FROM EDUCATION TO EMPLOYABILITY: COMPARATIVE ANALYSIS OF SKILL AND EDUCATION INTEGRATION IN INDIAN STATES

Abstract

The given paper focuses on the connection between education and skill development systems across the states of India. It influences the employability of marginalized populations that represent Scheduled cast (SC), Scheduled Tribes (ST), Other Backward Class (OBC), Persons with Disability (PWD), Muslims, Minorities, and EWS groups. This was achieved using the secondary data through Principal Component Analysis (PCA), where the first principal component (PC1) represents the performance of the state grants in a composite index manner. K-means was used to cluster states into three groups of levels of inclusion. A Fixed Effects Model of Panel data in 10 years showed different effects of educational and skill upgrading variables on employability, indicating the different challenges to meet across states. Although a $R^2 = 1$ points to an over-fitted process, the model has notable directional tendencies. The results highlight the necessity of purposeful, inclusive policies, stronger data governance, and inter-sectoral integration to achieve better labour market results concerning marginalised groups and ensure equitable development at the state level in India.

1. Introduction

The narrative of education in India has also taken a restructuring turn in recent years, with discussion increasingly on issues of equitable access not to education, but on skill building, employability, and alignment to the labour market. The differences between the educational infrastructure in Indian states can be attributed to different factors, such as capacity and governance, public education, teacher supply, urban-rural divide, and the enforcement of central government education plans. Such factors as the demographic dividend of India, including a higher skilled labour force, can be transformed into the driver of economic growth once the education system is fine-tuned to the constantly evolving demands of the workforce (Shukla, 2021). The present research article discusses the state of social and regional inequality in educational infrastructure in the states of India. The research also investigates the relationship between vocational preparation, education, and job quality as well as how advanced education and skill development affect employment rates, wage levels, and access to decent work (Kumar et al., 2019). National-level data on pre-primary education in India is not available, and the 75th Round of the National Sample Survey (NSS), organised by the National Statistical Office (NSO) between July 2017 and June 2018, marked the first attempt to collect comprehensive data on pre-primary school. With the help of those datasets, it is observed that children in rural households are significantly less likely to attend preschool as compared to urban students. This rural-urban gap narrows with an increase in household income and the educational attainment of the household head (Choudhury et al., 2024). Although systemic barriers still exist due to this, there comes the necessity to have evidence-based comparisons among the Indian states about the transition from education to employability. Existing studies emphasize the significance of properly targeting the youth vulnerable groups, as it is necessary to make vocational education and training (VET) systems inclusive and successful (Gupta & Datta, 2023). Pilot schemes using data-driven methods to reach vulnerable young people, such as students who have dropped out of school and the informally employed, have shown positive measurable benefits in terms of employability outcomes when training is contextualized and accessible in urban regions such as Mumbai.

Individual characteristics such as age, sex, level of education, and marital status show a positive relationship with wages as they align with the standard labour market theory. Among all predictors, the level of education is projected as the most influential factor affecting wages. General education, as well as technical education, has a positive relationship with wage income. Interestingly, in the migrant group, the relatively stronger impact of general education implies that migrants with identical educational qualifications earn higher wages than their non-migrant counterparts, presumably because the migrants have greater employment opportunities or are intrinsically more motivated. The type of occupation is also a subject of key importance in determining wages. As expected, workers in agriculture, fishery, and elementary occupations earn considerably lower wages compared to semi-skilled and highly skilled jobs. This is an indication of the necessity to expand accessibility of education and skill training, not only to match the demands of the labour market to ensure an equitable delivery of wages (Parida, 2019). The economic analysis of the development and production of skills has been mainly influenced by two common approaches, such as the supply-side and the demand-side standpoints. The supply side confirms the efficiency and public investment in Vocational Education and Training (VET) systems, whereas the demand side stresses the match between educational output and the labour market, and changing skills requirements. One of the key supply-side measures is the Rate of Return (ROR) analysis controlled by the World Bank. After analysis, it is found that there were relatively lower returns to Vocational and Technical Education (VTE) compared with primary and general secondary education, leading to a policy of disinvestment in VET in favour of general education. This approach received increased attention in the wake of the development of the Education for All movement. The diversity of national education systems and their cultural and institutional variability also reduces the transferability of these models. In much of the developing world, vocational education is poorly defined, overlapping with workplace training and informal apprenticeships. Moreover, ROR analysis supposes that per-pupil government spending can be interpreted as social investment, which should be compared based on individual (wage) returns or aggregate income returns. Regional labour markets and cultural contexts provide a fuller understanding of the varied ways in which education and training can contribute to equity, employability, and inclusive economic development (McGrath, S., & Yamada, 2023). This paper underlines a highly essential context-responsive, as well as comprehensive and labour market-oriented educational and skill training approaches. The issue is to bridge regional inequalities alongside socio-economic variations in access and outcomes to change the Indian demographic dividend into equal and sustainable economic growth.

Regarding Indian socio-economic typology, there are a number of important demographic groups that people use acronyms to denote. The other vulnerable groups defined by the Constitution of India as having a historical disadvantage and thus targeted in the formulation of affirmative action and welfare policies include the Scheduled Castes (SC) and the Scheduled Tribes (ST). Socially and educationally backward classes that also get some benefits under government schemes are termed as Other Backward Classes (OBC). Persons with Disabilities (PWD) is a group referring to those people with a physical, mental, sensorial or cognitive impairment which can impede full and effective engagement in the society. Further, Economically Weaker Sections (EWS) is also a category introduced to give reservation benefits to persons belonging to general category, on the basis of fulfilling a set of income limit and assets properties with the aims of ensuring equity in education and employment opportunities.

