = p-k+1$ and $v_2 = q-k+1$.

b. Estimation of Canonical Correlation Coefficients

To obtain the canonical correlation coefficients (ρ_k), the square root of the eigenvalue (λ) of the cross-covariance matrix is calculated using the formula (Hair et al., 2006).

$$\rho_k = \sqrt{\lambda_k} \quad (6)$$

where λ_k is the solution to the equation

$$|\Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} - \lambda I| = 0 \quad (7)$$

where $\Sigma_{XX} \Sigma_{YY}$ is the covariance matrix of the first and second variable sets, respectively, and $\Sigma_{XY} \Sigma_{YX}$ is the cross-covariance matrix.

c. Visualization of Canonical Correlation

d. Estimation of Canonical Coefficients and Formation of Canonical Variables

Canonical coefficients are obtained from the eigenvectors that satisfy (Johnson & Wichern, 2007).

$$\Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} \mathbf{a} = \lambda_a \quad (8)$$

$$\Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} \mathbf{b} = \lambda_b \quad (9)$$

A canonical variable is defined as

$$\mathbf{U} = \mathbf{a}^T \mathbf{X} \quad ; \quad \mathbf{V} = \mathbf{b}^T \mathbf{X} \quad (10)$$

e. Interpretation of Canonical Correlation Values and Coefficients

The canonical correlation coefficient (ρ) and the canonical coefficient of determination (ρ^2) are interpreted to measure the extent to which the variability among sets of variables can be explained collectively. Additionally, an analysis of the canonical coefficients is conducted to determine the relative contribution of each variable to the resulting canonical variables.

5. Drawing Conclusions

Once the entire analysis process is complete, synthesize the results to identify the dominant relationships between socioeconomic factors and national happiness based on canonical correlation.

RESULTS

Testing the Assumptions of Canonical Correlation Analysis

Before conducting a canonical correlation analysis to test the relationship between the set of happiness variables and the set of socioeconomic variables, assumption tests were first

performed. The classical assumption tests conducted in this study included tests for multicollinearity, linearity, and multicollinearity. The results of these three assumption tests are presented as follows.

The multicollinearity test was used to determine whether there were high linear relationships among the independent variables within a set of variables. The test was conducted by calculating the VIF for each variable in the happiness set and the socioeconomic set. The results of the multicollinearity test are presented in Table 1.

Table 1. Results of the Multicollinearity Test

Variable Set	Variable	VIF
Happiness	Ladder Score (X_1)	1.591
	Positive Affect (X_2)	1.362
	Negative Affect (X_3)	1.423
Socioeconomic	Log GDP per Capita (Y_1)	2.282
	Social Support (Y_2)	1.304
	Healthy Life Expectancy (Y_3)	4.432
	Freedom to Make Life Choices (Y_4)	3.820

Based on Table 1, all variables in both sets have VIF values < 10. Thus, it can be concluded that no multicollinearity issues were detected in the research data, so the canonical correlation analysis can proceed.

Next, a linearity test was conducted to ensure that the assumption of a linear relationship between the set of happiness variables and the set of socioeconomic variables was met. The results of the linearity test are presented in Table 2.

Table 2. Linearity Test Results

Regression Model	F-statistic	p-value	Notes
$X_1 \sim all Y$	2.560	0.112	Linear
$X_2 \sim all Y$	0.240	0.625	Linear
$X_3 \sim all Y$	0.177	0.675	Linear
$Y_1 \sim all X$	0.001	0.976	Linear
$Y_2 \sim all X$	0.152	0.698	Linear
$Y_3 \sim all X$	1.262	0.263	Linear
$Y_4 \sim all X$	0.784	0.378	Linear

The results in Table 2 show that all tested regression models have *p-values* > 0.05. Thus, this indicates that the relationship between each variable in one set and all variables in the other set is linear. No violations of the linearity assumption were found, so the assumption is met.

Next, a univariate normality test was conducted on each variable using the one-sample Kolmogorov-Smirnov test. The results of the normality test are presented in Table 3.

Table 3. Normality Test Results

Variable Set	Variable	Statistic D	p-value
Happiness	Ladder Score (X_1)	0.0597	0.617
	Positive Affect (X_2)	0.0597	0.617
	Negative Affect (X_3)	0.0597	0.628
Socioeconomic	Log GDP per Capita (Y_1)	0.0597	0.635
	Social Support (Y_2)	0.0596	0.614
	Healthy Life Expectancy (Y_3)	0.0596	0.617
	Freedom to Make Life Choices (Y_4)	0.0596	0.617

Based on Table 3, all variables in both sets have p-values > 0.05. This indicates that, in a univariate analysis, each variable is normally distributed. Thus, all necessary assumptions have been met, allowing canonical correlation analysis to proceed to test the relationships among the canonical variables.

