Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

What Draws Students In? A Psychometric Exploration of Students Choices in Higher Education

Girisha Jain
Easwari School of Liberal Arts, SRM University-AP
Ayesha Parveen Haroon
Easwari School of Liberal Arts, SRM University-AP
Prabakaran Ganeshan
SRM University- AP
Karthik Rajendran
rajendran.k@srmap.edu.in

SRM University- AP

Article

Keywords: Student Preferences, Psychometrics, Higher Education, Theory of Planned Behaviour, Equitable Access

Posted Date: June 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9455998/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

What Draws Students In? A Psychometric Exploration of Students Choices in Higher Education

Girisha Jain[1], Ayesha Parveen Haroon[1], Prabakaran Ganeshan[2], and Karthik Rajendran*[2]

[1] Easwari School of Liberal Arts, Department of Psychology, SRM University-AP, Andhra Pradesh-522240, India

[2] The NET ZERO Lab, Department of Environmental Science and Engineering, SRM University-AP, Andhra Pradesh-522240, India

*Corresponding author Email: rajendran.k@srmap.edu.in

Abstract

Student Preferences in choosing a higher education institution involves complex interaction of institutional, geographic, psychological and socio-economic factors. Gaining insight in these factors will advance equal access to quality education. Existing research focuses on individual factor with either prediction or forecasting instead of understanding the interaction between the factors. This study addresses the gap by developing a questionnaire based on Theory of Planned Behaviour (TPB). The questionnaire was developed and validated using Crohn Bach Alpha and Content, Construct validity. The content validity through expert evaluation showed acceptable Content Validity Index (CVI)=1. The construct validity analysed through Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) as well as Convergent and Discriminant Validity. The EFA factor loadings results were consistent with the TPB three factor structure with high factor loadings (0.61 – 0.88), KMO = 0.97 and Bartlett's test ($p<0.05$). The CFA demonstrated good model fit with CFI= 0.93, RMSEA = 0.079, GFI=0.90. The convergent and discriminant validity was confirmed with CR values > 0.70, AVE values > 0.50 and Fornell-Larcker, HTMT test range (<0.85). Overall, two items related to self-confidence and same behaviour were removed showing consistent low factor loadings and estimate in content and construct validity. The reliability of the questionnaire analysed through Crohn Bach Alpha ($\alpha = 0.96$) showed high internal consistency yielding good psychometric and theoretically coherent for assessing student preferences in higher education choices. This will lead to equal and accessible education opportunities contributing to SDG 4 (Quality Education) and SDG 5 (Gender Equality).

Keywords: Student Preferences, Psychometrics, Higher Education, Theory of Planned Behaviour, Equitable Access

1 Introduction

Deciding where to study is one of the most consequential choices in a student's life, shaped by interplay of factors (Lyu et al., 2023). Understanding student preferences in selecting an institution can promote equity in education. This ensures that all students, regardless of background, have fair access to opportunities and resources. While accessibility towards choice of institution has increased with globalization disparities in affordability, quality and equal participation continue to persist (S. K. Jha, 2025).

Affordability remains a significant barrier, with gaps in higher education enrollment between the poorest and richest, reaching between 60% or more in some countries. The highest poor enrollment gap is in countries like Bolivia (66%), Panama (74%), Mongolia (70%) (OECD, 2024). The inequality also exists in quality of university where the top income quartile is twice as likely to attend high ranked universities compared to those from the bottom quartile (UNESCO, 2022). In terms of equal participation gaps are present in gender representation, where women participation in tertiary education is overrepresented. The underrepresentation of women participation exists STEM (≈40%) compared to Arts, Humanities and Social Sciences (≈60%) , Medicine (≈70%) (UNESCO IESALC & THE, 2020).

The disparities can be addressed by understanding the factors shaping student preferences when choosing an institution. Recent research has studied student preferences through theoretical frameworks from diverse domains, including economics, sociology, migration studies, and geography. The push-pull theory explains mobility in terms of factors that push students away (e.g., limited opportunities) and factors that motivate i.e., pull them (e.g., better quality, career prospects) (Singh et al., 2022). The gravity model studies the bilateral migration flows where student flows are positively related to the size of the origin and destination systems and negatively related to the distance between the origin and destination (Van Bouwel & Veugelers, 2013). The Human capital theory views education as an investment, where students choose destination that maximizes future returns in skills and earnings(Becker, 1993). The network theory emphasizes the role of personal,

institutional and alumni networks in facilitating and sustaining mobility(Shields, 2013).

The preferences to choose a university is a cognitive process involving individual preferences to institution specific factors. Beyond the theories, a psychological lens in terms of motivation, either instrumental aspirations (e.g., higher salaries, institutional reputation) or intrinsic aspiration (e.g., experiencing new culture) is used to understand student preferences (Lyu et al., 2023). The geography of the place i.e., perception shaped by media and social narratives, social networks and desire for multicultural experiences also motivate students to mobilize (Beech, 2020).

In addition to the theories and concept highlighting institutional, geographic and psychological influences, socio-economic factors play a pivotal role. The socio demographic factors also known as factors that are individual specific contribute indirectly towards student preferences. Socio-economic factors like income level and birth order influences affordability, parents' education level influences quality and access. towards decision making in choosing higher education institutions (S. Jha & Kumar, 2017).

The interaction of institutional, geographic, psychological and socio-economic factors can be systematically analyzed through the Theory of Planned Behavior (TPB). TPB, the human decision process is widely applied to predict behavior in the constructs of attitude, subjective norms and perceived behavioral control (Ajzen, 2020). In the context on student preferences, attitude construct like institutional characteristic and intrinsic aspiration focuses on career prospects or cultural enrichment (Becker, 1993; Singh et al., 2022) . The subjective norms constructs aligning with the network theory understand the influences through family, peers or neighbors (Shields, 2013). The perceived behavioral control constructs highlights the resources available motivating to choose an institution (Van Bouwel & Veugelers, 2013).

Despite extensive research on student preferences, no holistic questionnaire exists that simultaneously captures institutional, geographic, psychological and social-economic dimensions of student preferences within a unified approach. Existing questionnaire focuses on isolated factors or predictive modeling, overlooking the interplay of

attitudes, social norms and perceived control emphasized in the Theory of Planned Behavior.

