PERINATAL JOURNAL

Original Article

Perinatal Journal 2026; 34(1):851-864

https://doi.org/10.57239/prn.26.03410087

Determinants of graduate leadership ability of university students in Zheng Zhou city: The mediating role of Self-Efficacy in higher education

Ziqian Peng[1,2]*, Ali Khatibi[2], Jacquline Tham[2]

[1]Shangqiu Medical College, No.666, Yingbin Avenue, Shangqiu City, Henan Province, China
[2]Postgraduate Centre, Management and Science University, University Drive, Off Persiaran Olahraga, Section 13, 40100 Shah Alam, Malaysia

Abstract

The results of the comprehensive data analysis showed that Personal preferences, interpersonal relationships, university programs, community atmosphere, policy formulation all had significant effects on graduate leadership ability and the corresponding hypotheses were valid, and the self-efficacy mediating effect also played a significant mediating role and the hypotheses were valid. According to the results of the research, based on the micro and macro based on the socio-ecological model, this was verified that the environment in which a graduate is placed can greatly influence the leadership development. It is recommended that a leadership curriculum be added to the development of the atmosphere for leadership development in the university. It is recommended that the community can increase the diversity of activities so that graduate can be more active and get exercise. Graduate students need to develop leadership skills in a variety of ways. The sample of this research was limited to 20 colleges and universities in Zhengzhou City, which is a relatively small sample size and may not adequately represent the actual situation of college graduates nationwide. Therefore, the general applicability and extensiveness of the findings need to be verified through further research.

Keywords: Social ecology, Self-Efficacy, Leadership, Graduates

1.Introduction

The main objective of this research is to explore and analyse in depth which key factors play a decisive role in the formation and development of leadership under the role of social-ecology in a group of graduates in Zhengzhou City, Henan Province, China. In this chapter, we will provide a detailed explanation of the quantitative analysis section and present the results visually in the form of graphs and charts. To ensure the accuracy and reliability of the findings, the authors collected data from a broadly representative sample. By analysis these data, the aim is to validate the study to verify a series of hypotheses.

The first part of the quantitative study focuses on the presentation of frequency analysis, measures of central tendency (e.g., mean, median, and plurality), and measures of dispersion (e.g., standard deviation and variance), which are tools that allow for a better understanding of the distribution and characteristics of the data. And in the second half of the quantitative study, the authors use inferential statistics, including confidence interval estimation and hypothesis testing, to estimate statistical parameters for the population as a whole and to test and judge the hypotheses proposed in the current research.

2. Sample characteristics

Participants in this study included graduates from 20 different universities in Zhengzhou City, Henan Province. These groups of graduates have several notable characteristics: firstly, the parties have just completed their university studies and are starting to enter society, and are therefore at a critical turning point where The parties need to develop and enhance their leadership skills. Secondly, these graduates have been able to clearly realise that leadership plays a crucial role and is an indispensable skill in their future career development and personal growth. After obtaining the consent and support of all final year counsellors, we explained in detail the purpose and significance of the questionnaire survey to the counsellors and graduates, and at the same time emphasized the voluntary nature of participation and the confidentiality of all data.

After a period of questionnaire collection, we received a total of 588 questionnaires, and after strict screening and auditing, we finally identified 588 questionnaires as valid questionnaires. Through in-

Correspondence: Ziqian Peng, Shangqiu Medical College,No.666, Yingbin Avenue, Shangqiu City, Henan Province, China, **e-mail:** 15236821999@139.com,
Received: November 30, 2025 **Accepted:** January 8, 2026

depth analysis of the data from these valid questionnaires, the researchers found that the group of graduates participating in this study showed great diversity. In order to better understand the characteristics of this group, the following subsections will detail the descriptive statistics of the 588 graduates in this study, which will help us to more accurately grasp the current state of leadership among the graduates as well as their needs and expectations for leadership development.

2.1 demographic information

The following paragraphs report demographic information on the 588 valid samples collected by the survey.

2.2.1 Gender

Table 1: Gender statistics

Gender	Frequency	Percentage(%)
Male	286	48.6
Female	302	51.4
Total	588	100

In this particular research study, the total number of male participants who took part in the survey was 286, which constituted 48.6 percent of the entire participant pool. On the other hand, the total number of female participants was 302, representing 51.4 percent of the entire group. When comparing the two figures, this becomes evident that the proportion of female participants is marginally higher than that of their male counterparts. However, the disparity between the two groups is not substantial enough to be considered statistically significant. The gender distribution observed in this survey exhibits a relatively balanced characteristic, which is noteworthy. This balanced gender distribution serves as a foundational backdrop for additional data analyses that may be conducted using this sample, this provides valuable context that can aid in a more comprehensive understanding and interpretation of the results that emerge from these analyses.

Moreover, the gender balance observed in this study is particularly important when considering the potential impact on the research outcomes. This is well-documented that gender differences can influence responses to various stimuli and questions

2.2.2 Graduation year

Table 2: Graduation year statistics

Graduation year	Frequency	Percentage(%)
2023	125	21.3
2024	264	44.9
2025	199	33.8
Total	588	100%

Among the students who successfully completed their academic pursuits and received their diplomas in the year 2024, the frequency of their representation within the study sample reached an impressive count of 264 individuals. This number constitutes a substantial 44.9% of the entire group of participants, marking this as the highest proportion observed across all the different graduation years included in the analysis. In contrast, the frequency of students who graduated in the preceding year of 2023 was notably lower, standing at 21.3%. Similarly, the cohort of students who graduated in the subsequent year of 2025 exhibited a frequency of 33.8%. Table 4.5 presents statistics on graduation years, covering three years: 2023, 2024, and 2025, with a total sample size of 588.