Testing the Relationship Between Pairs of Canonical Variates

The analysis of canonical variable pairs used Wilks' Lambda as the test criterion. A summary of the analysis results is presented in Table 4.

Table 4. Results of the Correlation Test Between Pairs of Canonical Variates

	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
1.	0.0790	42.29	12	315.14	<.0001
2.	0.5426	14.30	6	240	<.0001
3.	0.9351	4.20	2	121	0.0173

Based on Table 4, it can be seen that all significance values (Pr > F) are less than 0.05, indicating that all three canonical variable pairs are statistically significant. The first canonical pair has an F-value of 42.29 with a significance level of < 0.0001, followed by the second pair with an F-value of 14.30 and significance of < 0.0001, and the third pair with an F-value of 4.20 and significance of 0.0173.

Canonical Correlation Estimate

The results of the canonical correlation analysis are presented in Table 5 as follows.

Table 5. Results of the Canonical Correlation Analysis

	Canonical Correlation	Adjusted Canonical Correlation	Approximate Standard Error	Squared Canonical Correlation
1.	0.9243	0.9216	0.0130	0.8544
2.	0.6479	0.6389	0.0519	0.4198
3.	0.2547	0.2445	0.0836	0.0836

Based on Table 5, the canonical correlation coefficient (ρ^2) in the last column indicates the proportion of shared variance between pairs of canonical variables from the happiness set and the socioeconomic set. U_1, U_2, U_3 are, respectively, the first, second, and third canonical variables derived from the happiness set (X), while V_1, V_2, V_3 are the canonical variables from the socioeconomic set (Y). The ρ^2 value for the first canonical pair is 0.8544, indicating that 85,44% of the variance is shared between U_1 and V_1 . The second canonical pair has $\rho^2 = 0.4198$, indicating 41,98% of the shared variance between U_2 and V_2 , while the third pair has $\rho^2 = 0,0836$ equivalent to 8,39% of the shared variance between U_3 and V_3 . Thus, the first canonical pair is the most dominant, indicating a very strong canonical relationship and the most substantively meaningful one. These results confirm that only the first canonical pair makes a significant contribution to explaining the multidimensional relationship between socioeconomic factors and happiness. This is reinforced by the following plot visualization.

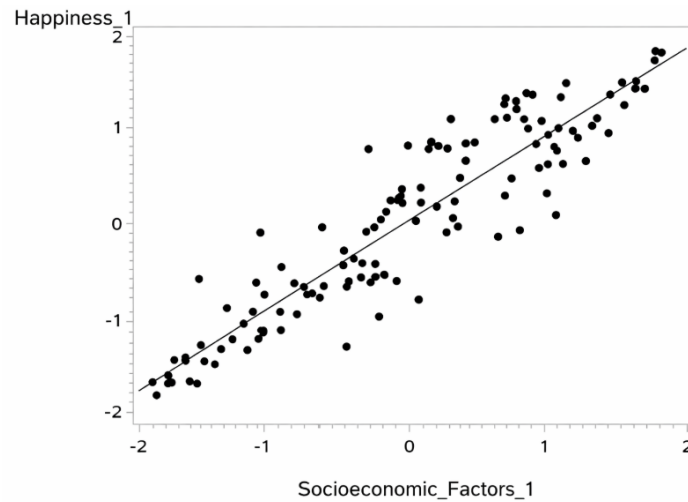


Figure 1. Canonical Correlation 1 Plot

Figure 1 shows the relationship between the first canonical variable of the socioeconomic factor and the first canonical variable of happiness. It appears that the data points are scattered along a strong and consistent linear pattern along the regression line, indicating a very high positive relationship between the two variables. The higher the socioeconomic factor score, the higher the happiness score, indicating that economic and social dimensions play an important role in explaining individual happiness levels. For comparison, a plot for the second canonical variable pair is presented below.