The aim of this study was to develop a questionnaire to understand the student preferences in terms of affordability, quality and equal participation. The objectives of the study are,

a) Developing a questionnaire using Theory of Planned Behavior to understand the student preferences in choosing a higher education institution.
b) Analyze validity and reliability of the developed questionnaire.
c) Modify the questionnaire using the analysis of validity and reliability

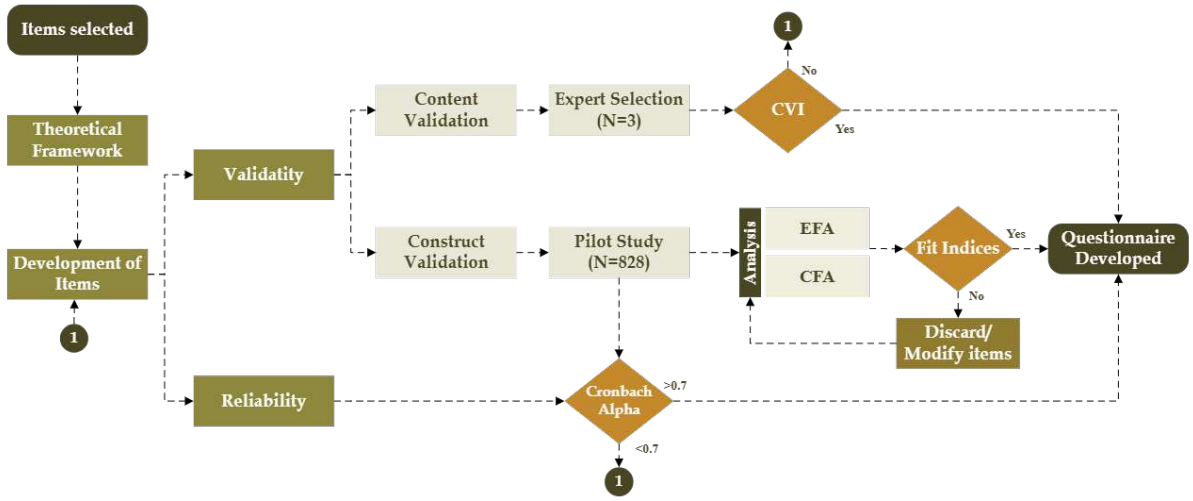


Figure 1

Methodology diagram depicting the process for questionnaire development.

Items for the questionnaire were initially selected and aligned with a theoretical framework, specifically the Theory of Planned Behavior. The developed items were then assessed for validity and reliability. Content validity was evaluated by three experts from relevant fields (student mobility, psychology, and language), and the Content Validity Index (CVI) was calculated. Construct validity was tested using data from a pilot study (N=828) through Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), with model fit indices. Reliability was measured using Cronbach's Alpha, with a threshold of ≥0.7. Items that did not meet the required standards were modified or discarded. The final questionnaire was developed using only the items that demonstrated acceptable validity and reliability.

2 Methodology

The questionnaire was developed to explore students' preferences and factors influencing decision-making when selecting higher education institutions. The items related to student preferences were selected from existing literature. Based on a Theory of Planned Behavior (theoretical framework) a questionnaire was developed and subsequently administered in a pilot study. The data from the pilot study were analyzed to assess the reliability and validity of the questionnaire.

Development of the Questionnaire

The development of the questionnaire was based on the factors influencing students' preferences when selecting a higher education institution. The factors were selected based on the previous research involving secondary data analysis (Jain et al., 2025). Other psychosocial factors were identified through a review of existing literature, using keywords such as 'higher education institutions', 'student diversity' and 'student mobility'. The factors like financial support, cost of migration, guidance of team and network support plays an important role in international student mobility (Batista et al., 2024). Rank being another important factor (Margaret D. M. Cullen, Andre P. Calitz, 2020). Despite that each study focuses on limited factors, this questionnaire was developed to understand students' preference wholistically in each dimension. The questionnaire was developed using Theory of Planned Behavior (TPB)(Ajzen, 2019).

Theory of Planned Behavior helps understand behavioral intention of a person in three core constructs (Attitude, Subjective Norms and Perceived Behavioral Control). The attitude construct identifies the behavior beliefs i.e., outcome like career opportunities and discipline either positively or negatively influencing student preferences. Whereas the subjective norms highlight the social influence from family, friends or neighbors in choosing a Higher Education Institution. Perceived Behavioral Control understand the internal and external constraints like, nearby home, accommodation and cost of living in choosing an institution (Ajzen, 2019). The demographic and social factors (Chandrasekhar & Sharma, 2014) related to student's institution preferences were collected through literature and each was mapped

153 together as per the TPB constructs. Items of the survey were developed using the core
154 constructs of TPB and factors identified through literature review.

155 Item Development for the Questionnaire

156 Students' preferences were influenced by many factors like gender, scholarship
157 availability, ranking, family, course availability, career opportunities, family or
158 friends influence, accommodation, cost of living and safety in campus. The factors
159 were included as per the TPB constructs. The question framed for the attitude
160 construct focused on important factors that influenced student in choosing a
161 university, "*Before joining this university, what factors did you consider important when*
162 *choosing a university*?". The social construct focused on, "*How much did the following*
163 *group influence your decision in choosing this university*?". Also, a question related to
164 likely social behavior was formulated, "*Most of your friends/seniors/classmate have also*
165 *chosen a university away from home*". The internal and external constraints of the
166 perceived behavioral control construct was formulated in terms of self-esteem and
167 importance of factors. The question formulated for self-esteem was "*I am confident that*
168 *studying in this university will lead to personal growth and success*". The question
169 formulated for external constraints was "*While choosing this university, what factors you*
170 *considered as important?*".

171 The formulated questionnaire was administered through the Likert scale (1-5) to
172 gather respondents' perception across the key constructs. The five-point scale was
173 employed instead of seven-point scale as the factors were considered from literature.
174 The Likert scale ranged from Very Unimportant (1) to Very Important (5), while few
175 questions administered No Influence at all (1) to Extreme Influence (5) and Very
176 Unlikely (1) to Very Likely (5) as per their formulation.

177 Sample and Data Collection

178 The questionnaire before administering to students was reviewed and accepted by the
179 Institutional Ethical Committee (IEC: SRMAP/IEC/2024-13) (Philip, 2024). All
180 methods were performed in accordance with the relevant guidelines and regulations.
181 Each participant provided informed consent through google form prior to
182 questionnaire administration. The questionnaire was administered to a sample of

undergraduate students, selected through convenience sampling from universities in Andhra Pradesh. A total of 828 samples were collected through online mode, and the demographic distribution was given in **Table 1**. The sample majorly comprised of 1st year undergraduate students of SRM University-AP and Government Medical Colleges of Andra Pradesh. The sample collected was later used in analysis for construct validity and reliability.