2. Literature Review

Educational infrastructure and outputs are fundamentally class, and capitalism is a class society where the bourgeoisie, who have economic and institutional power, are just appreciated. In such an environment, educational privilege becomes a luxury of the few, and those on the margins of society are left without access to vital educational tools, particularly in the developing world. Education affects the position in the economic system as well as guarantees social status and power. As a result, those higher in socioeconomic class are more likely to have access to better schools, better curriculum, and better institutional support, validating previous social stratification. Accordingly, while inequalities in educational achievements between Indian states and socio-economic groups may be interrelated to some extent by systemic inequities, locationality, and cultural capital, effective policy responses seeking to ensure equitable access and outcomes need to be place-based (Kumar et al., 2024). Public policy has a significant role to play in connecting education and the requirements of the labour market, and in responding to deep-seated social inequalities. In case of a fast-changing economy, such as India's, when education systems are aligned with the world of skills, such strata, including those hailing from marginalized communities, can bring to fruition the dreams of gainful employment they have. A few public policies that align with the Indian labour market and education are SIM 2015, NEP 2020, DDU-GKY, NAPS, SANKALP, and STRIVE. According to (Skillindiamission.in, 2025), Skill India Mission (SIM) is a well-planned and well-equipped institution which is in the nature of a new horizon in the sky of knowledge in the field of technical education. There is only one motive for the SIM, and that is to wake people up to the world of technical science, which is not scary and difficult. The New Education Policy 2020 seeks to overhaul the education system to meet the needs of the 21st century by making vocational education part of mainstream education. By 2025, it aims to make vocational training available to at least 50 per cent of students, from middle school, to reduce the skill gap, dignity of labour, and eliminate the societal stigma associated with vocational education (Education.gov.in, 2025). As per the statement of (Rao, 2023), the National Apprenticeship Promotion Scheme (NAPS) is an effort by the Government of India to boost Apprenticeship training in the country. It offers partial stipend support to incentivize employers, especially small businesses, to hire apprentices, especially in disadvantaged areas. It is focused on creating an industry-ready workforce through on-the-job training and aims to skill 46 lakh apprentices by 2025. The program also connects skill training at the industry level with jobs as well as national infrastructure projects, providing a monthly stipend to apprentices as a means to ease the school-to-work transition. The Skills Acquisition and Knowledge Awareness for Livelihood Promotion (SANKALP) programme has been implemented by the Ministry of Skill Development and Entrepreneurship (MSDE) since 2018. As part of India's wider policy approach to ensure education responds to the demand for labour services, SANKALP is instrumental in enhancing decentralized, demand-centric training systems. But, its application has experienced significant interstate differences on account of differences in governance capabilities, capacity of institutions, and public-private collaboration. It is believed that states such as Maharashtra and Gujarat have been able to do so effectively with the aid of strong local participation models (Rose & Vengatesh, 2022). As per the statement of (Swain & Swain, 2020), the DDU-GKY was initiated by the Government of India on Antyodaya Diwas, 25th September 2014, under the Ministry of Rural Development. Its main purpose is to support rural youth from poor families (especially those aged 15-35 years) so that they can be equipped for the industry with marketable training.

According to the India Skills Report 2022, only 46.2% of graduates were employable from the perspective of recruiters (Economictimes.indiatimes.com, 2021). Other countries, such as Germany, build their educational systems with vocational education and academic pathways, though in India's case, VET is often identified as being subordinate. A lack of support by the government through uneven expenditures on education and skill development across the states contributes to differences (Nepram et al., 2021). States such as Kerala and Tamil Nadu are ahead in this area with better alignment between skill-based and job-oriented programs, while states including Bihar and Jharkhand are severely underdeveloped when it comes to skill-based program alignment. There is growth in female enrolment in STEM disciplines, yet the number of women employed in core engineering sectors is about 14% of the workforce (Chowdhury, 2022). Female labour force participation in India is about 25% while the level of educational attainment is steadily increasing (Bano, 2025).

According to the report of the Ministry of Skill Development & Entrepreneurship (MSDE), the age dividend of India, with the youngest population of the world, and a median age of 28. India has an enviable

demographic advantage as India has a young population with little more than 65% below the age of 35, which is a challenge and an opportunity for skill development. The positive change is that the percentage of demand for school education from the youth population of 10-24 years for skill training has increased from just under 34% to 51.3% during the last ten years (Pib.gov.in, 2024). Furthermore, policies for women-specific entrepreneurship, including mentorship, financial inclusion, and access to support for business development, are imperative to mitigate gender disparity while improving inclusivity (Guelich & Guelich, 2024). A number of developments in India, such as Startup India, have demonstrated potential yet are not gender-specific in the mentorship and follow-up elements. An existing study confirmed that the review of rural skill training, as in NRLM schemes, women's income and agency were significantly improved with integrated gender-responsive planning (Shukla & Phookan, 2023). Another study noted the need for fiscal policies and labour market contextualization to address gender-differentiated impacts (Culotta, 2023). This analysis resonates with Gender and Development (GAD) Theory, which stresses addressing the structural inequalities and power relations that restrict women's access to resources, opportunities, and decision-making. This includes creating policies that transform systems to ensure these policies do not reinforce gender disparities in employability outcomes. The inclusion of gender-responsive frameworks, skill-building policies, institutional practices, and funding mechanisms could help create an inclusive and equitable workforce.

The recognition and directing efforts towards vulnerable youth groups are critical to enabling vocational education and training (VET) programs to reach the target groups. Vulnerable groups, including school dropouts, rural youth, tribal communities, and economically disadvantaged populations, experience exclusion from mainstream education and capacity-building initiatives due to their marginalized socio-economic conditions. VET programs that account for the socio-economic realities of youth lead to improved rates of participation, retention, and employment outcomes (Shi & Bangpan, 2022). Tribal youth have limited training opportunities in terms of a lack of access as a result of language barriers, congestion of training and services infrastructure in Jharkhand and Odisha. Initiatives that involved integrated language and local content as a component of the VET model have resulted in higher engagement rates of indigenous women developmentally, while suggesting the relevance of culturally based perspectives (Dagar, 2022). It has been observed that helpful labour market information offered to youth before enrolment improves linkages between training and employment (Chakravorty et al, 2024). Youth dropout rates and job placements improved significantly after they were provided with clear information about industry needs, course relevance, and wage information in Gujarat. One of the recurring financial barriers to completion is that vulnerable youth are often unable to continue their VET training, which necessitates targeted financial literacy and support systems (Choudhary & Jain, 2023). The stipend support and training on opening bank accounts have improved completion rates of different courses among female students in VET programs within Bihar (Mathur et al., 2022). A data-driven vision towards regional labour demand-based target gender disparities, education levels, and socio-economic status (Carranza & McKenzie, 2024). Therefore, a targeted region-wise strategy needs to be adopted in consideration of the intersecting vulnerabilities of the young generation to mainstream vocational training effectively and inclusively in India.

Existing studies offer numerous and useful perspectives on high-level issues of employability, such as curriculum deformity, exposure to industry, and lack of training in soft skills in the Indian Higher education space [(1) (16)]. Although these studies lack intersectionality of gender and regional disparities in higher education and vocational training systems. There is a clear gap in research exploring comprehensively how states in India operationalize inclusive gender sensitive frameworks within skill development policies. There is a lack of research examining long-term employment outcomes from vocational training connected to formal education or training systems. Therefore, there is a critical need for studies that mitigate these gaps by concentrating on state-specific and gender-sensitive analyses that align education with employability in a diversified socio-economic environment such as India.