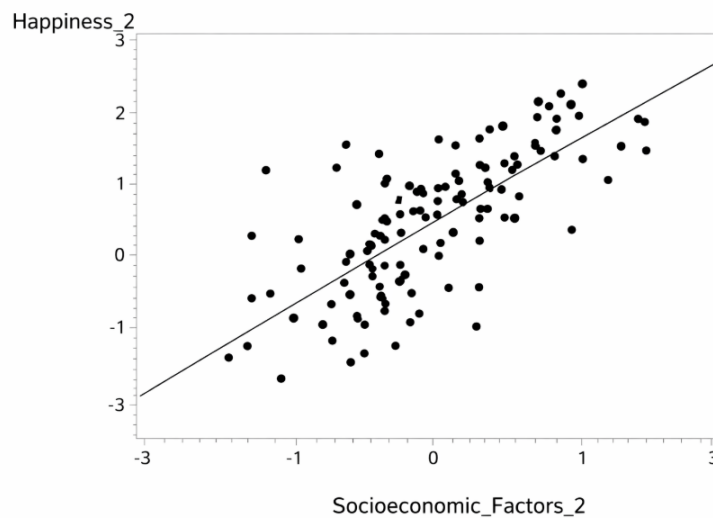


Figure 2. Canonical Correlation 2 Plot

Figure 2 shows a plot of the relationship between the second canonical variable of the socioeconomic factor and the second canonical variable of happiness. The data points still show a positive relationship pattern, but with greater dispersion and not as strong as in the first pair. The regression line remains upward, but many points are quite far from the line. This indicates that the relationship between the two second canonical variables is moderate, with a weaker correlation strength compared to the first pair. Next, the correlation plot of the third canonical variable is presented as follows.

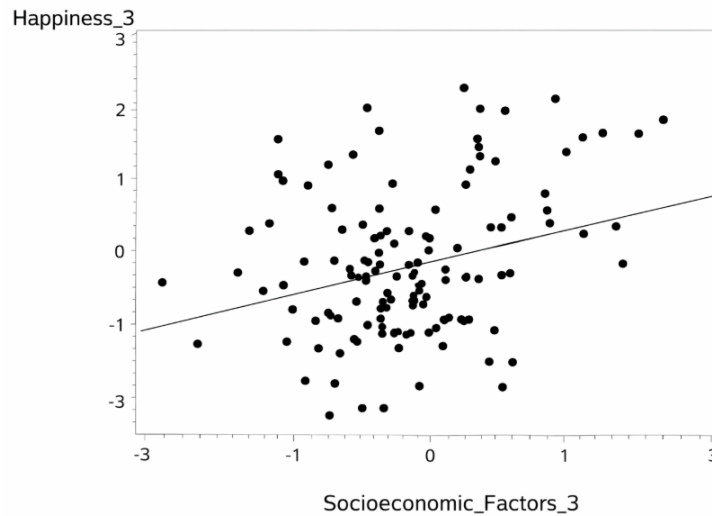


Figure 3. Canonical Correlation 3 Plot

Figure 3 illustrates the relationship between the third canonical variable of the socioeconomic factor and the third canonical variable of happiness. Based on the distribution of the data points, it is evident that the relationship between these two variables is very weak. The points are scattered randomly without following a clear linear pattern, and the regression line slopes only slightly upward, reflecting a very low positive correlation. Thus, it can be concluded that the third canonical pair does not make a significant contribution to explaining the relationship between the two groups of variables. Therefore, this pair may be considered non-priority in substantive interpretation and decision-making.

Canonical Coefficient

The estimated canonical coefficients for the happiness variable are presented in Table 6 as follows.

Table 6. Canonical Coefficients for the Happiness Variable

	Happiness_1	Happiness_2	Happiness_3
X_1	0.017	- 0.013	0.018
X_2	0.002	0.026	0.002
X_3	0.006	-0.002	-0.026

Based on the canonical coefficient estimates presented in the corresponding table columns, the transformation of the manifest variables into canonical happiness variables (U_i) is performed through linear operations. This synthesis process essentially yields a composite variable that captures the general dimension of happiness, as shown in the following equations.

$$U_1 = 0.017X_1 + 0.002X_2 + 0.006X_3 \quad (8)$$

$$U_2 = -0.013X_1 + 0.026X_2 - 0.002X_3 \quad (9)$$

$$U_3 = 0.018X_1 + 0.002X_2 - 0.026X_3 \quad (10)$$

Similarly, the estimates of the canonical coefficients for the socioeconomic factors are presented in Table 7 as follows.