Table 1

Sociodemographic information for the overall sample data (N=828)

Variable	Count	Mean	SD	Min	25%	50%	75%	Max
Age	828	18.72	1.14	16	18	18	19	23

	Variable	n	%
Personal Details	Gender		
	Female	417	50.4
	Male	411	49.6
	Nature of Family		
	Nuclear	643	77.7
	Joint	185	22.3
	Siblings		
	One	606	73.2
	None	97	11.7
	Two & more than two	125	15.1
	Birth Order		
	Oldest	374	51.2
	Youngest	318	43.5
	Middle	39	5.3
	Family Income		
	≤ 10 Lakh	593	71.6
	>10 & <20 Lakh	165	19.9
	>20 & <30 Lakh	38	4.6
	≥30 Lakh	32	3.9
	Religion		
	Hindu	730	88.2
	Muslim	48	5.8
	Christian	45	5.4
	Others	5	0.6
Parents Education & Occupation	Fathers Education		
	PhD	33	4.2
	Postgraduate	258	32.3
	Undergraduate	162	20.5
	Diploma	60	7.6
	12	83	10.5
	10 (SSC) & Below 10	196	24.7
	Mothers Education		
	PhD	20	2.6
	Postgraduate	193	24.7

	Undergraduate	160	20.5
	Diploma	24	3.1
	12	109	13.9
	10 (SSC) & Below 10	275	35.2
	Fathers Employment	796	
	Government Employee	213	26.8
	Agriculture Sector	178	22.4
	Entrepreneur	106	13.3
	Industrial Sector	92	11.6
	Service Provider	67	8.4
	Unemployed	36	4.5
	Health Sector	35	4.4
	Education Sector	34	4.3
	IT Sector	23	2.9
	Deceased	8	1
	Home maker	4	0.5
	Mothers Employment	789	
	Government Employee	105	13.3
	Agriculture Sector	16	2
	Entrepreneur	15	1.9
	Industrial Sector	4	0.5
	Service Provider	11	1.4
	Unemployed	20	2.5
	Health Sector	15	1.9
	Education Sector	34	4.3
	IT Sector	7	0.8
	Deceased	8	1
	Home maker	554	70.2
Academic Details	**Stream of Study**		
	Medicine	439	53
	STEM	369	44.6
	Others	20	2.4
	Year of Study		
	I	811	98
	II	7	0.8
	III	6	0.7
	IV	4	0.5
	Lived away from home during school		
	Yes	223	26.9
	No	605	73.1

Validity and Reliability

The questionnaire was assessed on validity and reliability to measure the accuracy and consistency of the items respectively. The validity addresses the meaningfulness and accuracy of what it's intended to measure, whereas reliability addresses if the questionnaire was yielding similar results under consistent conditions. The validity of the questionnaire was measured using content validity and construct validity. The other construct validity parameters like convergent, discriminant and criterion

validity weren't measured due to lack of questionnaire related to student preferences in choosing higher education.

2.1.1 Content Validity

Content validity of the questionnaire was measured to test the aspects covered. Three subject matter experts from the field of student mobility, psychology and language reviewed the items in questionnaire on relevance and clarity aspect. Informed consent was obtained from all the experts for the content validation process. The experts were mailed content validation form with clear instructions to administer relevance and clarity from 1 (Irrelevance/lack of clarity) – 4 (Very Relevant/ Clear). The relevance and clarity scores were calculated using Content Validity Index (CVI). Each item was scored by the experts using Item-CVI (I-CVI) showed in the Equation ***(1)*** and the average I-CVI was calculated for the Scale-CVI (S-CVI/Ave) depicted in the Equation **(2)**. The proportion relevant is the average relevance rating by individual expert i.e., (S-CVI/Ave) was calculated using Equation **(3)**. The S-CVI/Ave was compared to the universal agreement S-CVI/UA method i.e., scoring each item 3 or 4 showed in the Equation **(4)** (Yusoff, 2019).

$$I - CVI = \frac{Agreed\ item}{Number\ of\ expert} \quad \textbf{(1)}$$

$$S - CVI/_{Ave} = \sum \frac{I - CVI\ Scores}{Number\ of\ item} \quad \textbf{(2)}$$

$$S - CVI/_{Ave} = \sum \frac{Proportion\ relevant\ rating}{Number\ of\ expert} \quad \textbf{(3)}$$

$$S - CVI/_{UA} = \sum \frac{Sum\ of\ UA\ scores}{Number\ of\ item} \quad \textbf{(4)}$$

where, agreed item means the experts scoring either 3 or 4, proportion relevant rating was scored by average of relevance rating by each expert. The Universal Agreement (UA) scores were calculated using the average of maximum score achieved in each item. The results of the CVI were attached in Supplementary Table 1.

The suggestions from experts related to including of "safety in campus" factor as well as wording changes from housewife to home maker were incorporated and the updated questionnaire was attached in **Appendix A**.

2.1.2 Construct Validity

Construct Validity measures the theoretical construct of the questionnaire through factor analysis. Factor Analysis helps in determining the items in the questionnaire aligns with the theoretical constructs. The constructs of the questionnaire were identified through Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). EFA identified potential factor structures without prior assumptions whereas CFA tested the hypothesized factor structure based on TPB (Tavakol & Wetzel, 2020).

The EFA and CFA were analyzed utilizing 828 data randomly split into half. The analysis was performed using semopy version 2.3.11 package in Python Version 3.11.9 (Igolkina & Meshcheryakov, 2020). EFA and CFA were represented as $X = \Lambda F + \varepsilon$, where X is vector of the observed variable, Λ denotes factor loadings, F is the latent structure and ε is the error term. In EFA, eigen values were utilized to determine the optimal number of factors to retain. Additionally, scree plot was used to determine the optimal number of factors using the 'elbow point' where the curve formed an angle resembling an elbow. Later, factor loadings for each item were analyzed using orthogonal rotation (assuming factors are uncorrelated) i.e., varimax rotation making the factor structure cleaner. The loadings were then represented in a heat map for further factor analysis. The CFA was conducted based on the TPB constructs and the factor loadings were estimated to assess the relationship between questionnaire items and TPB constructs. The results of the CFA were illustrated using a path diagram.

Both EFA and CFA were analyzed to check the goodness of fit. The EFA's goodness of fit was analyzed using Kaiser Meyer Olkin (KMO) test and Bartlett's Test of Sphericity. The KMO test determines the suitability of the data for factor analysis whereas the Bartlett's Test of Sphericity checks the data's correlation to justify using factor analysis. The CFA's goodness of fit was determined using tests like Chi-Square Statistics, Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), Goodness of Fit Index (GFI), Adjusted GFI (AGFI), Akaike Information Criterion (AIC), Bayesian Information Criteria (BIC) and Tucker Lewis Index (TIC). The test Chi-square, RMSEA and GFI finds the absolute fit, i.e., how well the model

253 reproduces the observed data. The test CFI and TLI determines the incremental fit,
254 i.e., how much better the model is compared to baseline model. The other tests like
255 AIC, BIC and AGFI were used to avoid overfitting. The description for both the tests
256 were shared in supplementary Table 4.

257 The convergent and discriminant validity indicators were identified within and
258 between the constructs. The convergent validity was established by calculating
259 Composite Reliability (CR) and Average Variance Extracted (AVE) with each TPB
260 construct. The discriminant validity was examined using Fornell-Larcker Matrix
261 calculated through √AVE values exceeding inter-construct correlations and
262 Heterotrait-Monotrait (HTMT) (Cheung et al., 2024).