2.2.3 Zhengzhou university graduates

Table 3: Zhengzhou university graduate's statistics

Zhengzhou university graduates	Frequency	Percentage (%)
Yes	588	100%
No	0	0%
Total	588	100%

Upon conducting a comprehensive and meticulous examination of the data that has been presented in the table mentioned earlier, this becomes abundantly clear that within the 588 valid samples that have been analyzed, each and every one of them represents a graduate who has completed their educational journey at institutions located within the city of Zhengzhou. There are no exceptions or instances of graduates from universities situated outside of Zhengzhou within this particular dataset. This specific aspect serves to highlight and underscore the fact that the scope and focus of this survey are exclusively concentrated on the demographic of individuals who have completed their educational

pursuits at universities based in Zhengzhou. As a result, the findings and insights that have been derived from this study can be considered as a true and accurate reflection of the current state of affairs pertaining to the leadership capabilities, experiences, and professional achievements of graduates who hail from the academic institutions located in Zhengzhou.

Additionally, the homogeneity in the sample set regarding the graduation institution eliminates potential confounding factors that could arise from comparing graduates from diverse educational backgrounds. This controlled environment facilitates a clearer and more accurate assessment of the impact of Zhengzhou-based universities on shaping the leadership capabilities of their students.

2.2.4 Political profile

Table 4: Political profile statistics

Political profile	Frequency	Percentage(%)
Communist party member	87	14.8
Communist Youth League member	225	38.3
Masses	276	46.9
Total	588	100%

From the data situation, we know that the Communist Party members: frequency 87, accounting for 14.8 percent. Communist Party members are often advanced members who have been cultivated and examined by the organisation in colleges and universities. In terms of leadership training, the Party organisation may have provided opportunities for theoretical learning and practical exercises, etc. The parties may be able to play an exemplary role in team leading and responsibility taking among the graduates' group, and The parties may have a stronger sense of organisational co-ordination and leadership.

Members of the Communist Youth League: frequency 225, accounting for 38.3%. The Communist Youth League is the assistant and reserve army of the Party. Members of the Communist Youth League receive education and guidance from the organisation, and in participating in various campus activities, The parties have the opportunity to exercise their communication and collaboration skills, which is helpful for the initial development of leadership.

Mass frequency 276, accounting for 46.9%, the highest proportion. Although graduates from the masses have not joined party organisations, the parties may develop leadership through other campus activities and social practice, but The parties may be relatively less likely to obtain systematic leadership training resources.

2.2.5 Served as a student leader

Table 5: Served as some student leader statistics

Served as a student leader	Frequency	Percentage(%)
Yes	391	66.5
No	197	33.5
Total	588	100%

According to the table, the number of graduates who have served as student leaders is 391, accounting for 66.5%. During the period of serving as a student leader, graduates have more opportunities to organise activities, co-ordinate interpersonal relationships, and handle affairs, etc. These experiences can effectively exercise leadership-related skills such as communication, organization and decision-making. In their subsequent career development or social activities, the parties may be able to play a better leadership role with their accumulated experience.

Not having served as a student leader the frequency is 197, accounting for 33.5%. Although this group of graduates lacks the direct training brought by the experience of student cadres, the parties may improve their leadership skills by participating in competitions, club activities, social practice, etc., but the way and focus of the training may be different from that of the students who have served as student cadres, which indicates that there are diversified paths for the development of leadership skills.

3 Descriptive statistics

Table 4.9 provides descriptive statistical analyses of participants' attitudes and perceptions on different dimensions of a research. The sample size was 588 and covered seven different dimensions: personal preferences, interpersonal relationship, university

programme, community atmosphere, policy formulation, self-efficacy, graduate leadership ability.

Table 6: Descriptive statistics

Variables	N	Minimum	Maximum	Mean	Standard deviation
PP	588	1.33	5	3.35	0.939
IR	588	1.33	5	3.31	0.969
UP	588	1.14	5	3.36	0.969
CA	588	1.20	5	3.32	0.973
PF	588	1.38	5	3.31	0.991
SE	588	1.14	5	3.35	0.975
GLA	588	1.50	5	3.31	0.990

The sample size for each variable under consideration in this research has been determined to be 588, which suggests that the data collection process has been quite thorough and representative.

For the variable PP, which represents personal preference, the sample size (N) is 588. The minimum value is 1.33, and the maximum is 5. This indicates that the scores for personal preference fall within a range from a relatively low value of 1.33 to a maximum of 5. The mean value is 3.35, suggesting that on average, respondents' personal preference scores lean towards the middle - higher end of the scale. The standard deviation is 0.939, which implies a moderate spread of data points around the mean. A standard deviation of this magnitude indicates that while there is some variability in personal preference scores,this is not overly dispersed.

The variable IR, standing for interpersonal relationship, also has a sample size of 588. The minimum and maximum values are 1.33 and 5 respectively, the same range as PP in this regard. The mean of 3.31 shows that, on average, the scores for interpersonal relationship are slightly lower than that of PP but still in the middle range. With a standard deviation of 0.969, this has a bit more variability compared to PP. This suggests that there is a relatively wider spread of scores for interpersonal relationship among the respondents, indicating more diversity in people's perceptions or experiences related to interpersonal relationships.

Regarding UP, which represents the university program variable, the sample size remains 588. The minimum value is 1.14, and the maximum is 5. The mean score is 3.36, which is slightly higher than the means of some other variables like IR and PF. This indicates that, on average, respondents rate the university program relatively favorably. The standard deviation of 0.969 is close to those of other variables, signifying a comparable level of variability in responses. So, while there is some dispersion in how people evaluate the university program,this is within a similar range as other factors.

The variable CA, labeled as community atmosphere, has a sample size of 588. The minimum value is 1.20, and the maximum is 5. The mean value of 3.32 shows that, on average, the scores related to this community - related variable are in the middle range. The standard deviation is 0.973, which is in line with the variability seen in other variables. This suggests that there is a moderate spread of opinions or experiences regarding the community - related aspect represented by CA among the respondents.

For PF, which stands for policy formulation, the sample size is 588. The minimum value is 1.38, and the maximum is 5. The mean score is 3.31, similar to that of IR. This implies that, on average, respondents' evaluations of policy formulation are at a comparable level to interpersonal relationship scores. The standard deviation of 0.991 is relatively high among these variables, indicating a greater spread of scores. This means there is more diversity in people's views on policy formulation, with some respondents having significantly different opinions compared to others.