The literature reviewed stresses the important link between access to education, skill formation, and socio-economic development in the Indian context. It signifies how education and skill training of the Indian ecosystem are strongly embedded in structural contingencies of class, region, gender, and institutional capability. Educational resources and achievements are inherently determined by socio-economic hierarchies, and systemic inequities are proven to be persistent for generations and over geographical areas. Although the Central government has been trying to ensure that its flagship initiatives, such as SIM, NEP 2020, DDU-GKY, NAPS, SANKALP, and STRIVE, are aimed at the convergence of education with the labour market (Kumar, 2022) (Dewan & Khan, 2019). These initiatives have been operationalized with

varying effectiveness across states because of differences in governance capacities, infrastructure, and local partnership models. The extent to which vocational training is integrated into mainstream education and employment avenues varies across the country, with states like Kerala, Gujarat, and Maharashtra being ahead of others like Bihar and Jharkhand. The intersection of gender, region, and socio-economic status has not been adequately addressed in impacting access to vocational education. Accordingly, comparative, state-level, and intersectional research are needed to ascertain how education and skill training systems can be made more inclusive, gender-sensitive, and can adapt to regional labour market demands.

3. Methodology

Comparative Study Design

This study adopts a mixed-methods and comparative study approach to explore how different Indian states implement education and skill development, with a focus on inclusive and gender-responsive employability results. The analysis has also combined secondary quantitative data with qualitative document analysis to ensure both statistical authenticity and contextual depth. The states selected for comparison have presented varying levels of performance in education-employment linkage, ranging from higher-performing regions like Karnataka to lower-performing states such as Bihar, ensuring geographic, economic, and institutional diversity.

Quantitative Approach

Secondary quantitative data is presented from verified data sources, and this includes the All-India Survey on Higher Education (AISHE) for insights into enrolment patterns, institutional characteristics, and the availability of vocational education within higher education systems. The AISHE is a survey platform to cover all the educational institutions in India that offer higher education. Indicators of the student teacher ratio, institution density, gross enrolment, and the data obtained by AISHE are also considered in terms of gender parity index (Aishe.gov.in, 2025). The information on Enrolment, certification is also analysed with the help of the National Skill Development Corporation (NSDC) and placement records to achieve the goal of skill training programs. It is a public-private partnership organization in skill development in India, which is vital for the successful implementation of the government-funded scheme. NSDC also coordinates and improves the private sector initiative for skill development by providing them with financial support (Singla & Singh, 2020). AISHE is the combined initiatives of the Ministry of Human Resource Development, Inter-University Centre of the Planning Commission and National Informatics Centre (NIC) which is an important comprehensive source of information on higher education and is equipped with various indicators such as growth, access, and equity with emphasis on the diversity and the development of the higher education system, student profile, and the structure of the education network. These are important control variables in the panel data regression, which help to examine how patterns of educational infrastructure and participation impact trends in employability across the states in India over time. The Difference-in-Differences (DiD) approach, especially in the context of examining state-level interventions, implementation of DDU-GKY, STRIVE, or SANKALP schemes. These interventions, implemented over time and across different states, form the foundation of a treatment and control group to estimate a quasi-experiment DiD model. State-level data from the NFHS, the PLFS, and UDISE+ form a vital ingredient in the case of Education, Skill, and Employability composite indices that are assessed through PCA or Min-Max Normalization. NFHS also offers disaggregated socio-economic and demographic indicators pertinent to education access and employability (Girotra et al., 2023). PLFS provides key indicators of youth unemployment, labour force participation, and distribution by type of work (Abdul-Razak, S., & Sahoo, 2021), and UDISE+ supplies extensive school-level data concerning infrastructure, teacher qualifications, and availability of vocational training (Pritam & Shaikh, 2024). These multi-dimensional factors, normalized with PCA or normalization, facilitate robust index construction to compare them across states. The quantitative analysis is conducted with the help of R software, and this software is used to clean, organize, and analyze the gathered data. Descriptive statistics have helped to summarize basic trends, and regression analysis has been used to examine the relationships between education variables and employment outcomes across states.

Qualitative Approach

In parallel, qualitative analysis is also focused on reviewing the policy and institutional documents related to higher education and skill development. State-oriented plans, policy papers from the NSDC, Ministry of Skill Development and Entrepreneurship, and reports by national sources such as NITI Aayog and IHD are examined to understand the governance architecture, funding process, and implementation strategies. National Institution for Transforming India (NITI) Aayog is a policy think tank, which was established in

2015 to replace the Planning Commission. It offers both practical and visionary insights to the Government of India. This initiative of the government helps to shape India's development narrative and aligns all the policies with sustainable goals in a vibrant way (Hans, 2023). Particular attention has been given to public-private partnerships, state skill development missions, and bureaucratic capacity. Budget documents and program evaluations are also assessed to trace resource allocations and delivery models. Content analysis is used to identify different themes such as industry collaboration, interdepartmental coordination, institutional fragmentation, and inclusivity in design and implementation. Analysis from this document review is coded and thematically analysed to complement and explain the quantitative findings.