Table 7. Cannonic Coefficients for Socioeconomic Factors

	Socioeconomic Factors 1	Socioeconomic Factors 2	Socioeconomic Factors 3
Y_1	0.011	-0.002	-0.026
Y_2	0.006	0.023	0.004
Y_3	0.004	-0.011	-0.012
Y_4	0.005	-0.006	0.038

Based on the canonical coefficient estimates presented in the corresponding table columns, the transformation of manifest variables into canonical variables for economic and social factors (V_i) is performed through linear operations. This synthesis process essentially yields composite variables that capture the common dimensions of economic and social factors, as shown in the following equations.

$$V_1 = 0.011Y_1 + 0.006Y_2 + 0.004Y_3 + 0.005Y_4 \tag{11}$$

$$V_2 = -0.002Y_1 + 0.023Y_2 - 0.011Y_3 - 0.006Y_4 \tag{12}$$

$$V_3 = -0.026Y_1 + 0.004Y_2 - 0.012Y_3 + 0.038Y_4 \tag{13}$$

The absolute values of the canonical coefficients in the U_i and V_i equations serve as indicators of the relative importance of each predictor's contribution. Operationally, coefficients with larger absolute values indicate a more dominant influence in shaping the canonical variables U_i (happiness) and V_i (socioeconomic factors), while the sign of the coefficient (+/-) determines the direction of the relationship.

Effect Size and Contribution of Variables

In addition to statistical significance, effect size calculations are essential for estimating the strength of a relationship in substantive terms. The squared canonical correlation (ρ^2) in canonical correlation analysis is an effect size indicator that reflects the proportion of shared variance between pairs of canonical variables. Cohen's guidelines (1988) categorize effect sizes as small ($\rho^2 \approx 0.15$), moderate ($\rho^2 \approx 0.36$), and large ($\rho^2 \approx 0.65$) (Lovakov & Agadullina, 2021). A summary of the effect sizes for the three canonical functions based on the established criteria is presented in Table 8.

Table 8. Effect Size and Substantive Interpretation of Canonical Functions

Canonical function	ρ^2	Effect Size	Interpretation
1	0.854	> 0.65	Very Large
2	0.420	> 0.36	Large
3	0.084	< 0.15	Very Small

Based on Table 8, the first canonical function has a very large effect size with $\rho^2 = 0.854$, meaning that 85.4% of the shared variance between the socioeconomic set and the happiness set can be explained by this canonical pair. The effect size of the second canonical function is classified as large, while the third function shows a small effect size. Although all three canonical functions reached statistical significance, these results indicate that only the first function provides a substantive contribution with sufficient practical power. The second and third functions provide only a marginal contribution to understanding the relationship between the two sets of variables.

Variables with the highest contribution to the formation of canonical dimensions were identified using standardized canonical loadings, which are the simple correlation coefficient between the manifest variables and their corresponding canonical variables. A threshold of ≥ 0.50 serves as the criterion for a variable's substantive contribution (Afthanorhan et al., 2020). Table 9 and Table 10 present the standardized loadings for each group of variables.

Table 9. Standardized Canonical Loadings for the Happiness Scale

	Function 1	Function 2	Function 3	Interpretation
X_1	0.971	-0.089	0.220	Substantive (F1)
X_2	0.573	0.816	0.072	Substantive (F1, F2)
X_3	0.731	-0.012	-0.683	Substantive (F1, F3)

All happiness variables in the first canonical function yielded standardized loadings above 0.50, with the ladder score (X_1) contributing the most compared to the other variables. Thus, the first function reflects a comprehensive construct of happiness that is specifically underpinned by the dominance of the evaluative aspects of life. The second canonical function is dominated by positive affect (X_2), indicating that this function better represents the positive affective dimension of happiness. Meanwhile, the third canonical function includes contributions from and negative affect (X_3), where X_3 represents the negative affective dimension with an inverse relationship.

Table 10. Standardized Canonical Loadings for the Socioeconomic Scale

	Function 1	Function 2	Function 3	Interpretation
Y_1	0.938	-0.135	-0.275	Substantive (F1)
Y_2	0.672	0.721	0.108	Substantive (F1, F2)
Y_3	0.878	-0.336	0.071	Substantive (F1)
Y_4	0.871	-0.278	0.403	Substantive (F1)

All indicators in the socioeconomic domain within the first canonical function yielded standardized loadings greater than 0.50, with social support and the log of GDP per capita emerging as the two variables with the strongest contributions. This confirms that social support and economic prosperity are the main pillars in the relationship between socioeconomic conditions and overall happiness. The second canonical function only had a substantive contribution from freedom to make life choices, indicating that this function specifically represents a dimension of individual freedom that is relatively independent of economic-material factors. The third canonical function did not show a substantive contribution from any variable.