263 **2.1.3 Reliability**

264 Reliability measures consistency, i.e., measurements stable and repeatable of the
265 questionnaire with different methods like, Internal Consistency, Test-Retest
266 reliability, Inter-Rater Reliability, Parallel-Forms Reliability and Split -half reliability.
267 As the test was preliminary and exploratory internal consistency was measured using
268 Cronbach's alpha (α) shown in Equation **(5)**. The Cronbach alpha indicates agreement
269 between items scoring between 0 to 1, the higher the score the higher agreement
270 among the items.

$$\alpha = \frac{N \times \bar{c}}{\bar{v} + (N-1) \times \bar{c}} \quad \textbf{(5)}$$

271 where, N is number of items, $\bar{c}$ & $\bar{v}$ represents mean covariance between items and
272 mean item variance respectively.

273 **2.1.4 Demographic and Socioeconomic Analysis**

274 Additionally, the study analysed a sample of 1,589 Indian undergraduate students, in
275 which 1,499 were selected due to missing rows in parents' education. Data collected
276 had demographic variables like personal (gender, birth order, sibling count, and
277 academic stream) and parental background (family income, employment, and
278 education). The mobility was defined in a binary outcome Movers (students attending
279 university in a different state, n=384 (24.2%)) and Stayers (students attending
280 university in their home state, n= 1,205). Birth order, Socio economic status (SES)

categorised in the order highlighted in Appendix A and was analysed using Chi-square (χ^2) and Cramér's V for effect size. Spearman's rank correlation (ρ) was utilized to test for ordinal variables like income and education. Logistic Regression was employed to calculate Odds Ratios (OR), determining which demographics significantly predict mobility when controlling for other household factors.

3 Results

The developed questionnaire was administered to sample and validated for content and construct1. The face validity was conducted with 20 students to determine if the questionnaire is measuring what it's intended to measure before content and construct validity. Later, validity and reliability were analysed through field experts and data collected from sample.

Content Validity

The relevance and clarity of the questionnaire items were assessed by three subject matter experts across student mobility, psychology, and language, using a 4-point rating scale. All items received scores of either 3 (Relevant/Clear) or 4 (Highly Relevant/Very Clear), indicating strong content validity. Notably, most segments—particularly those related to Attitude and Subjective Norms—were consistently rated 4 across both domains. The Control Beliefs segment showed slightly lower consensus, where items pertaining to "confidence" received a score of 3 in both relevance and clarity from the psychology expert. Additionally, the language expert rated clarity as 3 for items in the "social circle" subdomain.

The Item-CVI (I-CVI) values were as per the agreeable standard across all items was calculated. The total item CVI was scored in a Scale-CVI average (S-CVI/Ave) with 1.0, highlighting acceptable CVI values for content validity (Yusoff, 2019). The Universal Agreement method (S-CVI/UA) also yielded satisfactory results, the CVI score from each expert is attached in the Supplementary data Table 1. The Expert's suggested revisions, such as including "safety in campus" and updating terminology in demographic variable from "housewife" to "home maker" were incorporated. After the expert suggestion and acceptable CVI score, the questionnaire was then administered for construct validity.

Construct Validity

Construct validity of the questionnaire was examined through both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) to determine whether the items aligned meaningfully with the theoretical framework of the Theory of Planned Behavior (TPB). The developed questionnaire was administered with a total of n=828 data for construct validity analysis.

3.1.1 Exploratory Factor Analysis

EFA was conducted on a random split subset of the dataset (n=414). The data was analyzed using Eigen value and plotted as the scree plot in ***Figure* 2.** The scree plot highlighted an elbow point at 3, which shows the need of three factors consistent with TPB's core constructs. The factor loadings as per the three factors was represented in the factor loadings heatmap in **Figure 3,** the items clustering revealed strong convergence around the TPB constructs. For example, items including Academic Reputation, Carrer Opportunities and Discipline showed consistently high loadings (0.88) on Factor 1 (corresponding to be behavioral beliefs in TPB constructs). Factor 2 (corresponding to be Normative beliefs in TPB construct) showing a cluster of loadings like Family, Friends and Seniors in the range 0.61 to 0.83. Factor 3 (corresponding to be Perceived Behavioral Control in TPB construct), clustering items like, scholarship availability, accommodation & cost of living and ranking. The factor structure exhibited minimal cross loadings, confirming dimensional integrity and theoretical coherence.

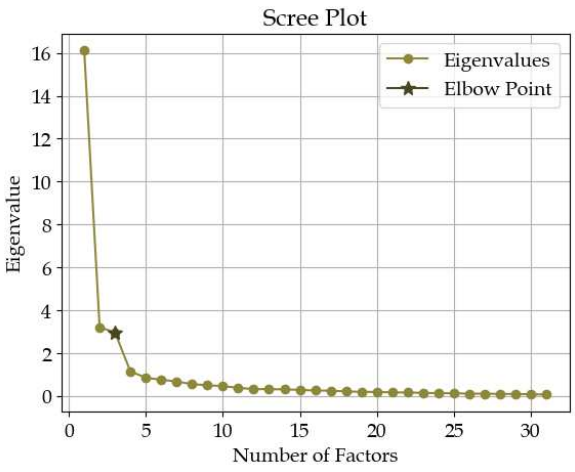


Figure 2

Scree plot depicting the eigenvalues used to determine the optimal number of factors to retain in the model.

The goodness of fit was obtained using tests like KMO and Bartlett's Test of Sphericity. The KMO was 0.97 showed that the data was appropriate for factor analysis and the

Bartlett's Test of Sphericity was $p<0.05$ showing the data was sufficiently correlated for factor analysis.