4. Results

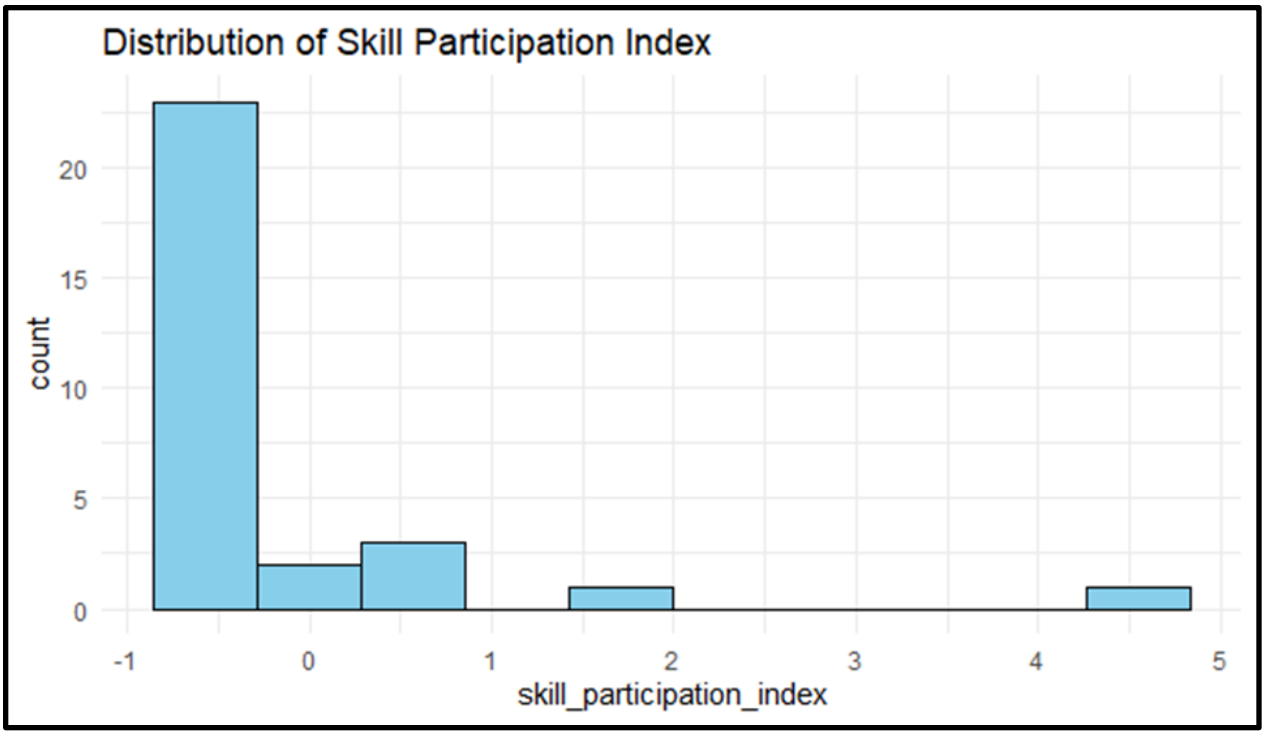
Descriptive statistics

Variable	Vars	N	Mean	SD	Median
Sl. No.	1	30	15.50	8.80	15.50
Programme*	2	30	15.50	8.80	15.50
Scheduled Caste (Male)	3	30	62,059.83	204,853.38	881.00
Scheduled Caste (Female)	4	30	58,982.00	200,155.00	703.50
Scheduled Tribe (Male)	5	30	25,479.10	88,612.38	286.50
Scheduled Tribe (Female)	6	30	25,190.77	88,426.40	219.00
OBC (Male)	7	30	146,691.13	415,054.26	2,684.50
OBC (Female)	8	30	139,374.43	418,541.29	2,547.50
PWD (Male)	9	30	802.27	2,299.46	12.50
PWD (Female)	10	30	663.07	1,939.58	8.50
Muslim (Male)	11	30	19,366.13	51,618.99	335.00
Muslim (Female)	12	30	20,721.57	66,098.02	377.00
Other Minority Communities (Male)	13	30	7,130.77	17,140.88	374.00
Other Minority Communities (Female)	14	30	7,764.53	19,453.55	402.50
EWS (Male)	15	30	4,776.20	11,288.54	144.50
EWS (Female)	16	30	4,649.60	11,942.89	121.00
SC_Total	17	30	121,041.83	404,721.73	1,602.00
ST_Total	18	30	50,669.87	176,931.76	483.00
OBC_Total	19	30	286,065.57	831,693.23	6,041.50
PWD_Total	20	30	1,465.33	4,221.69	23.00
Muslim_Total	21	30	40,087.70	116,922.23	759.50
Minority_Total	22	30	14,895.30	36,259.78	728.50

EWS_Total	23	30	9,425.80	23,022.13	254.00
skill_participation_index	24	30	0.00	1.00	-0.35
total_participation	25	30	522,186.07	1,579,339.10	10,234.00

Panel Data Regression

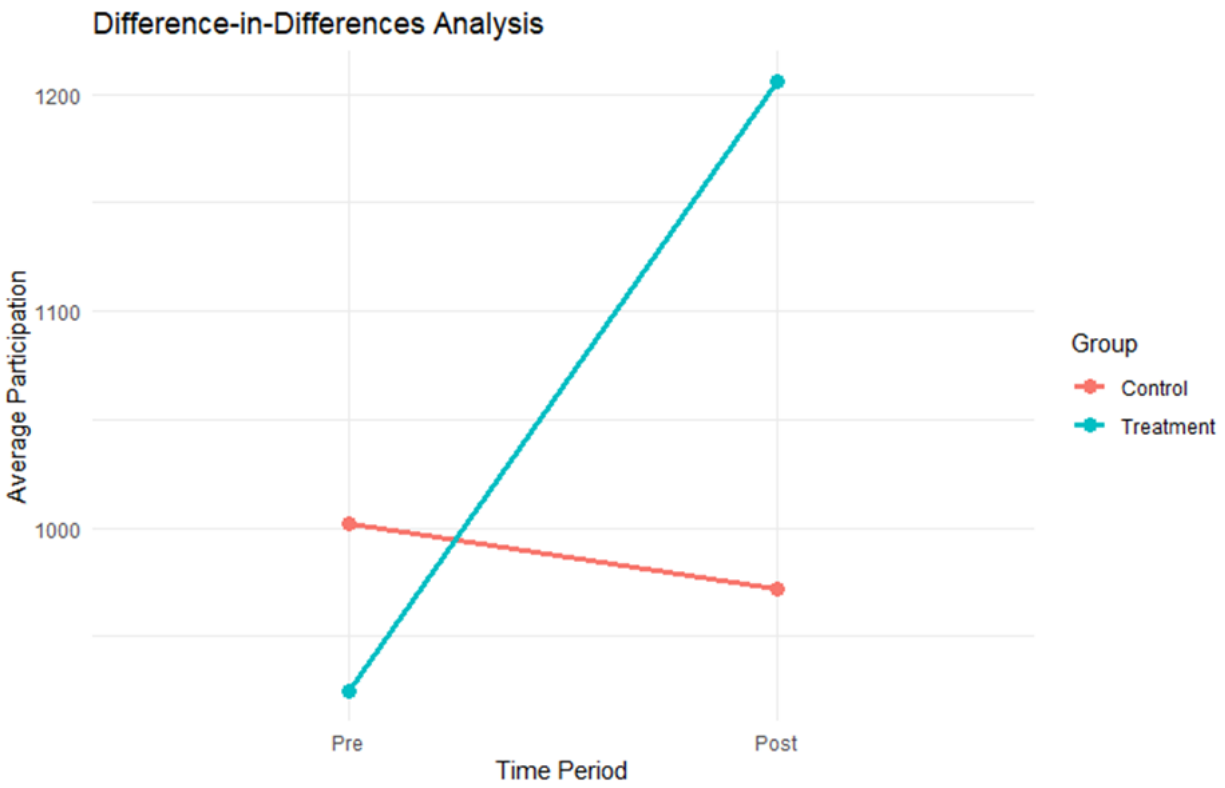
	SC_Total	ST_Total	OBC_Total	Pwd_Total	Muslim_Total	Minority_Total	EWS_Total
SC_Total	1.000000	0.993386	0.987640	0.980640	0.985813	0.884899	0.953161
ST_Total	0.993386	1.000000	0.971050	0.978428	0.977675	0.845215	0.930987
OBC_Total	0.987640	0.971050	1.000000	0.995405	0.992484	0.945405	0.984473
Pwd_Total	0.980640	0.978428	0.995405	1.000000	0.992484	0.929741	0.979767
Muslim_Total	0.985813	0.977675	0.992484	0.992484	1.000000	0.935231	0.975274
Minority_Total	0.884899	0.845215	0.945405	0.929741	0.935231	1.000000	0.966994
EWS_Total	0.953161	0.930987	0.984473	0.979767	0.975274	0.966994	1.000000