DISCUSSION

This study confirms that the logarithm of GDP per capita and social support are dominant contributors to the relationship between socioeconomic factors and overall happiness. These findings are consistent with the World Happiness Report (Helliwell et al., 2023), which consistently identifies GDP and social support as the strongest predictors of the ladder score across countries. Empirical support also comes from the study by Berisha et al. (2025) in Southeast European countries from 2010 to 2021, which used the Ordinary Least Squares (OLS) method to confirm that social support, freedom to make life choices, and emotional well-being are crucial in shaping happiness. Meanwhile, the study by Bruna and Rungo (2020) on Latin American and U.S. residents demonstrates the marginal impact of income and family support on happiness. This consistency across methods and regions indicates that the central role of GDP and social support is robust and independent of specific analytical approaches.

Another interesting finding is that the affective dimensions of happiness (positive affective and negative affective) only emerge substantively in the second and third canonical functions, with

a much smaller contribution. This confirms the affective evaluation theory that life evaluation is more responsive to material conditions than emotional well-being (Dance et al., 2025). Cross-national variability in the second and third functions indicates contextual heterogeneity. In collectivist countries like Vietnam, happiness remains high despite low GDP due to strong social solidarity (Tran, 2020). This reinforces that culture moderates the relationship between socioeconomic factors and happiness; however, the absence of cultural variables in the model constitutes a limitation of the study. The implications of these findings are recommendations for optimal, integrated policies that incorporate socioeconomic and cultural factors through community-based social protection, the development of public spaces, and education in social-emotional skills. Methodologically, this study highlights the relevance of canonical correlation in public policy analysis, offering a multidimensional analytical framework that has long been overlooked in the happiness economics literature, which is dominated by regression analysis.

CONCLUSION

Canonical correlation analysis demonstrates that the first canonical function explains the relationship between socioeconomic factors and overall happiness very strongly and significantly ($\rho^2 = 85.44\%$), with social support and log GDP per capita as the dominant contributors in the socioeconomic set, while the ladder score dominates in the happiness set. The second and third canonical functions have limited contributions, indicating that the affective dimension of happiness is more autonomous and influenced by factors outside the scope of this study. These findings confirm that economic well-being and the strengthening of social capital are complementary as the foundation of national happiness.

Theoretically, this study contributes to the development of the happiness literature by treating happiness and its determinants symmetrically through a multidimensional approach. Methodologically, this study implements canonical correlation, a method still rarely applied in the economics of happiness, as a public policy analysis tool capable of capturing the complexity of multivariate relationships. The practical implications of this study are the availability of an empirical basis for designing integrative policies that combine economic growth agendas with the strengthening of social capital. The limitations of this study primarily lie in its cross-sectional design, which does not allow for causal inferences; the use of aggregate-level data, which is prone to the ecological fallacy; and the exclusion of cultural variables, thereby preventing a comprehensive exploration of cross-national differences. Future research is advised to use a longitudinal approach, multilevel analysis with individual-level data, and to integrate cultural dimensions and other contextual variables to gain a more comprehensive understanding of the determinants of happiness.

AI ACKNOWLEDGMENT

The authors declare that no generative AI or AI-assisted technology was used in any form during the writing, drafting, or editing of this manuscript. The entire content of the article is entirely the authors' original work, and the authors assume full responsibility for the content in accordance with COPE guidelines.

INFORMED CONSENT

The authors state that they have obtained informed consent from all participants involved in this study. Participants were provided with a sufficient explanation of the study's objectives, procedures, and their rights, and they agreed to participate voluntarily.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest in this study. All research processes and the writing of this article were conducted objectively and independently, without influence from any party.