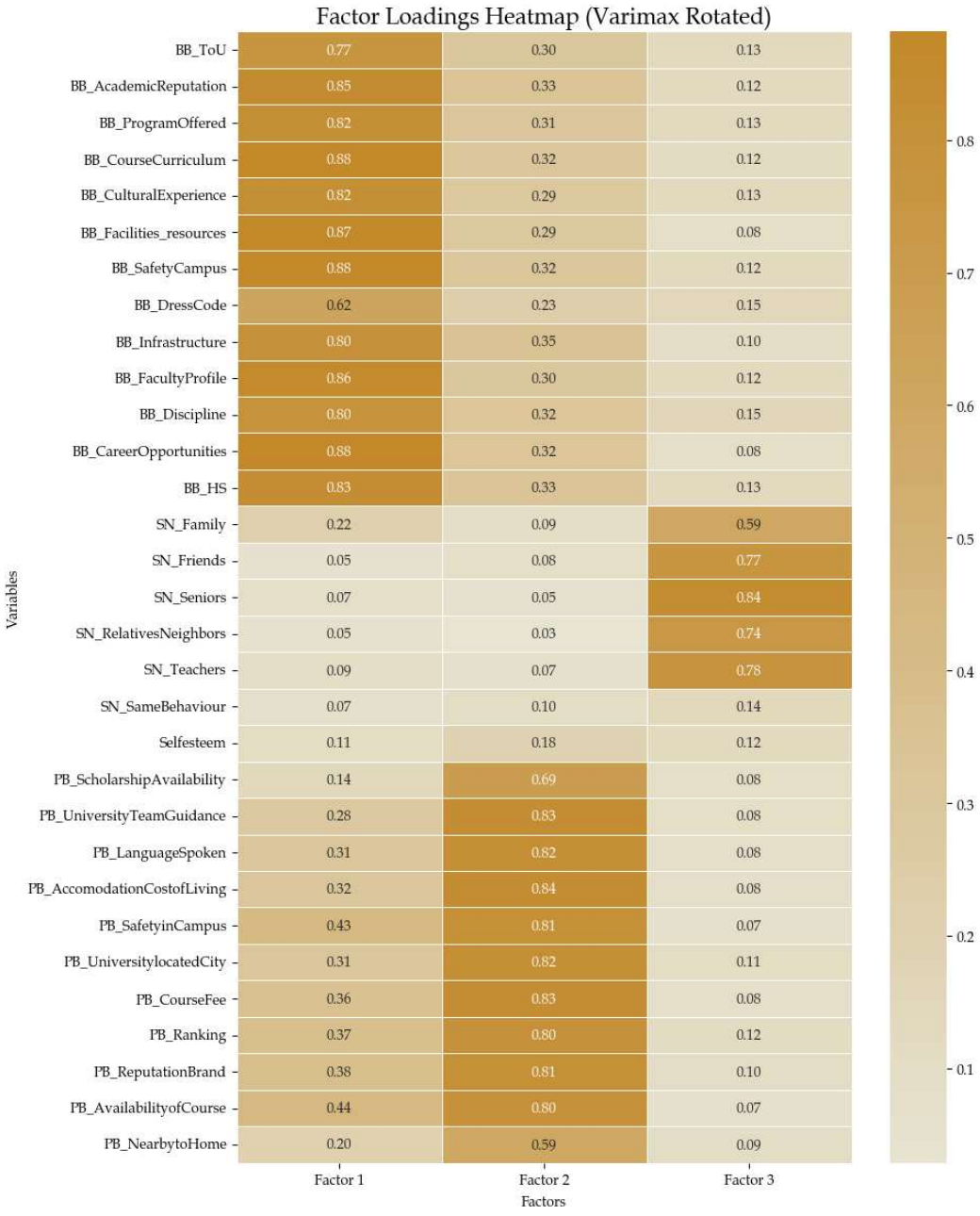


Figure 3

Heatmap of Factor Loadings from Exploratory Factor Analysis with Varimax Rotation

This figure illustrates the standardized factor loadings of observed variables on three extracted factors following varimax rotation. Darker shades represent stronger loadings, indicating a higher correlation between the variable and the respective factors.

3.1.2 Confirmatory Factor Analysis

The remaining data set (n=414) was randomly split for CFA analysis; theory of planned behavior (theoretical framework) was employed. The model was separated using three constructs Behavioral Beliefs, Normative Beliefs and Perceived Behavioral Control. Behavioral Beliefs included 13 observed variables reflecting institutional characteristics and perceived benefits. Normative beliefs included five social influencing variables whereas Perceived Behavioral Control included 10 variables including internal confidence and external constraints.

The estimates in the behavioral construct showed that all the variables were above Λ = 0.6, indicating robust relationships. The variables, same behavior and self-confidence were having estimate Λ = 0.2 i.e., showing low relationship with the normative beliefs construct. The two variables showing low estimates were discarded and the remaining variables and constructs were depicted in **Figure 4**. The CFA results indicated a strong goodness of fit for the proposed model with Chi-square (p>0.05), CFI = 0.93, RMSEA = 0.079, GFI =0.9, AGFI=0.89, AIC=115.5, BIC=361.08 and TLI = 0.92. The questionnaire's content and construct validity were examined, later the reliability of the questionnaire was assessed.

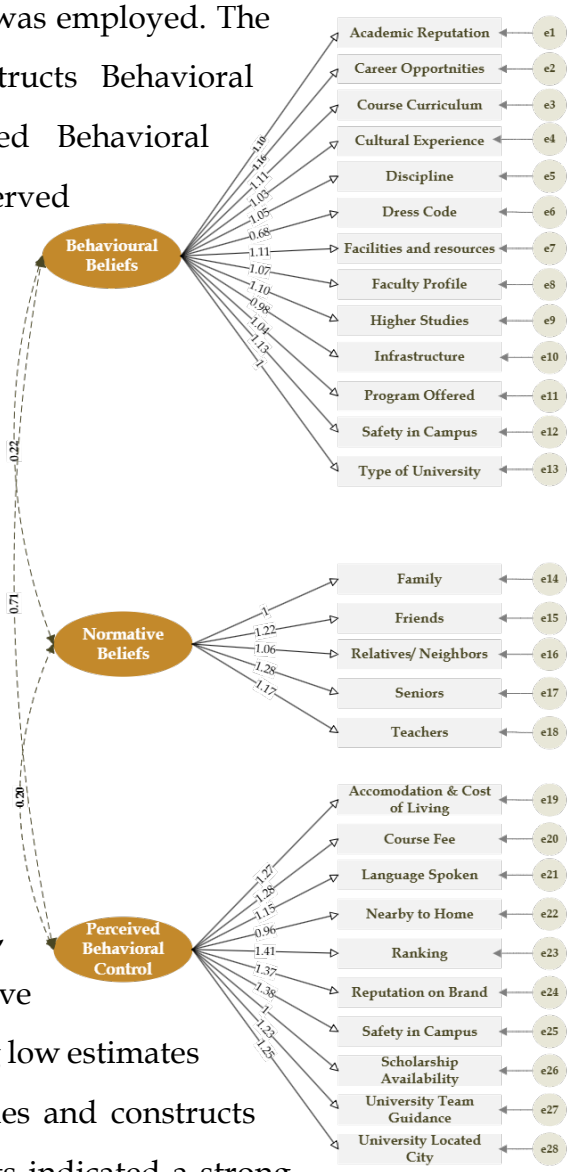


Figure 4

Confirmatory Factor Analysis Model of Behavioral, Normative, and Perceived Behavioral Control Beliefs

This figure illustrates the CFA model assessing the measurement structure of three latent constructs: Behavioral Beliefs, Normative Beliefs, and Perceived Behavioral Control. Each construct is measured

by multiple observed indicators with standardized factor loadings and associated error terms (e_1–e_{28}). The model evaluates the adequacy of the proposed factor structure aligned with the Theory of Planned Behavior.