The variable SE, representing self - efficacy, has a sample size of 588. The minimum value is 1.14, and the maximum is 5. The mean value of 3.35 is the same as that of PP, showing that, on average, respondents' self - efficacy scores are at a similar level to personal preference scores. The standard deviation of 0.975 indicates a moderate amount of variability in self -

efficacy scores. This means that while there is some difference in how individuals perceive their self - efficacy, this is within a reasonable range compared to other variables.

The variable GLA, denoting graduate leadership ability, has a sample size of 588. The minimum value is 1.50, and the maximum is 5. The mean value of 3.31 is slightly lower than some other variables. This suggests that, on average, respondents rate graduate leadership ability somewhat less favorably compared to variables like UP and PP. The standard deviation of 0.990 is relatively high, indicating a relatively wide spread of scores. This shows that there is a significant amount of diversity in how people assess graduate leadership ability, with a notable range of opinions among the respondents.

In summary, the data collected and analyzed in this study provide a foundational understanding of the leadership development of graduates in Zhengzhou City. These insights can serve as a valuable starting point for further research and can be integrated with other factors to conduct a more comprehensive analysis of the influencing factors and potential strategies for enhancing the leadership capabilities of college students.

4 Dimensions of the variables

4.1 Exploratory factor analysis

Exploratory Factor Analysis (EFA) is a multivariate statistical method that has become a fundamental tool in the development and validation of psychological theories and measures. Researchers must make a number of thoughtful and evidence-based methodological decisions when conducting EFA, and at each decision point there are multiple options, each with their own advantages and disadvantages (Watkins, 2018; Jam et al., 2025).

Data were analysed for rigorous reliability and validity using SPSS software. The initial internal consistency alpha coefficient exceeded the threshold of 0.7 and the entries all had high reliability. The (KMO) measure was close to 0.9, confirming the validity of the questionnaire (Zhang et al.,2024).

The Bartlett's test of sphericity value was significant (p = 0.000) and the factor analysis was appropriate (Ibikunle & Smith,2021).

The p-value of the Bartlett's spherical test was <0.001, indicating that the strength of the correlation between the items was good enough for EFA.All factor loadings in the one-way model were >0.3.The factor loadings indicated that the infrastructure could be considered as one-way domains in the developmental study.The KMO <0.8 verified the sample adequacy, and the p-value of Bartlett's test was <0.001, indicating that the strength of the relationship between the items was good enough to allow extended analyses (EFA) on this data. indicating that the strength of the relationship between the items was good enough to allow for extended analyses (EFA) to be conducted on this data (Dar et al., 2021).

Table 7: KMO and Bartlett's test results

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.965
Bartlett's Test of Sphericity	Approx. Chi-Square	18314.388
	df	1081
	Sig.	.000

Kaiser - Meyer - Olkin sampling fitness measure (KMO value) is 0.965, KMO value is between 0 - 1, the closer to 1, the stronger the correlation between the variables, the more suitable the data for factor analysis, usually 0.9 and above means sampling adequacy is 'very good', the KMO value of this data is 0.965, indicating a high degree of correlation between the variables, a good basis for factor analysis. 0.965, indicating a high degree of correlation between the variables and a good basis for factor analysis.In the Bartlett's spherical test, the approximate chi-square (Approx. Chi - Square) is as high as 18314.388, which is free. is as high as 18314.388, the degree of freedom (df) is 1081, and the significance (Sig.) is 0.000. When the significance is less than 0.05, it means that the data is not a unit matrix, i.e., there is a significant correlation between the variables, and analyses that rely on the assumption of correlation of variables, such as factor analysis, can be carried out. In summary, from these two key test indicators, the data are excellent in terms of fitness and are very suitable for multivariate statistical operations such as factor

analysis, which can provide a reliable data basis for subsequent analyses such as in-depth mining of the potential structure of variables and extraction of public factors.

In summary, both the high KMO value and the significant result of Bartlett's test of sphericity strongly support the suitability of the data for factor analysis.

5 Structural Equation Modelling (SEM)

5.1 Confirmatory Factor Analysis (CFA) via individual measurement model

In practice, if the model is designed to be too complex, or if the available sample size is insufficient to support the complexity of the model,this may result in a poor model fit, which in turn prevents accurate and valid parameter estimates from being obtained. Further if the number of questions included in the model is too small and the sample size is relatively large, the high degree of correlation that may exist between the variables in the data may cause the standard errors of the parameter estimates to become too large, thus affecting the reliability of the statistical tests. In some cases, parameter estimates may not be derived, or the model fit indicators may be unusually good, often suggesting that the model may have over fitted the data, i.e., the model is too finely tuned to the current data set and loses its ability to generalise to new data.

Therefore, crucial to carefully consider the complexity of the SEM model and ensure that this aligns with the available sample size. This involves striking a balance between a model that is sufficiently complex to capture the underlying relationships in the data, and one that is simple enough to avoid over fitting. Additionally, attention should be paid to the number of questions included in the model, as this can impact the correlation between variables and the reliability of the parameter estimates. Overall, a thoughtful approach to model design and validation is essential for ensuring the accuracy and validity of the results obtained from SEM analysis.

AGFI ≥ 0.85 indicates an acceptable fit, which means that the scale maintains validity and is free of unnecessary complexity, making this more suitable for practical assessment. RMSEA ≤ 0.08 indicates an acceptable fit, with lower RMSEA values indicating minimal unexplained variance, which demonstrates that the scale provides a stable and generalisable measure (Wang et al., 2025).

With regard to the validity of the tool, the following criteria should be met: NFI > 0.8, CFI > 0.8, GFI > 0.8, CMIN/DF < 3 (Li et al.,2024).

SRMR value < 0.05; RMSEA value < 0.08 proved good model fit; GFI and AGFI values > 0.80. NFI, IFI, TLI and CFI values ≥ 0.90, and CMIN/DF values between 2.0 and 5.0. Convergent validity was evaluated using Composite Reliability (CR) and Average Variance Extracted (AVE), and CR ≥ 0.70 and AVE ≥ 0.50 indicated acceptable convergent validity (Ding et al,2022).