REFERENCES

- Anas, M., Umar, N. F., & Harum, A. (2022). Faktor-faktor yang mempengaruhi kebahagiaan siswa. *JURKAM: Jurnal Konseling Andi Matappa*, 6(1), 51-64. <https://doi.org/10.31100/jurkam.v6i1.2123>
- Afthanorhan, A., Awang, Z., & Aimran, N. (2020). An extensive comparison of CB-SEM and PLS-SEM for reliability and validity. *International Journal of Data and Network Science*, 4(4), 357-364. <https://doi.org/10.5267/j.ijdns.2020.9.003>
- Agustina, W., & Wulandari, W. (2023). Collaborative governance dalam meningkatkan indeks kebahagiaan masyarakat (index of happiness) di Kota Bandung. *Jurnal Pemerintahan dan Politik*, 8(3), 173-182. <https://doi.org/10.36982/jpg.v8i3.3164>
- Angela, A. L. (2018). Analisis indeks kebahagiaan di Indonesia. *Equator Journal of Management and Entrepreneurship*, 6(1), 1-14. <https://doi.org/10.26418/ejme.v6i1.28791>
- Bae, E., Hur, J., Kim, J., Kwon, J. S., & Lee, J. (2020). Multi-group analysis using generalized additive kernel canonical correlation analysis. *Scientific Reports*, 10, Article 12624. <https://doi.org/10.1038/s41598-020-69575-x>
- Berisha, B., Qehaja, D., Hoti, A., & Gara, A. (2026). Wealth gap and happiness: A Southeast European (SEE) perspective. *International Journal of Sociology and Social Policy*, 46(1-2), 261-273.
- Bruna, F., & Rungo, P. (2020). A note on the concavity of the happiness function in family support. *Economics Bulletin*, 40(2), 1122-1131.
- Carr, A. (2003). *Positive psychology: The science of happiness and human strengths*. Routledge. <https://doi.org/10.4324/9780203506035>
- Cimpoeru, S. (2023). Impact of income and unemployment on happiness: Panel data evidence for European countries. *Applied Economics Letters*, 30(8), 1047-1051. <https://doi.org/10.1080/13504851.2022.2035666>
- Dance, C., Lybecker, D., & Szczygiel, N. (2025). Well-being and life satisfaction versus income. In *Vocabulary for sustainable consumption and lifestyles* (pp. 238-242). Routledge.
- Fukushima, S., Uchida, Y., & Takemura, K. (2021). Do you feel happy when other members look happy? Moderating effect of community-level social capital on interconnection of happiness. *International Journal of Psychology*, 56(5), 642-653. <https://doi.org/10.1002/ijop.12744>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2006). *Multivariate data analysis* (6th ed.). Pearson Prentice Hall.
- Helliwell, J., Layard, R., & Sachs, J. (2019). *World happiness report 2019*. Sustainable Development Solutions Network.
- Helliwell, J. F., Layard, R., & Sachs, J. D. (2023). *World happiness report 2023*. Sustainable Development Solutions Network.
- Indah, S., & Magriasti, L. (2023). Peran indeks kebahagiaan (human happiness index) dalam perumusan kebijakan publik di Indonesia. *Jurnal Pendidikan Tambusai*, 7(3), 30001-30012. <https://doi.org/10.31004/jptam.v7i3.11842>
- Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis* (6th ed.). Pearson.
- Kumalasari, D. A., & Yasa, I. G. W. M. (2020). Pengaruh faktor-faktor yang mempengaruhi tingkat kebahagiaan negara di dunia. *E-Jurnal Ekonomi Pembangunan Universitas Udayana*, 9(5), 963-992.

- Lovakov, A., & Agadullina, E. R. (2021). Empirically derived guidelines for effect size interpretation in social psychology. *European Journal of Social Psychology, 51*(3), 485–504. <https://doi.org/10.1002/ejsp.2752>
- OECD. (2013). *Guidelines on measuring subjective well-being*. OECD Publishing.
- Pillai, S. E., & Hu, W. (2024). Privacy-preserving gross domestic product (GDP) calculation using Paillier encryption and differential privacy. In *Proceedings of the 2024 ACM Southeast Conference* (pp. 182–187). <https://doi.org/10.1145/3603287.3651188>
- Puritz, C., Ness-Cohn, E., & Braun, R. (2023). fasano.franceschini.test: An implementation of a multivariate KS test in R. *The R Journal, 15*(3), 159–171. <https://doi.org/10.32614/RJ-2023-067>
- Salmeron-Gomez, R., Garcia-Garcia, C. B., & Garcia-Perez, J. (2025). A redefined variance inflation factor: Overcoming the limitations of the variance inflation factor. *Computational Economics, 65*(1), 337–363. <https://doi.org/10.1007/s10614-024-10575-8>
- Sherry, A., & Henson, R. K. (2005). Canonical correlation analysis: Use of composite variables for investigating relationships. *Psychological Methods, 10*(3), 381–391. <https://doi.org/10.1037/1082-989X.10.3.381>
- Tran, T. P. (2020). Ho Chi Minh thought on solidarity and the social sustainable development in Vietnam. *E3S Web of Conferences, 164*, Article 12004. <https://doi.org/10.1051/e3sconf/202016412004>