3.1.3 Convergent and Discriminant Validity

The convergent and discriminant validity ensures the TPB constructs (Behavioral beliefs, Normative Beliefs, Perceived Behavioral Control) were internally consistent and conceptually distinct. The values for each construct showed acceptable convergent and discriminant validity displayed in the **Table 2.** The CR values above 0.70 and AVE values above 0.50 were considered applicable (Hair et al., 2019). The three constructs met these criteria, with CR values ranging from 0.685 to 0.730 and AVE values ranging from 0.650 to 0.709, indicating that the indicators effectively represent their respective constructs. The discriminant validity assessed by Fornell-Larcker criterion and the HTMT ratio showed that √AVE for each construct exceeded its correlations with other constructs, satisfying the fornell-larcker threshold (Fornell & Larcker, 1981). Additionally, HTMT values ranged from 0.205 to 0.681, remaining below the conservative cutoff of 0.85 (Henseler et al., 2015). The TPB construct demonstrates both convergent and discriminant validity, thereby supporting the structural integrity and theoretical distinctiveness of the measurement model.

Table 2: Convergent and Discriminant Validity of TPB Constructs

Construct	CR	AVE	√AVE	Max HTMT	Max Correlation
Behavioural Beliefs	0.701	0.709	0.842	0.658	0.680
Normative Beliefs	0.685	0.650	0.806	0.292	0.420
Perceived Behavioural	0.730	0.691	0.831	0.681	0.680

Note: *Composite Reliability (CR) and Average Variance Extracted (AVE) were calculated for each latent construct. All CR values exceeded the recommended 0.70 threshold, and AVE values were above the 0.50 benchmark, indicating strong convergent validity. √AVE values exceeded the maximum inter-construct correlations and HTMT ratios were below 0.85, confirming discriminant validity. (Fornell & Larcker, 1981; Hair et al., 2019)*

406 Reliability

407 The internal consistency of the developed questionnaire was evaluated using
408 Cronbach's alpha (α). The internal consistency was analyzed with the three constructs
409 of the TPB. The Behavioral beliefs and Perceived Behavioral Control showed strong
410 internal consistency of items α = 0.98 and 0.96 respectively, whereas the internal
411 consistency of social norms was moderate with α = 0.79. The overall consistency of the
412 questionnaire developed was α=0.96, showing strong internal consistency among the
413 questionnaire items. The analysis validates the results for the questionnaire
414 development which was further explained within existing research and theoretical
415 frameworks.

416

417 Demographic Drivers of Student Mobility

418 The personal demographic variables like birth order were a highly significant
419 predictor of interstate mobility ($\chi^2 = 31.23, df = 3, p$ =0.000001). Youngest children
420 exhibit the lowest mobility rate (17%), while middle children are the most mobile
421 (35.8%). Pairwise comparisons show that oldest children (28.3% mobility) are nearly
422 twice as likely to move as their youngest siblings (OR = 1.92, p<0.001). This birth order
423 effect remains significant even when controlling for total sibling count and family
424 income. The socio-economic status i.e., family income was identified as the single
425 strongest demographic predictor of mobility ($\chi^2 = 95.44,\ p$ =0.000001). The mobility
426 pattern in students with income level 30 lakhs and above category were 50% mobile.

427 The parental demographics impact was also present where parental education
428 independently reinforces the income effect. Children of parents with a PhD move at
429 roughly four times the rate of children whose parents have less than a 10th-grade
430 education (41–44% to 11–16%). Logistic regression highlighted in **Table 3** show that
431 parent education remains a significant predictor even after controlling for total family
432 wealth (OR = 1.15, p<0.001).

433

434

435 **Table 3:** Logistic Regression Predicting Mobility by Demographic Variables

Predictor	**Coefficient (β)**	**Odds Ratio (OR)**	**p-value**
Birth Order: Oldest	+0.563	1.757	0.0001***
Family Income	+0.478	1.613	0.0000***
Mother's Education	+0.175	1.191	0.0000***
Sibling Count	+0.713	2.040	0.0000***
Gender (Female)	-0.059	0.943	0.6455 ns

436 **Note: Reference categories are youngest child and Male. Significance: **p < 0.001, ns = not*
437 *significant*

438 **4 Discussion**

439 The questionnaire was validated and reliable to understand the decision-making
440 patterns of students in choosing an institution. The results of validity and reliability
441 for the questionnaire development were consistent and aligned with the Theory of
442 Planned Behavior (TPB) (theoretical framework). The factor loadings were strongly
443 and theoretically aligned with the three TPB constructs, further supported by robust
444 model fit indices. In particular, the validity and reliability of weaker aligned variables
445 self-esteem(CVI=3, Λ_{EFA}=0.12, Λ_{CFA}=0.24) and similar behaviour patterns(CVI=3,
446 Λ_{EFA}=0.14 , Λ_{CFA}=0.23) contributing to weaker psychometric outcomes. The overall
447 poor fit variables were discarded, and the final scale was attached in **Appendix A.**

448 In the initial, face validity was administered as the students already experienced the
449 behavior of choosing a university. Face validity ensured students could understand
450 the questions and their decision-making behavior clearly. Later, Confirmatory Factor
451 Analysis (CFA) was conducted using Maximum Likelihood (ML) estimation, which is
452 appropriate for large samples and continuous indicators. Meanwhile, the Exploratory
453 Factor Analysis (EFA) employed Principal Axis Factoring (PAF), a least-squares-based
454 estimator well-suited for non-normal psychological data. Varimax rotation, an
455 orthogonal technique, was used in EFA to simplify factor interpretation while

maintaining uncorrelated constructs. Since CFA does not involve rotation, the predefined structure was directly validated. These estimator and rotation choices align with best practices for psychometric modelling in social sciences, enhancing the methodological rigor of the validation process (Kyriazos & Poga-Kyriazou, 2023). The convergent and discriminant validity established psychometric thresholds showing theoretical distinctiveness and internal consistency of the questionnaire (Cheung et al., 2024).

The content validity and construct validity yielded similar results where CVI score and factor analysis both represented good relationship between variables except the same behavior and self-confidence variables. While internal validity was analyzed other forms of validity were not assessed. In particular, criterion related validity wasn't measured as there were no existing scale understanding students' preferences in choosing an institution. The internal consistency reliability measured using Cronbach alpha for each TPB construct revealed consistent results with validity. The inter rater as well as test-retest reliability were excluded as the questionnaire was a single time self-administration (Cooper, 2023).

The findings validates the applicability of the TPB framework for exploring institutional decision-making among students(Ajzen, 2020). The strong construct validity and internal consistency support the argument that students' choices are significantly shaped by perceived benefits (behavioural beliefs) (Van Bouwel & Veugelers, 2013), social expectations (normative beliefs) (S. Jha & Kumar, 2017), and external or internal constraints (perceived behavioural control) (Beine et al., 2014). Importantly, the validated instrument provides a comprehensive tool for educational institutions seeking to understand student behaviour and refine their outreach strategies accordingly.

Practically, the questionnaire can aid universities and policymakers in evaluating the factors that attract students, such as scholarship availability, academic reputation, or accommodation concerns. Additionally, the inclusion of social and psychological dimensions enables a holistic understanding of student preferences beyond economic and academic factors(S. Jha & Kumar, 2017). The questionnaire would also help us

bring equity in educational choices due to differences in gender, socio-economic constraints(Acai et al., 2022).