Aspect	**Description**
Research Objective	To assess the effect of social disaggregated indicators (SC_Total, ST_Total, OBC_Total, PWD_Total, Muslim_Total, Minority_Total, EWS_Total) on Skill Participation Index (SPI) across 3 Indian states over 10 years (n = 3, T = 10, N = 30).
Dependent Variable	Skill Participation Index (SPI)
Independent Variables	SC_Total, ST_Total, OBC_Total, PWD_Total, Muslim_Total, Minority_Total, EWS_Total
Analytical Method	Fixed Effects Panel Regression
Data Transformation	PCA and Min-Max Normalization used to construct SPI and policy direction indices
Purpose of Method	Controls for time-invariant state-specific heterogeneity (e.g., governance, institutional differences, historical exclusion)

R-squared	1.00 (indicating a perfect fit, but raises concerns of overfitting/multicollinearity)
F-statistic	5.33×10^{32} (extremely high and significant at $p < 2.22e\text{-}16$)
Multicollinearity Indicators	- SC_Total correlated 0.99 with ST_Total - OBC_Total correlated 0.995 with PWD_Total - All major variables had correlations > 0.95
Model Weaknesses	- Severe multicollinearity - Possible redundancy due to overlapping social group indicators - Overfitting due to perfect fit
Histogram Findings	SPI distribution is highly right-skewed; most values lie between -1 and 0, indicating low participation in most states
Index Outliers	Few high-performing states with SPI values of 4 and 5
Proposed Solution	Reduce redundancy by creating a composite "Inclusion Index" using PCA; use VIF diagnostics to detect and manage multicollinearity
Policy Implication	While targeting disadvantaged groups is comprehensive, inefficiencies and double-counting reduce effectiveness. State-specific strategies are necessary.
Interpretation of F-value	High F-statistic confirms joint significance of independent variables but also suggests overfitting or multicollinearity concerns

Difference-in-Differences (DiD) Linear Regression Model



Component	**Value / Description**
Method	Difference-in-Differences (DiD) – Quasi-experimental approach
Groups Compared	Treatment Group: Received intervention Control Group: Did not receive intervention
Pre-Treatment Observation	Control group had higher average participation than the treatment group
Post-Treatment Observation	Treatment Group: Large increase in participation Control Group: Small decrease in participation
Interaction Term (Treatment × Post)	Coefficient: 310.07 - Standard Error: 94.36 - t-value: 3.286 - p-value: 0.00176
Interpretation of Interaction	Treatment group saw 310-unit higher participation increase post-intervention compared to control

Baseline Control Group Mean	1001.67 ($p < 2e-16$) – Mean participation in control group before treatment
Treatment Group Difference at Baseline	-76.67 – Treatment group had 76.67 units lower participation at baseline than control group (had there been no treatment)
Post-Treatment Coefficient	-29.47 – Suggests a general decline in control group participation post-intervention, but not statistically significant (p = 0.66046)
Model Diagnostics	- Residual Std. Dev.: 182.7 - Multiple R^2: 0.2693 (≈27% variance explained) - Adjusted R^2: 0.2302 (moderate fit)
F-statistic	6.88 (df = 3, 56) - p-value: 0.0005017 Model is statistically significant
Program Effectiveness	Results show positive and statistically significant causal effect of the intervention on participation in the treatment group

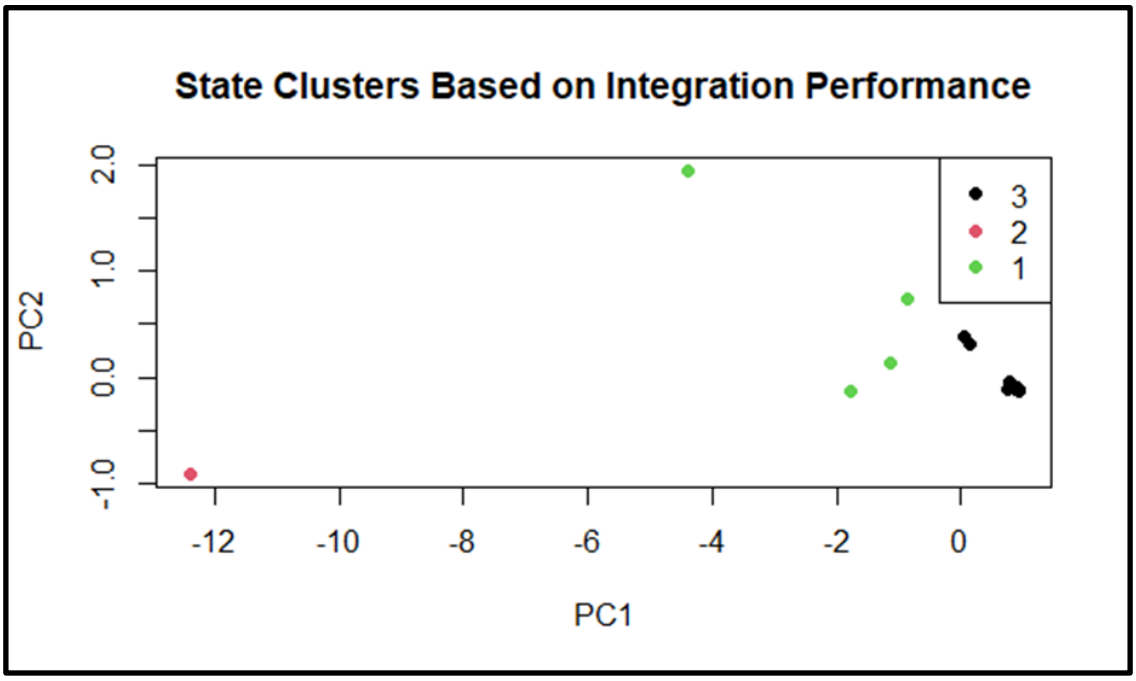
Principal Component Analysis (PCA)