5.1.1 Confirmatory factor analysis of personal preference

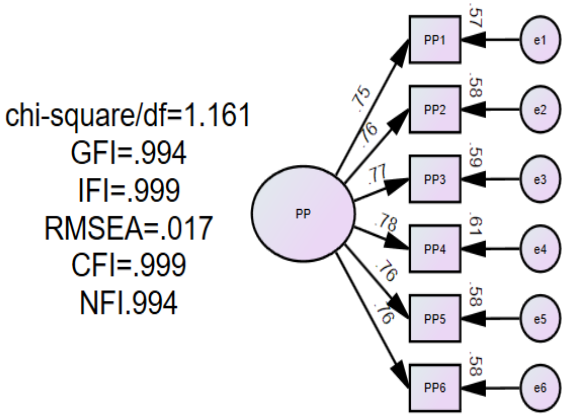

Figure 1: Personal preferences measurement model

Table 8: Personal preferences model fitting test

Test index	Fitting standard	Model fit index	Result
Chi-square/df	<3	1.161	Achieved
RMSEA	<0.10	0.017	Achieved
CFI	>0.90	0.999	Achieved
NFI	>0.80	0.994	Achieved
GFI	>0.80	0.994	Achieved
IFI	>0.90	0.999	Achieved

The figure presents data for several fit metrics for the PP measurement model. The chi-square degrees of

freedom ratio is 1.161, which is below the fit criterion of less than 3 and meets the standard. The RMSEA is 0.017, which is less than 0.10 according to the standard and meets the standard. The CFI, NFI, GFI, and IFI are 0.999, 0.994, 0.994, 0.999, respectively, which are all higher than their respective fitting standards (CFI, IFI needs to be greater than 0.90; NFI, GFI needs to be greater than 0.80), and all of them are in compliance. Overall, most of the indicators show that the model fits well.

Table 9: Personal preferences convergent validity test table

Name	Factor loading	Measurement error	AVE	CR
PP1	0.755	-	0.59	0.88
PP2	0.761	0.055		
PP3	0.77	0.056		
PP4	0.782	0.056		
PP5	0.762	0.056		
PP6	0.764	0.055		

This table presents the results of a convergent validity test for the "PP" construct. The factor loadings for all items (PP1 - PP6) range from 0.755 to 0.782, which are relatively high. Generally, factor loadings above 0.7 are considered strong, indicating that each item is strongly related to the underlying construct.

The measurement error values for PP2, PP3, PP4, PP5, and PP6 are all around 0.055 - 0.056, which are relatively low, suggesting that the measurement of these items is precise.

The Average Variance Extracted (AVE) is 0.59, which exceeds the recommended threshold of 0.5, indicating that the construct explains more variance in its items than the error variance. The Composite Reliability (CR) is 0.88, well above the commonly - used threshold of 0.7, which further confirms the internal consistency and reliability of the "PP" construct. Overall, these results provide strong evidence for the convergent validity of the "PP" construct.

5.1.2 Confirmatory factor analysis of interpersonal relationship

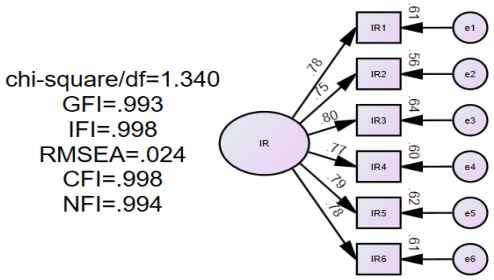

Figure 2: Interpersonal relationships measurement model

Table 10: Interpersonal relationships model fitting test

Test index	Fitting standard	Model fit index	Result
Chi-square/df	<3	1.34	Achieved
RMSEA	<0.10	0.024	Achieved
CFI	>0.90	0.998	Achieved
NFI	>0.80	0.994	Achieved
GFI	>0.80	0.993	Achieved
IFI	>0.90	0.998	Achieved

This is the result of the IR model fit test. The ratio of chi-square degrees of freedom is 1.340, which is less than the criterion of 3, indicating that the model is well fitted on this index; the RMSEA is 0.024, which is much lower than the criterion of less than 0.10, showing that the model error is small and the degree of fit is high; the CFI, NFI, CFI, NF, GFI, IFI are 0.998, 0.994, 0.993, 0.998, respectively, which are all higher than their respective standards of greater than 0.90 (CFI, IFI) or greater than 0.80 (NFI, GFI). The results of all indicators are 'Achieved', indicating that the IR model meets the good standard in these key fitting indicators, and the overall fitting effect is excellent.

Table 11: Interpersonal relationships convergent validity test table

Name	Factor loading	Measurement error	AVE	CR
IR1	0.778	-	0.61	0.9
IR2	0.746	0.05		
IR3	0.798	0.051		
IR4	0.773	0.051		
IR5	0.791	0.052		
IR6	0.78	0.051		

This table shows the results of a convergent validity test for the "IR" construct. The factor loadings for items IR1 - IR6 range from 0.746 to 0.798, which are quite high. Typically, factor loadings above 0.7 signify a strong relationship between each item and the underlying construct, suggesting that these items effectively measure what the "IR" construct intends to capture. The measurement error values for IR2, IR3, IR4, IR5, and IR6 are all in the range of 0.05 - 0.052, indicating relatively low measurement errors and thus precise measurement of these items. The Average Variance Extracted (AVE) is 0.61, surpassing the recommended threshold of 0.5. This implies that the "IR" construct explains more variance in its items than the error variance. The Composite Reliability (CR) is 0.90, well above the commonly - used benchmark of 0.7, further validating the internal consistency and reliability of the "IR" construct. Overall, these findings strongly support the convergent validity of the "IR" construct.