This study was limited by the use of convenience sampling from institutions in Andhra Pradesh, which may affect generalizability (Bujang et al., 2024). Future research should validate the instrument in diverse geographic, institutional and cultural settings, explore predictive validity by linking preference scores to actual enrolment behaviours to gather deeper insights into student motivations.

5 Conclusion

This study developed and validated a Theory of Planned Behavior based questionnaire to capture student preferences in selecting higher education institutions. The questionnaire showed strong content validity, with expert reviews ensuring relevance and clarity. Construct validity, assessed through EFA and CFA, confirmed the alignment of items with the three TPB construct Behavioral Beliefs, Normative Beliefs, and Perceived Behavioral Control while achieving satisfactory model fit indices. Convergent and discriminant validity metrics further support the theoretical distinctiveness and measurement integrity of the constructs. Reliability analysis showed high internal consistency across all constructs and the overall questionnaire. The validated questionnaire offers a robust tool for examining the multifaceted factors influencing student decision-making, with potential applications in institutional policy design, market research, and equity-focused educational planning. Future research could apply this tool across different genders, disciplines, mobility patterns, cultures, and socio-economic groups, enabling meaningful cross-country comparisons and richer insights into higher education choices.

Appendix A

Pathways of Choice: Student Mobility Questionnaire

Psychosocial Factors Influencing Student Mobility in Indian Higher Education Institutions

1. Participant Information

Name:	_____
Age:	☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20 ☐ 21 ☐ 22 ☐ 23
Gender	☐ Male ☐ Female ☐ Other
Family Type	☐ Nuclear ☐ Joint
Total Family Members	☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ More than 5
Number of Siblings	☐ Only Child ☐ One ☐ Two ☐ More than Two
Birth Order	☐ Oldest ☐ Middle ☐ Youngest ☐ Only Child
Annual Family Income	☐ ≤ ₹10 Lakh ☐ ₹10–20 Lakh ☐ ₹20–30 Lakh ☐ ≥ ₹30 Lakh
Religion	☐ Hindu ☐ Muslim ☐ Christian ☐ Sikh ☐ Jain ☐ Other:
Parents' Residence (City, State)	_____

Parents' Education Background

	Mother	Father
Below 10	☐	☐
10 (SSC)	☐	☐
12 (HSC)	☐	☐
Diploma	☐	☐
Undergraduate	☐	☐
Postgraduate	☐	☐
Ph.D.	☐	☐

Parents' Occupation

	Mother	Father
Government Employee	☐	☐
IT Sector	☐	☐
Industrial Sector	☐	☐
Entrepreneur	☐	☐
Healthcare Sector	☐	☐
Service Provider	☐	☐
Agriculture Sector	☐	☐
Education Sector	☐	☐
Unemployed	☐	☐
Deceased	☐	☐
Housewife	☐	☐

2. Academic Details

Which stream are you studying?

☐ STEM ☐ Arts ☐ Humanities & Social Sciences ☐ Management ☐ Medicine ☐ Other

Year of Study: ☐ I Year ☐ II Year ☐ III Year ☐ IV Year

What is name and location of your university:

_____

You are residing in?

☐ Institution Hostel ☐ Private Hostel ☐ Paying Guest (PG) ☐ Relative ☐ Home

Did you live away from home during school years? ☐ Yes ☐ No

How many languages can you speak?

☐ 1 ☐ 2 ☐ 3 ☐ More than 3

Did social media influence your university choice? ☐ Yes ☐ No

Do you know ranking bodies like NIRF, QS, THE, Outlook, etc.?

☐ Yes ☐ A few ☐ No

Is your institute providing scholarship for your education?

☐ Yes ☐ No

If yes, what is the type of Scholarship?

☐ Full Tuition Fee ☐ Partial Tuition Fee

Scholarship provided or supported by

☐ University ☐ Private Body (Industry, NGO, etc.) ☐ Government Body

3. Survey Questionnaire

1. What was the three main reasons for you to choose this university?

_____

2. Before joining this university, what factors did you consider important when choosing a university?

1: Very Unimportant, 2: Unimportant, 3: Neutral, 4: Important, 5: Very Important

Factor	1	2	3	4	5
Type of Institution (State/Central/Private)	☐	☐	☐	☐	☐
Academic Reputation	☐	☐	☐	☐	☐
Program Offered	☐	☐	☐	☐	☐
Course Curriculum	☐	☐	☐	☐	☐
Cultural Experience/ Social Environment	☐	☐	☐	☐	☐
Facilities and Resources (lab, funds, technology, etc.)	☐	☐	☐	☐	☐
Safety in Campus	☐	☐	☐	☐	☐
Dress Code	☐	☐	☐	☐	☐
Infrastructure	☐	☐	☐	☐	☐
Faculty Profile	☐	☐	☐	☐	☐
Discipline	☐	☐	☐	☐	☐
Career Opportunities	☐	☐	☐	☐	☐
Higher Studies Opportunity (SAP/ Pathway/ Integrated/Dual Degree)	☐	☐	☐	☐	☐

3. **How much did the following group influence your decision in choosing this university?**

1: No Influence, 2: Very Little Influence, 3: Moderate Influence, 4: Significant Influence, 5: Extreme Influence

Group	1	2	3	4	5
Family	☐	☐	☐	☐	☐
Friends	☐	☐	☐	☐	☐
Seniors	☐	☐	☐	☐	☐
Relatives/Neighbours	☐	☐	☐	☐	☐
Teachers	☐	☐	☐	☐	☐

4. **While choosing this university, what factors did you consider important?**

1: Very Unimportant, 2: Unimportant, 3: Neutral, 4: Important, 5: Very Important

Factor	1	2	3	4	5
Availability of Scholarship	☐	☐	☐	☐	☐
Guidance by University Team	☐	☐	☐	☐	☐
Language Spoken	☐	☐	☐	☐	☐
Accommodation and Cost of Living	☐	☐	☐	☐	☐
Safety in Campus	☐	☐	☐	☐	☐
University located in City	☐	☐	☐	☐	☐
Course Fee	☐	☐	☐	☐	☐
Ranking of the Institution	☐	☐	☐	☐	☐
Reputation/Brand	☐	☐	☐	☐	☐
Availability of Course	☐	☐	☐	☐	☐
Nearby to Home	☐	☐	☐	☐	☐

References

Acai, A., Mercer-Mapstone, L., & Guitman, R. (2022). Mind the (gender) gap: engaging students as partners to promote gender equity in higher education. *Teaching in Higher Education*, *27*(1), 18–38. https://doi.org/10.1080/13562517.2019.1696296

Ajzen, I. (2019). TPB Questionnaire Construction Constructing a Theory of Planned Behaviour Questionnaire. *University of Massachusetts Amherst*, 1–7. http://people.umass.edu/~aizen/pdf/tpb.measurement.pdf

Ajzen, I. (2020). The theory of planned behavior: Frequently asked questions. *Human Behavior and Emerging Technologies*, *2*(4), 314–324. https://doi.org/10.1002/hbe2.195

Batista, C., Costa, D. M., Freitas, P., Lima, G., & Reis, A. B. (2024). What Matters for the Decision to Study Abroad? A Lab-in-The-Field Experiment in Cape Verde. In *SSRN Electronic Journal* (Issue 17096). https://doi.org/10.2139/ssrn.4883855

Becker, G. S. (1993). *Human Capital: A Theoretical and empirical analysis with special reference to education* (Third). University of Chicago Press.