Component	Standard Deviation	Proportion of Variance (%)	Cumulative Variance (%)	Interpretation
PC1	2.6026	96.77%	96.77%	Highly informative
PC2	(not specified)	2.82%	99.59%	Low contribution
PC3–PC7	(not specified)	< 0.25% each	~0.41% total	Likely noise or minor fluctuations

Insight	Implication
PC1 alone accounts for 96.77% of the total data variance	Suggests that socio-economic patterns are effectively one-dimensional

PC2 contributes just 2.82%, and PCs 3–7 add negligible variance	Including them in models may add complexity without meaningful improvement
Recommend using PC1 as a composite socio-economic index	Simplifies further analysis (e.g., regression, clustering) without loss of explanatory power
PCA provides a transparent and efficient dimensionality reduction strategy	Ideal for visualizations and econometric modeling using large, correlated datasets

Cluster Analysis



Cluster	Cluster Color	Key Characteristics	State Performance Level	Examples of Socio-Economic Indicators	Interpretation
Cluster 1	Green	States with low values across all socio-economic indicators: SC, ST, OBC, Minority populations	Low-performing / Least integrated	SC_Total, ST_Total, OBC_Total, and Minority_Total are all low	Indicates poor policy penetration, data underrepresentation, and low integration outcomes
Cluster 2	Red	Contains only one outlier state with	High-performing /	SC_Total = 2.17 million,	Suggests a large, diverse, better-

		very high totals in all marginalized groups	Diverse data-rich	OBC_Total = 4.26 million, Minority_Total = 154,707, Muslim_Total = 598,802	reported state with complex integration needs
Cluster 3	Black	States with moderate socio-economic disadvantage; reflective of partial policy inclusion and infrastructure development	Medium-performing / Transitional	SC_Total = 314,143, ST_Total = 128,915	Shows signs of policy intervention success, relatively better integration outcomes

5. Discussion and Findings

Discussion

This study aimed to understand how gender-responsive and inclusive education opens the gateways of employment in India. Various datasets are drawn from different sources such as AISHE, NSDC, NFHS, PLFS, and UDISE+. After the application of inferential statistics through SPSS and descriptive statistics, the study offers critical insight into gender, caste, and minority-based correlations and disparities in access to higher education. The Indian government has taken multiple important steps to implement skill development among youth. Through the "Skill India Mission" and "Samagra Siksha", the government projected the vocational courses as essential elements from secondary schooling onward (Hans, 2023). The analysis indicates large variations in the absorption of education, skill formation, and employment generation programmes across Indian states, and especially for the weaker sections of Indian society. Principal Component Analysis (PCA) indicates that a single principal component (PC1) represents 96.77% of the variance, indicating that the covariances of selected fundamental socio-economic indicators such as SC, ST, OBC, PWD, minority, EWS, and Muslim population are strong and can be encapsulated into a single composite index. This index was employed for the regional grouping of the states according to their integration ability. Three separate clusters were identified in the analysis. Cluster 1 represents states with low minority representation and low inclusive development performance. These can be states that have low population or reporting, or weaker implementation systems. These states include Nagaland, Mizoram, Sikkim, and Arunachal Pradesh. Cluster 2 is an anomalous cluster due to the high percentage representation within it across all categories, which may reflect a very crowded state with good data availability and a high burden of socio-economic inequality. The cluster 2 is an outlier cluster that consists of Uttar Pradesh as the only state with incredible indicator values around the concept of marginalized groups. Cluster 3 has all middle values of integration and very symmetric outreach to disadvantaged territories. This Cluster indicates states such as Jharkhand, Odisha, and Chhattisgarh. The clustering results reveal disparities in the state-wise movement towards more balanced growth and skill-education-employment alignment in India. These inequities emphasize the importance of multilevel policy processes, particularly offering targeted resources and facilitated capacity-building to lagging States explained in Cluster 1. High capacity or complex states explained in Cluster 2 need more granular, disaggregated policy instruments. The results also highlight the need for strong systematized data collection and performance monitoring to support implementation and accountability. Altogether, PCA and cluster analysis present a set of data-based landscapes to plan for more equitable and context-specific development interventions.

Findings

Dimension Affects Richness and Composite Index Efficiency

The PCA emerged with the first principal component (PC1) to capture 96.77% total variance, implying that the multiple socio-economic variables such as SC_Total, ST_Total, OBC_Total, PWD_Total, Muslim_Total, Minority_Total, and EWS_Total are highly interrelated. This is an indication that these variables move together and they indeed measure one common structure, the extent of social exclusion or inclusion at a state level. As other components explain only marginal fractions of variance, such as PC2 explains 2.82% more, the application of PC1 as a summary measure for integration performance is statistically appropriate and ergonomic. This reduction to a lower-dimensional space does not lead to a significant loss of structural information and allows for easy, meaningful comparison between states.

State-Level Clustering Reveals Sharp Disparities

The states were clustered into three clusters with different traits by the k-means clustering method based on the PC1 and PC2 scores. Cluster 1 includes states with extremely low concentration of under-represented groups in all variables, such as average SC is 8,183, ST is 2,209, and OBC is 27,412. These states may be low-population, or their low level of reporting indicates a need for focused administrative attention and capacity-building. Cluster 2 is dominated by high averages such as SC 2.17 million, OBC 4.26 million, and Muslims 598,802, representing overcrowded states with severe integration requirements. It is likely that these states host the maximum burden of the national inclusion task and need multi-level, scalable interventions to bring equity in education, skill formation, and employment. Cluster 3 indicates an equitable middle, with “middle of the road” values for each of the eight variables, such as SC 314,143, OBC 907,228, and EWS 33,587 centered with the average levels of participation framework. These states can act as models and provide ideas about replicable, scalable policies.

Policy and Equity Interventions

The implications of these clustering results are important for state-level policy. Cluster 1 states need serious preparatory work for outreach, data strengthening, and high-level institutional support. They also have difficulty accessing education and employment opportunities for marginalized populations. Cluster 2 states are statistically overrepresented, yet they remain at-risk sites for overgeneralized interventions, because of their variety and size, the same-size interventions are insufficient. Instead, hyperlocal targeting and incremental program design can work better. Cluster 3 states have achieved a level of policy success from which best practice can be built and transferred. For all clusters, the exercise highlights the importance of strong data ecosystems, decentralised governance, and focused, budgeted effort for achieving inclusive development. The need-based segmentation provides a basis for need knowledge by policymakers, which links access to education, skill formation, and employability for the deprived sections of society with the local needs of the state.