5.1.3 Confirmatory factor analysis of university programmer

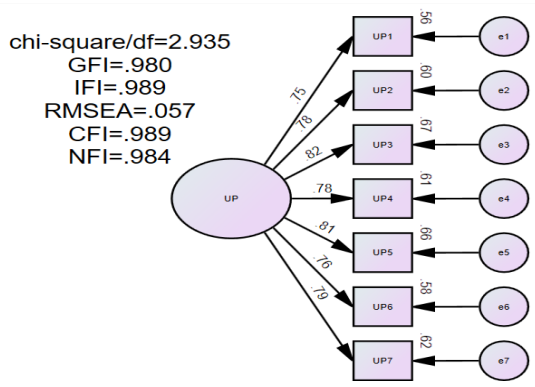

Figure 3：University programmer measurement model

Table 12: University programme model fitting test

Test index	Fitting standard	Model fit index	Result
Chi-square/df	<3	2.935	Achieved
RMSEA	<0.10	0.057	Achieved
CFI	>0.90	0.989	Achieved
NFI	>0.80	0.984	Achieved
GFI	>0.80	0.98	Achieved
IFI	>0.90	0.989	Achieved

This table presents the results of a model fitting test for the "UP" model. Several key fit indices are evaluated against their respective fitting standards. The Chi - square/df ratio is 2.935, which meets the standard of being less than 3, indicating a reasonable model - data fit in terms of the ratio of the chi - square statistic to degrees of freedom. The RMSEA is 0.057, well below the threshold of 0.10, suggesting a good approximation of the population model. The CFI and IFI both have values of 0.989, exceeding the standard of 0.90, which implies that the proposed model fits the data significantly better than a baseline model. The NFI is 0.984 and the GFI is 0.980, both surpassing their respective standards of 0.80. Overall, all the evaluated fit indices meet or exceed their recommended thresholds, strongly indicating that the "UP" model provides an excellent fit to the data.

Table 13: University programme convergent validity test table

Name	Factor loading	Measurement error	AVE	CR
UP1	0.747	-	0.61	0.92
UP2	0.778	0.054		
UP3	0.817	0.057		
UP4	0.781	0.055		
UP5	0.811	0.056		
UP6	0.763	0.055		
UP7	0.79	0.055		

This table shows the results of a convergent validity test for the "UP" construct. The factor loadings for items UP1 - UP7 range from 0.747 to 0.817. Generally, factor loadings above 0.7 indicate a strong relationship between each item and the underlying construct, suggesting that these items effectively measure the "UP" construct. The measurement error values for UP2, UP3, UP4, UP5, UP6, and UP7 are in the range of 0.054 - 0.057, which are relatively low, signifying precise measurement of these items. The Average Variance Extracted (AVE) is 0.61, exceeding the commonly - recommended threshold of 0.5. This means that the "UP" construct explains more variance in its items than the error variance. The Composite Reliability (CR) is 0.92, well above the typical benchmark of 0.7, further confirming the internal consistency and reliability of the "UP" construct. Overall, these results strongly support the convergent validity of the "UP" construct.

5.1.4 Confirmatory factor analysis of community atmosphere

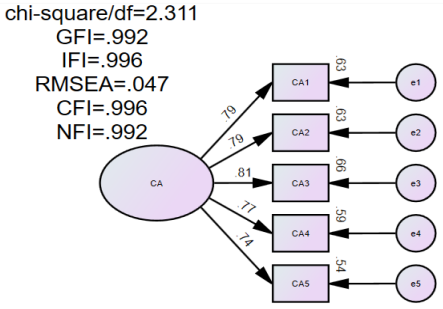

Figure 4: Community atmosphere measurement model

Table 14: Community atmosphere model fitting test

Test index	Fitting standard	Model fit index	Result
Chi-square/df	<3	2.311	Achieved
RMSEA	<0.10	0.047	Achieved
CFI	>0.90	0.996	Achieved
NFI	>0.80	0.992	Achieved
GFI	>0.80	0.992	Achieved
IFI	>0.90	0.996	Achieved

This table presents the results of a model fitting test for the "CA" model. The evaluation is based on several key fit indices compared to their respective fitting standards. The Chi - square/df ratio is 2.31, which is less than the recommended 3, indicating a reasonable fit between the model and the data in terms of the relationship between the chi - square statistic and degrees of freedom. The RMSEA is 0.047, well below the threshold of 0.10, suggesting a very good approximation of the population model. The CFI and IFI both have extremely high values of 0.996, far exceeding the standard of 0.90, which implies that the proposed model fits the data much better than a baseline model. The NFI and GFI are both 0.992, surpassing the 0.80 standard. Overall, all the fit indices meet or exceed their respective thresholds, providing strong evidence that the "CA" model has an excellent fit to the data.

This is the CA Convergent Validity Test table involving five variables CA1 - CA5. Factor loading (Factor loading) is between 0.74 - 0.81, indicating that the variables are more closely related to the latent variables. Measurement error (Measurement error) ranged from 0.049 - 0.051, which is a small value, reflecting the high precision of the measurement.

Table15: Community atmosphere convergent validity test table

Name	Factor loading	Measurement error	AVE	CR
CA1	0.791	-	0.6090	0.8860
CA2	0.792	0.050		
CA3	0.813	0.050		
CA4	0.765	0.051		
CA5	0.738	0.049		

The Average Extracted Variance (AVE) was 0.6090, which is greater than the standard of 0.5, indicating good model aggregation validity. The combined reliability (CR) is 0.8860, which is higher than the criterion of 0.7, implying that the model reliability is high. Overall, the model performs well in terms of convergent validity and reliability, and the variables can effectively reflect the characteristics of latent variables.