Beech, S. E. (2020). The geographies of international student mobility. Spaces, places and decision making. In *Hungarian Educational Research Journal* (Vol. 10, Issue 1). https://doi.org/10.1556/063.2020.00011

Beine, M., Noël, R., & Ragot, L. (2014). Determinants of the international mobility of students. *Economics of Education Review*, *41*(2014), 40–54. https://doi.org/10.1016/j.econedurev.2014.03.003

Bujang, M. A., Omar, E. D., Foo, D. H. P., & Hon, Y. K. (2024). Sample size determination for conducting a pilot study to assess reliability of a questionnaire. *Restorative Dentistry and Endodontics*, *49*(1), 1–8. https://doi.org/10.5395/rde.2024.49.e3

Chandrasekhar, S., & Sharma, A. (2014). On the internal mobility of Indians : Knowledge gaps and emerging concerns. *Indira Gandhi Institute of Development*

Research (IGIDR), September 2012.

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. In *Asia Pacific Journal of Management* (Vol. 41, Issue 2). Springer US. https://doi.org/10.1007/s10490-023-09871-y

Cooper, C. (2023). An Introduction to Psychometrics and Psychological Assessment. In *An Introduction to Psychometrics and Psychological Assessment*. SAGE Publications. https://doi.org/10.4324/9781003240181

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1), 39. https://doi.org/10.2307/3151312

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Igolkina, A. A., & Meshcheryakov, G. (2020). semopy: A Python Package for Structural Equation Modeling. *Structural Equation Modeling*, *27*(6), 2.3.11. https://doi.org/10.1080/10705511.2019.1704289

Jain, G., Haroon, A. P., Ganeshan, P., & Rajendran, K. (2025). Factors affecting student diversity in higher education institutions: a data-driven case study from India. *Discover Education*, *4*(1). https://doi.org/10.1007/s44217-025-00895-8

Jha, S. K. (2025). From state to market: Shifts in access and quality in Indian higher education. *International Journal of Educational Development*, *114*(December 2024). https://doi.org/10.1016/j.ijedudev.2025.103240

Jha, S., & Kumar, S. (2017). Socio-economic determinants of inter-state student mobility in India. *High. Educ. Future*, *4*(2), 166–185.

Kyriazos, T., & Poga-Kyriazou, M. (2023). Applied Psychometrics: Estimator Considerations in Commonly Encountered Conditions in CFA, SEM, and EFA Practice. *Psychology*, *14*(05), 799–828. https://doi.org/10.4236/psych.2023.145043

Lyu, J., Wei, Y., & Zheng, X. (2023). An Analysis on Motivation of International Students Who Come to China for Higher Education. In *Proceedings of the 2022 International Conference on Science Education and Art Appreciation (SEAA 2022)* (Vol. 1). Atlantis Press SARL. https://doi.org/10.2991/978-2-494069-05-3

Margaret D. M. Cullen, Andre P. Calitz, W. K. (2020). The Importance of University Rankings for Students University of Choice: A South African Perspective. In *Higher Education Marketing in Africa: Explorations into Student Choice* (pp. 315–342). Palgrave macmilan. https://doi.org/https://doi.org/10.1007/978-3-030-39379-3_12

OECD. (2024). *Education at a Glance 2024: OECD Indicators*. https://doi.org/https://doi.org/10.1787/c00cad36-en

Philip, M. (2024). Questionnaire, rating scale and checklist – How do they differ? *Methods in Psychology*, *10*(April), 100145. https://doi.org/10.1016/j.metip.2024.100145

Shields, R. (2013). Globalization and International Student Mobility: A Network Analysis. *Comparative Education Review*, *57*(4), 609–636. https://doi.org/10.1086/671752

Singh, R., Manchanda, N., & Mishra, R. (2022). Internal student migration in India: Impact of the COVID-19 crisis. *Asian and Pacific Migration Journal*, *31*(4), 454–477. https://doi.org/10.1177/01171968231154590

Tavakol, M., & Wetzel, A. (2020). Factor Analysis: a means for theory and instrument development in support of construct validity. *International Journal of Medical*

Education, *11*, 245–247. https://doi.org/10.5116/ijme.5f96.0f4a

UNESCO. (2022). Higher Education Global Data Report. In *3rd UNESCO World Higher Education Conference (WHCE2022)* (Issue May). https://creativecommons.org/licenses/by-sa/3.0/igo

UNESCO IESALC & THE. (2020). Gender equality: How global universities are performing: Part 1. In *UNESCO IESALC*. http://creativecommons.org/licenses/by-sa/3.0/igo/

Van Bouwel, L., & Veugelers, R. (2013). The determinants of student mobility in Europe: the quality dimension. *European Journal of Higher Education*, *3*(2), 172–190. https://doi.org/10.1080/21568235.2013.772345

Yusoff, M. S. B. (2019). ABC of Content Validation and Content Validity Index Calculation. *Education in Medicine Journal*, *11*(2), 49–54. https://doi.org/10.21315/eimj2019.11.2.6

Acknowledgements

This work acknowledges the invaluable support of the institutions and students that facilitated and enabled the data collection process. Their contributions in providing access, resources, and coordination were essential to the successful completion of this study. I also deeply grateful to the experts - Prof. S Irudaya Rajan (Chairman of the International Institute of Migration and Development (IIMAD), India), Dr. Bijeta Mishra (Social Psychologist) and Ms. Apurba Ganguly (Research Scholar, Department of Literature and Languages) for their insightful guidance and expertise, which significantly enriched the quality of the work.

CRediT authorship contribution statement

Ms. Girisha Jain: Methodology, Analysis, Writing – Review & Original Draft, Visualization.

Dr. Ayesha Parveen Haroon: Writing – review & editing, Supervision

Dr. Prabakaran Ganeshan: Writing – review & editing

Dr. Karthik Rajendran: Conceptualization, Writing – review & editing, Supervision & Project administration

Data Availability Statement

The datasets analyzed and the code used for the study are available in https://github.com/Girishajain/StudentPreferences.git

Fundings

This research received grant from SRM University-AP through the Research Grant under Grant No. SRMAP/URG/SEED/2025-26/034.

Declarations

Conflict of Interest

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed Consent

Informed consent was obtained from all individual participants prior to administering the questionnaire.

Ethics Approval

The study involved human participants. Approval was obtained from the Institutional Ethical Committee at SRM University-AP (IEC: SRMAP/IEC/2024-13)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- QDSupplementryGJ.docx