6. Conclusion

The present study aimed to assess the extent to which Indian states have been successful in integrating education and skill development mechanisms to improve employability. Based on secondary data from important inclusion categories such as SC, ST, OBC, PWD, Muslims, other minority communities, and EWS, the study used PCA and K-means clustering for data reduction and robust state grouping reflecting the degree of inclusiveness. PC1 has much greater driver capacity, explaining 96.77% of the variance of an entire dataset. This means that the socio-economic factors are strongly correlated and indicate a common dimension, probably related to a structural inclusion of vulnerable communities in education and skill-building. Cluster 1 represents states with inadequate coverage in terms of inclusion categories, reflecting small population bases or systemic exclusion. These states call for immediate policy attention to enhance outreach, institutional preparedness, and basic data collection capabilities. Cluster 2 consists of states that exhibited the most average representation across all categories. They are frequently the largest states by population, have diverse demographic populations with varied needs. Although they can get ample support from the Centre and the states, their scale requires hyper-local planning and execution. Cluster 3 includes states that are somewhat in between. They do not have the extremely high and low numbers present in the other two clusters. These states find themselves in a very mixed situation, sometimes with relatively good governance and moderate implementation structures that are functional to a reasonable extent, and also relatively inclusive infrastructure. They can be critical learning models, both for underachieving and pressure-cooked states. Documenting and replicating their successes can help individual countries to make substantial advances in their ability to integrate education and skill development.

The findings of this study essentially reinforce the idea that a uniform approach for educational and skill integration in India is inadequate. The heterogeneity of states in terms of populations, institutional capabilities, and historical traditions of inclusion demands a fragmented and variegated policy. Policies need to be sensitive to not only the demands of large demographics but also their particular education and labor market contexts. To build better data-driven governance from the state to the district. Without good, disaggregated, and real-time data determine where the gaps are or measure progress properly. Second, inter-sectoral coordination is critical. Aggregate platforms corresponding with school education, job training, and industry demand can facilitate the transition from schools to jobs. Third, equity should fall into design and delivery. Positive discrimination for SCs, STs, OBCs, PWDs, Muslims, and EWS should become part of the mainstream programs and should not be seen as separate welfare measures. Inclusion has to go from being the exception, which is targeted, to a principle that cuts across. The acquisition of skills must no longer be confined to technical and vocational education and training or further education and training, but must be included in the curricula at an early stage to encourage critical thinking, adaptability, and digital literacy. The New Education Policy (NEP 2020) has the potential to operationalize this shift, if its successful localization is possible.

7. Summary

The proposed research focuses on education and skills development integration in different states of India with the application of sophisticated quantitative tools, such as Fixed Effects Panel Regression, Difference-in-Differences (DiD) regression, Principal Component Analysis (PCA) and Cluster Analysis. The study forges ahead with an emphasis on the impact of socio-economic and demographic variables that include Scheduled Castes (SC), Scheduled Tribes (ST), Other Backwards Classes (OBC), Persons with Disabilities (PWD), Minority and Muslim Populations, and Economically Weaker Sections (EWS) on the Skill Participation Index (SPI) in the span of a decade in Indian states. The Fixed Effects model shows that even though a strong statistical fit is provided ($R^2 = 1.00$) due to uncontrolled multicollinearity among predictors, the results cannot be utilized in the inference. This uses PCA, with the first principal component (PC1) being able to justify a total of 96.77 percent of the variation, thus presenting the multidimensional data using a simplified and powerful representation. A 3-state cluster with PCA scores as a variable reveals that states that do not impact policies much and have low socio-economic indicators, such as Nagaland, Sikkim cluster into Cluster 1, outperforming Uttar Pradesh (Cluster 2), and decently integrated states with persistent issues, such as Jharkhand and Odisha cluster in the middle, in Cluster 3. The DiD regression also confirms the success of targeted interventions in terms of using skills. The causal effect of the program was confirmed in that the treatment group had higher participation after the intervention, but it was statistically significant. However, the research points to the usefulness of evidence-based methods in policy assessment. Applications of PCA-derived indices and DiD approaches are deemed helpful in the realignment of policies that can mirror the genuine distribution of inclusion, making educational and skill development planning a lot more strategic and intentional, across states of India. The methods are applied to model trends of participation over time, evaluate causal effects of treatments, simplify multivariate data by dimensionality reduction, and cluster states based on performance. This quantitative mixed method offers strong, policy-relevant conclusions with regard to structural inequality and program adequacy.

8. References

Abdul-Razak, S., & Sahoo, S. (2021). Using the rotational panel data from the Periodic Labour Force Survey: A cautionary tale. *Available at SSRN 3932499*. https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=3932499

Aishe.gov.in (2025). Introduction. Available at: https://aishe.gov.in/about-department/introduction/ [Accessed on 15 the July, 2025]

Bano, T. (2025). Female Labour Force Participation and Economic Growth in India: Analysis of the current scenario. *Idealistic Journal of Advanced Research in Progressive Spectrums (IJARPS) eISSN–2583-6986*, *4*(03), 109-115. https://journal.ijarps.org/index.php/IJARPS/article/view/722

Carranza, E., & McKenzie, D. (2024). Job training and job search assistance policies in developing countries. *Journal of Economic Perspectives*, *38*(1), 221-244. https://doi.org/10.1257/jep.38.1.221

Chakravorty, B., Arulampalam, W., Bhatiya, A. Y., Imbert, C., & Rathelot, R. (2024). Can information about jobs improve the effectiveness of vocational training? Experimental evidence from India. *Journal of Development Economics*, *169*, 103273. https://doi.org/10.1016/j.jdeveco.2024.103273

Choudhary, H., & Jain, H. (2023). Addressing financial exclusion through financial literacy training programs: a systematic literature review. *Empirical Research in Vocational Education and Training*, *15*(1), 8. https://doi.org/10.1186/s40461-023-00147-9

Choudhury, P. K., Joshi, R., & Kumar, A. (2023). Regional and socioeconomic inequalities in access to pre-primary education in India: evidence from a recent household survey. *International Journal of Child Care and Education Policy*, *17*(1), 13. https://link.springer.com/content/pdf/10.1186/s40723-023-00117-4.pdf

Chowdhury, D. (2022). Making higher education institutions gender-sensitive: Visions and voices from the Indian education system. In *Leading change in gender and diversity in higher education from margins to mainstream* (pp. 239-258). Routledge. DOI: 10.4324/9781003286943-19

Culotta, F. (2023). Gender-responsive regional fiscal policies: The labour market. In *Measuring Gender Equality: A Multidisciplinary Analysis of Some EU Countries* (pp. 177-205). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-41486-2_8

Dagar, P. (2022). Vocational education and training for indigenous women in India: toward a participatory planning approach. *International Journal of Training Research*, *20*(1), 43-57. DOI: 10.1080/14480220.2021.1959379