5.1.5 Confirmatory factor analysis of policy formulation

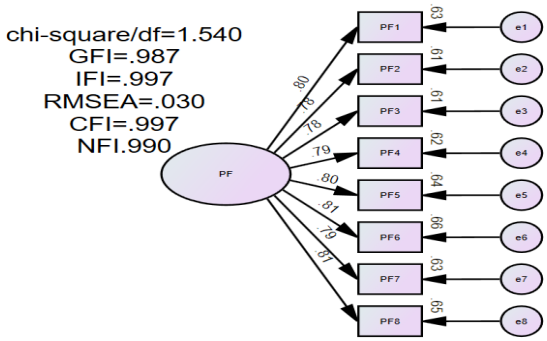

Figure 5：Policy formulation measurement model

Table 16: Policy formulation model fitting test

Test index	Fitting standard	Model fit index	Result
Chi-square/df	<3	1.54	Achieved
RMSEA	<0.10	0.03	Achieved
CFI	>0.90	0.997	Achieved
NFI	>0.80	0.99	Achieved
GFI	>0.80	0.987	Achieved
IFI	>0.90	0.997	Achieved

This is the result of the fitting test of the PF measurement model. The ratio of chi-square degrees of freedom (Chi - square/df) is 1.540, which is less than the criterion of 3, indicating that the model fits well; the root mean square error approximation RMSEA is 0.030, which is much lower than the criterion of less than 0.10, indicating that the model error is relatively small; the CFI is 0.997, the NFI is 0.990, the GFI is 0.987, and the IFI is 0.997, IFI should be greater than 0.90, NFI, GFI should be greater than 0.80. The results of all the indicators are 'Achieved', indicating that the PF model meets the good standard in these key fitting indicators.

Table 17: Policy formulation convergent validity test table

Name	Factor loading	Measurement error	AVE	CR
PF1	0.80	-	0.6322	0.9322
PF2	0.78	0.046		
PF3	0.78	0.046		
PF4	0.79	0.045		
PF5	0.80	0.046		
PF6	0.81	0.045		
PF7	0.79	0.047		
PF8	0.81	0.046		

This is the PF convergent validity test scale involving eight variables from PF1 - PF8. Factor loading (Factor loading) is between 0.78 - 0.81, indicating that the variables are strongly associated with the latent variables. Measurement error (Measurement error) ranged from 0.045 - 0.047, which is a small value, indicating a high degree of measurement precision.

The average extracted variance value (AVE) was 0.6322, which is greater than the standard of 0.5, indicating good convergent validity. The combined reliability (CR) is 0.9322 , which is higher than the criterion of 0.7, indicating high model reliability. Overall, the model performed well in terms of both convergent validity and reliability.

5.1.6 Confirmatory factor analysis of self-efficacy

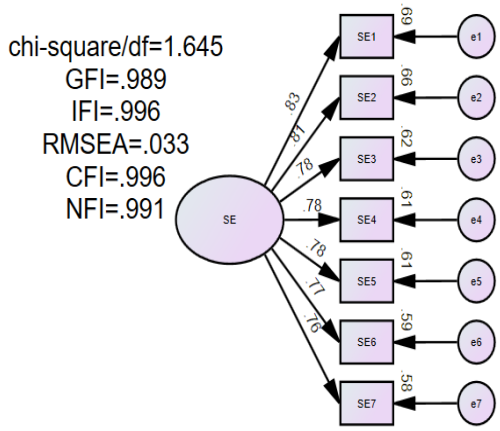

Figure 6：Self-efficacy measurement model

Table 18: Self-efficacy model fitting test

Test index	Fitting standard	Model fit index	Result
Chi-square/df	<3	1.645	Achieved
RMSEA	<0.10	0.33	Achieved
CFI	>0.90	0.996	Achieved
NFI	>0.80	0.991	Achieved
GFI	>0.80	0.989	Achieved
IFI	>0.90	0.996	Achieved

This table shows the results of a model fitting test for the "SE" model. The Chi - square/df ratio is 1.645, which is well below the fitting standard of less than 3, indicating a good fit between the model and the data in terms of the relationship between the chi - square statistic and degrees of freedom. The RMSEA is 0.33, which, although this meets the standard of being less than 0.10, is relatively high compared to typical good

- fitting models, suggesting some room for improvement in the approximation of the population model. However, the CFI, IFI, NFI, and GFI are all quite high, with values of 0.996, 0.996, 0.991, and 0.989 respectively, all exceeding their respective fitting standards (> 0.90 for CFI and IFI, > 0.80 for NFI and GFI). Overall, while the RMSEA shows a slight concern, the other indices indicate that the "SE" model generally provides a good fit to the data.

Table19: Self-efficacy convergent validity test table

Name	Factor loading	Measurement error	AVE	CR
SE1	0.831	-	0.629	0.932
SE2	0.814	0.041		
SE3	0.784	0.043		
SE4	0.781	0.042		
SE5	0.78	0.042		
SE6	0.77	0.043		
SE7	0.761	0.042		

This table presents the results of a convergent validity test for the "SE" construct. The factor loadings for items SE1 - SE7 range from 0.761 to 0.831. Factor loadings above 0.7 generally signify a strong relationship between each item and the underlying construct, indicating that these items effectively measure the "SE" construct. The measurement error values for SE2, SE3, SE4, SE5, SE6, and SE7 are in the range of 0.041 - 0.043, which are relatively low, suggesting precise measurement of these items. The Average Variance Extracted (AVE) is 0.629, exceeding the commonly - recommended threshold of 0.5. This means that the "SE" construct explains more variance in its items than the error variance. The Composite Reliability (CR) is 0.932, well above the typical benchmark of 0.7, further confirming the internal consistency and reliability of the "SE" construct. Overall, these results strongly support the convergent validity of the "SE" construct.

5.1.7 Confirmatory factor graduate leadership ability

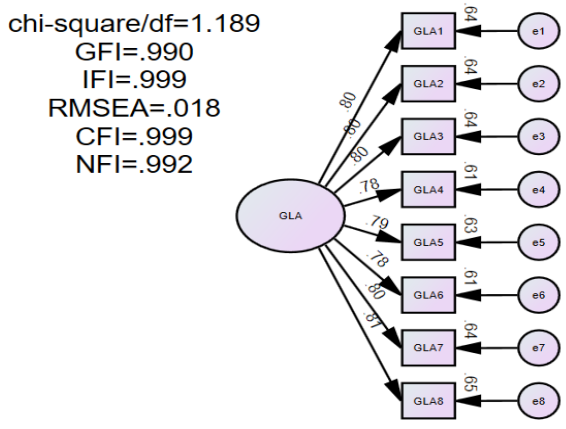

Figure 7: Graduate leadership ability measurement model

Table 20: Graduate leadership ability model fitting test

Test index	Fitting standard	Model fit index	Result
Chi-square/df	<3	1.189	Achieved
RMSEA	<0.10	0.018	Achieved
CFI	>0.90	0.999	Achieved
NFI	>0.80	0.992	Achieved
GFI	>0.80	0.99	Achieved
IFI	>0.90	0.999	Achieved

This is the result of the fit test of the GLA model. The Chi - square/df is 1.189, which is less than the criterion of 3, showing that the model fits well; the RMSEA is 0.018, which is much lower than the criterion of less than 0.10, indicating that the model error is extremely small; the CFI, IFI are 0.999, the NFI is 0.992, and the GFI is 0.990, which are all significantly higher than their respective corresponding fit criteria CFI, GFI, and GLA. Comparative Fit Index CFI, IFI are 0.999, NFI is 0.992, and GFI is 0.990, which are all significantly higher than CFI, IFI need to be greater than 0.90, NFI, GFI need to be greater than 0.80. The results of all indicators are 'Achieved', indicating that the GLA model meets the good standard in these key fitting indicators, the overall fitting effect is excellent, and the model fits the data well.

Table 21: Graduate leadership ability convergent validity test table

Name	Factor loading	Measurement error	AVE	CR
GLA1	0.800	-	0.630	0.948
GLA2	0.802	0.045		
GLA3	0.803	0.044		
GLA4	0.783	0.043		
GLA5	0.793	0.044		
GLA6	0.782	0.044		
GLA7	0.802	0.043		
GLA8	0.805	0.045		

This table shows the results of a convergent validity test for the "GLA" construct. The factor loadings for items GLA1 - GLA8 range from 0.782 to 0.805. Since factor loadings above 0.7 generally indicate a strong connection between each item and the underlying construct, these values suggest that the items effectively measure the "GLA" construct. The measurement error values for GLA2 - GLA8 are in the

range of 0.043 - 0.045, which are relatively low, signifying precise measurement of these items. The Average Variance Extracted (AVE) is 0.630, surpassing the commonly - used threshold of 0.5. This means that the "GLA" construct explains more variance in its items than the error variance. The Composite Reliability (CR) is 0.948, well above the typical benchmark of 0.7, further validating the internal consistency and reliability of the "GLA" construct. Overall, these findings strongly support the convergent validity of the "GLA" construct.

6 Conclusion

The research begins by reviewing the importance of leadership, pointing out that in the context of globalisation, people with leadership skills are crucial for the development of organisations and societies. The study also looks at the motivation of social sciences and liberal arts students to engage in student leadership and the students' deficiencies in leadership skills. These deficiencies are manifested in their lack of leadership behaviours to bring exemplary impact to the organisation. Therefore, the development of student-centred learning and the leadership behaviours of student leaders can be mutually reinforcing, and the use of 'student-centred learning' strategies can enhance students' leadership behaviours.Based on the above background, this study applies social ecology theory to guide the enhancement of leadership skills of college graduates. Social ecology theory emphasises the interaction between the individual and the environment, and suggests that the development of the individual is influenced by multiple levels of environmental factors. In this study, we explored the effects of personal reasons, interpersonal factors, college environment, community environment, and policies on students leadership enhancement and how these factors act on students self-efficacy.By analysing the data, we found that the social ecology theory has a significant role in explaining and enhancing the leadership competence of college graduates, while the theory also positively affects students' self-efficacy.

Future research can further expand the sample range to improve the representativeness and accuracy of the data. Also, a variety of research methods, such as interviews and observations, can be used to obtain more comprehensive and in-depth data. In addition, future research can also explore other factors that may affect leadership enhancement, such as personal traits and cultural background, in order to further improve the theoretical system and practical strategies of leadership development.

In conclusion, this research used social ecology theory to guide the enhancement of leadership competence of college graduates and achieved certain research results. However, due to the limitations and shortcomings in sample selection, data collection, etc., future research still needs to be further deepened and improved. We believe that as the study continues to deepen, we will be able to better understand the nature and laws of leadership development and provide more effective strategies and methods for the cultivation of talents in colleges and universities.

References

Arnall, J.R., Johnson, P., Lee, J.J., Linder, M.A., Lund, N.L., & Satpathy, S. (2014). Student Perspectives on Student Leadership Development Programs. American Journal of Pharmaceutical Education, 78.

Baharum, H., Ismail, A., Awang, Z., McKenna, L., Ibrahim, R., Mohamed, Z., & Hassan, N. H. (2023). Validating an instrument for measuring newly graduated nurses' adaptation. International journal of environmental research and public health, 20(4), 2860.

Barella, Y., Fergina, A., Mustami, M. K., Rahman, U., & Alajaili, H. M. A. (2024). Quantitative methods in scientific research. Jurnal Pendidikan Sosiologi Dan Humaniora, 15(1), 281-287.

Beck, D., & Purcell, R. (2020). Leadership in the Community. In Community Development for Social Change (pp. 144-147). Routledge.

Breen, J. M., & Gleason, M. C. (2022). Setting the stage: An overview of articles in this issue. New Directions for Student Leadership, 2022(176).

Budiman, A. S., & Hajarisman, N. (2024). Analisis Mediasi Multipel Paralel Kausal Step pada Data Stunting menurut Kabupaten/Kota. Jurnal Riset Statistika, 31-40.

Carter, A. R., Sanderson, E., Hammerton, G.,

Richmond, R. C., Davey Smith, G., Heron, J., ... & Howe, L. D. (2021). Mendelian randomisation for mediation analysis: current methods and challenges for implementation. European journal of epidemiology, 36(5), 465-478.