Dewan, S., & Khan, L. (2019). Breaking the cycle of vulnerability education, skills and employability for Indian youth. *JustJobs Network & UNICEF India, available at: https://demo. careleaverscommunity. org/wp content/files/reports_BreakingtheCycleofVulnerability. pdf*. https://justjobsnetwork.org/files/breaking-the-cycle-of-vulnerability-education-skills-and-employability-for-indian-youth_oct_2019.pdf

Economictimes.indiatimes.com (2021). *Indiatimes, "India's employability improves marginally to 46.2% from 45.97% last year", (2021)*. Available at: https://economictimes.indiatimes.com/jobs/indias-employability-improves-marginally-to-46-2-up-from-45-97-last-year/articleshow/88186260.cms?from=mdr

Education.gov.in (2025). Available at: https://www.education.gov.in/shikshakparv/docs/background_note_Reimagining_Vocational_Education_Skill_building_revised.pdf [Accessed on 26th July, 2025]

Girotra, S., Malik, M., Roy, S., & Basu, S. (2023). Utilization and determinants of adequate quality antenatal care services in India: evidence from the National Family Health Survey (NFHS-5)(2019-21). *BMC Pregnancy and Childbirth*, *23*(1), 800. https://doi.org/10.1186/s12884-023-06117-z

Guelich, U., & Guelich, H. M. (2024). Gender-responsive policies to support women start-ups to make a difference in the world. *Journal of Infrastructure, Policy and Development*, *8*(15), 10228. https://doi.org/10.24294/jipd10228

Gupta, P., & Datta, A. (2023). The role of accurate identification of vulnerable youth in vocational education and training systems for improved employability: Insights from experimental data. *Data in Brief*, *48*, 109258. https://doi.org/10.1016/j.dib.2023.109258

Hans, V. (2023). Niti Aayog. *Niti Aayog (December 5, 2023)*. https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=4653934

Kumar, A., Sharma, P., Patel, R., Gupta, A., & Singh, V. (2024). Educational Inequality and Its Impact on Social and Economic Opportunities in Rural India. *International Journal of Humanities, Management and Social Science (IJ-HuMaSS)*, *7*(2), 87-96. https://lamintang.org/journal/index.php/ij-humass/article/download/697/521

Kumar, R., Mandava, S., & Gopanapalli, V. S. (2019). Vocational training in India: determinants of participation and effect on wages. *Empirical Research in Vocational Education and Training*, *11*(1), 1-17. https://doi.org/10.1186/s40461-019-0078-y

Kumar, S. (2022). The skilling imperative in India: The bridge between women and work. *Observer Research Foundation Issue Brief*, (529), 3-18. https://www.orfonline.org/public/uploads/posts/pdf/20230814002546.pdf

Mathur, A., Sharan, M., Chakraborty, S., & Mullick, S. (2022). Technical and vocational education and training: Examining changing conditions in India. *Environmental Sciences Proceedings*, *15*(1), 31. https://doi.org/10.3390/environsciproc2022015031

McGrath, S., & Yamada, S. (2023). Skills for development and vocational education and training: Current and emergent trends. *International Journal of Educational Development*, *102*, 102853. https://doi.org/10.1016/j.ijedudev.2023.102853

Nepram, D., Singh, S. P., & Jaman, S. (2021). The effect of government expenditure on unemployment in India: A state level analysis. *The Journal of Asian Finance, Economics and Business*, *8*(3), 763-769. doi:10.13106/jafeb.2021.vol8.no3.0763

Parida, J. K. (2019). Rural-urban migration, urbanization, and wage differentials in urban India. In *Internal migration, urbanization and poverty in Asia: Dynamics and interrelationships* (pp. 189-218). Singapore: Springer Singapore. https://library.oapen.org/bitstream/handle/20.500.12657/50064/1/978-981-13-1537-4.pdf#page=194

Pib.gov.in (2024). Available at: https://static.pib.gov.in/WriteReadData/specificdocs/documents/2024/sep/doc2024911392001.pdf [Accessed on 26th July, 2025]

Pritam, B. P., & Shaikh, Z. (2024). Functional integration and utilization of technologies in school education with reference to regional disparity in India. *Discover Global Society*, *2*(1), 62. https://link.springer.com/content/pdf/10.1007/s44282-024-00087-z.pdf

Rao, C. R. National Apprenticeship Training Scheme: A Study of BOPT (ER) Performance (2023). https://doi.org/10.55306/GRJSSM.2023.120

Rose, A. M., & Vengatesh, T. P. CHAPTER–54 COMPARATIVE REVIEW OF STATE-LEVEL IMPLEMENTATION VARIATIONS IN THE SANKALP SCHEME. *Green Entrepreneurship Startups*, 396 (2022). https://www.researchgate.net/profile/Dr-j-4/publication/392810822_Role_of_SDGs_in_Entreprenuerial_Development/links/6853d6fedf3fa4286a41822b/Role-of-SDGs-in-Entreprenuerial-Development.pdf#page=433

Shi, Y., & Bangpan, M. (2022). Young people's participation experiences of technical and vocational education and training interventions in low-and middle-income countries: a systematic review of qualitative evidence. *Empirical research in vocational education and training*, *14*(1), 1-42. https://link.springer.com/content/pdf/10.1186/s40461-022-00136-4.pdf

Shukla, A., & Phookan, J. Gender Responsive Rural Livelihood Initiatives in India: A Review (2023). https://apijgs.com/fwdresearchjournal/01Gender%20Responsive%20Rural%20Livelihood%20Initiatives%20in%20India.pdf

Shukla, S. (2021). Employability of students in Indian higher education: challenges and opportunities. *Bhatt, AK, & Smita,.(2021). Employability of Students in Indian Higher Education: Challenges and Opportunities. Asian Journal of Economics, Finance and Management*, *3*(4), 10-17. https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=3842482

Singla, R., & Singh, B. PERFORMANCE EVALUATION OF NATIONAL SKILL DEVELOPMENT CORPORATION: AN ANALYTICAL STUDY (2020). https://www.academia.edu/download/114246571/download1.pdf

Skillindiamission.in (2025). Home. Available at: https://skillindiamission.in/home [Accessed on 15 th July, 2025]

Skillindiamission.in (2025). *Our Mission*. Available at: https://skillindiamission.in/ [Accessed on 26th July, 2025]

Swain, A., & Swain, S. (2020). Skill development in India: Challenges & opportunities. *International Journal of Scientific Research and Engineering Development*, *3*(6), 238-245. https://www.academia.edu/download/65090051/IJSRED_V3I6P23.pdf