Correia-Harker, B.P., & Dugan, J.P. (2020). Beyond Knowledge and Skills: Exploring Leadership Motivation as a Critical Construct for Student Leadership Development. Journal of College Student Development, 61, 299 - 316.

Coşkun, Ö., Timurçin, U., Kıyak, Y. S., & Budakoğlu, I. İ. (2023). Validation of IFMSA social accountability assessment tool: exploratory and confirmatory factor analysis. BMC Medical Education, 23(1), 138.

Didiano, T. J., Simpson, A. E., & Reeve, D. (2021, July). A Leadership-Development Ecosystem for Engineering Graduate Students. In 2021 ASEE Virtual Annual Conference Content Access.

Din,W.A., & Bidin,S.J A.（2014）.CONFIRMATORY FACTOR ANALYSIS OF THE REVISION STRATEGIES QUESTIONNAIRE IN MALAYSIAN ESL WRITERS'CONTEXT.

Ding, W., Chen, J., Liu, J., Lin, B., Li, S., Li, F., ... & Li, J. (2022). Development and validation of the health education adherence scale for stroke patients: a cross-sectional study. BMC neurology, 22(1), 69.

Gleason, M. C., & Edelman, L. E. (2022). Leveraging your programmatic "why": Articulating graduate leadership program value through mission alignment. New Directions for Student Leadership, 2022(176), 19-29.

Güvendir, M. A., & Özkan, Y. Ö. (2022). Item removal strategies conducted in exploratory factor analysis: A comparative study. International Journal of Assessment Tools in Education, 9(1), 165-180.

Hadie, S. N. H., Ismail, I. M., & Yusoff, M. S. B. (2023). Development and validation of Digi-MEE Instrument measuring online learning environments in medical education. Pakistan Journal of Medical Sciences, 39(6), 1573.

Hashemian Moghadam, A., Nemati-Vakilabad, R., Imashi, R., Yaghoobi Saghezchi, R., Dolat Abadi, P., Jamshidinia, M., & Mirzaei, A. (2024). The psychometric properties of the Persian version of the innovation support inventory (ISI-12) in clinical nurses: a methodological cross-sectional study. BMC nursing, 23(1), 699.

Hayes, K., Carter, T., Cook, P., Twaddell, E., & Buma, B. (2022). Supporting graduate field leadership through community-sourced advice, action, and policy. Ecosphere, 13(9), e4247.

Hazra, M.S., Mishra1, M.A., Gulati1, M.I., Sapra1, M.T., & Firdaus1, M.R. (2018). Motivation and Leadership Skills. International Journal of Indian Psychology.

Huang, S., & Reynolds, M. (2022). Facts that Influence College Students' Reading Motivation. ATHENS JOURNAL OF EDUCATION.

Izah, S. C., Sylva, L., & Hait, M. (2023). Cronbach's alpha: A cornerstone in ensuring reliability and validity in environmental health assessment. ES Energy & Environment, 23, 1057.

Latina, E., & Leadership, V.O. (2020). ELEVATING LATINA VOICES OF LEADERSHIP : Latina Student Leaders ' Beliefs , Approaches , and Influences to Leadership.

Li, S., Zhou, Z., Cai, F., Cheng, L., Du, W., Bai, H., ... & Zhang, J. (2024). Assessing reliability and validity of the Chi nese version of the Depression Change Exp ectancy Scale and its application in college students.

Linares, P. F., Higueras, E. R., Martin-Ferreres, M. L., Torre, M. Á. C., Capellades, L. W., & Fernández-Puebla, A. G. (2020). Dimensions of leadership in undergraduate nursing students. Validation of a tool. Nurse education today, 95, 104576.

Perruci, G., & Schwartz, S. W. (2002, September). Leadership for what? A humanistic approach to leadership development. In Art of Management and Organisation Conference, King's College, London, September (pp. 2-5).

Ritmak, N., Rattanawong, W., & Vongmanee, V. (2023). A New Dimension of Health Sustainability Model after Pandemic Crisis Using Structural Equation Model. Sustainability, 15(2), 1616.

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. Anesthesia & analgesia, 126(5), 1763-1768.

Jam, F. A., Ali, I., Albishri, N., Mammadov, A., & Mohapatra, A. K. (2025). How does the adoption of digital technologies in supply chain management enhance supply chain

performance? A mediated and moderated model. Technological Forecasting and Social Change, 219, 124225.

Skalicky, J., Warr Pedersen, K., Van der Meer, J., Fuglsang, S., Dawson, P., & Stewart, S. (2018). A framework for developing and supporting student leadership in higher education. Studies in Higher Education, 45, 100 - 116.

T. Kárász, J., Nagybányai Nagy, O., Széll, K., & Takács, S. (2022). Cronbach-alfa: vele vagy nélküle?. Magyar Pszichológiai Szemle, 77(1), 81-98.

Tan, W.C., & Koh, H.C. (2023). Applying Structural Equation Modelling in Education Research. 2023 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE), 1-8.

Voelkel, R. H., Johnson, C. W., & Gilbert, K. A. (2016). Use of immersive simulations to enhance graduate student learning: Implications for educational leadership programs. Online Journal of Distance Learning Administration, 19(2), 1-16.

Wang, Y. M., Hung, C. H., Li, Y. C., Ho, Y. C., & Huang, C. Y. (2025). Evaluating the construct validity of the health promotion literacy scale: a confirmatory factor analysis in Taiwan's university social responsibility context. BMC Medical Education, 25(1), 428.

Winton, S. L., Palmer, S., & Hughes, P. J. (2018). Developing leadership for increasing complexity: A review of online graduate leadership programs. Journal of Leadership Education, 17(1), 162-176.

Xie, M., Wang, A., Zhang, Z., Wang, K., Yu, Y., Ma, X., ... & Ke, J. (2024). Development and validation of the perceived symptom manageability scale among people living with the human immunodeficiency virus. BMC psychology, 12(1), 172.

Zheng, S., & Cao, Y. (2022). Correlation analysis for different types of variables and relationship between different correlation coefficients. Biom Biostat Int J, 11(4), 127-129.