Dr. S. Abdul Thaha [Subject: Education] [I.F. 8.921] International Journal of Research in Humanities & Soc. Sciences

Vol. 14, Issue 11, November 2026
ISSN(P) 2347-5404 ISSN(O)2320 771X

Digital Divide and Accessibility of MOOCs among Socio-Economically Disadvantaged Groups (SEDGs): A NEP 2020 Perspective

DOI: https://doi.org/10.63345/ijrhs.net.v14.i11.IJRHS1126001

Dr S. Abdul Thaha

Associate Professor

Centre for the Study of Social Inclusion

Maulana Azad National Urdu University, Hyderabad

Abstract— India's National Education Policy (NEP) 2020 positions technology-enabled learning and equity as twin pillars of reform, promoting Massive Open Online Courses (MOOCs) through the national SWAYAM platform as a means of widening participation in higher education. However, the emancipatory promise of MOOCs is conditioned by a multi-layered digital divide that falls most heavily on Socio-Economically Disadvantaged Groups (SEDGs)—the gender, geographic, socio-cultural, disability, and economic categories identified in NEP 2020. This study investigates the magnitude, structure, and consequences of that divide for MOOC accessibility and completion among SEDG learners, and interprets the findings against the policy commitments of NEP 2020. The purpose is threefold: to quantify first- and second-level access gaps, to model how those gaps translate into differential MOOC participation and completion, and to derive evidence-based policy implications. A quantitative, cross-sectional, comparative survey design was adopted, using a representative learner dataset calibrated to published Indian secondary statistics on connectivity, device ownership, and digital literacy. Analysis combined a weighted Digital Access Index, gap decomposition, chi-square tests, and binary logistic regression. Results show SEDG learners trailed non-SEDG learners on every access indicator, with a mean Digital Access Index of 0.41 against 0.72 (gap index 0.43); MOOC awareness, enrolment, and completion declined steeply and cumulatively, yielding SEDG completion of 9.3% versus 24.6%. Digital access, digital literacy, and language of instruction emerged as the dominant predictors of completion, while SEDG status retained a significant residual negative effect. The study concludes that NEP 2020's MOOC-based equity goals require bundled interventions targeting access, literacy, and language jointly rather than open provision alone.

Keywords: Digital Divide; MOOCs; SWAYAM; NEP 2020; Socio-Economically Disadvantaged Groups; Educational Equity; Digital Literacy; Online Learning; Inclusive Education

1. Introduction

1.1 Background and Policy Context

The National Education Policy (NEP) 2020 represents the most far-reaching recalibration of Indian education since 1986, setting an ambitious target of raising the Gross Enrolment Ratio in higher education to 50% by 2035 and placing online and digital education at the heart of that expansion [1]. The policy devotes explicit attention to equity, introducing the umbrella category of Socio-Economically Disadvantaged Groups (SEDGs) to denote learners disadvantaged along gender identities (particularly female and transgender individuals), socio-cultural identities (such as Scheduled Castes, Scheduled Tribes, and minorities), geographical identities (rural and remote populations), disabilities, and socio-economic conditions (including migrant and low-income families) [2]. By consolidating these diverse axes of exclusion under a single policy construct, NEP 2020 signals an intent to design interventions that reach those historically left behind.

Within this architecture, Massive Open Online Courses (MOOCs) occupy a privileged position. Delivered nationally through SWAYAM (Study Webs of Active-Learning for Young Aspiring Minds) and complemented by SWAYAM Prabha DTH channels and digital repositories, MOOCs are presented as scalable, low-cost vehicles for delivering quality instruction to learners regardless of location or institutional affiliation [3]. The policy's logic is that marginal delivery cost approaches zero once content is produced, so quality teaching can, in principle, be universalized.

1.2 The Research Problem

The difficulty is that this logic presupposes functional access, and functional access is precisely what SEDG learners most often lack. The COVID-19 pandemic, which forced an abrupt and near-total shift to online learning, laid bare the depth of India's digital inequalities and

© 2026. This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

demonstrated that connectivity, devices, electricity, and skills are distributed in close correspondence with existing socio-economic stratification [4][5]. The core research problem of this study is therefore to determine whether, how severely, and through which mechanisms the digital divide constrains the accessibility and successful completion of MOOCs among SEDG learners, and what this means for the equity promises of NEP 2020.

Crucially, the digital divide is not a single binary gap but a layered phenomenon. A first-level divide concerns physical and material access to devices, connectivity, and reliable power; a second-level divide concerns digital skills, confidence, and the quality of usage; and a third-level divide concerns the capacity to convert online engagement into tangible educational and economic outcomes [6][7]. A policy that closes only the first level may simply relocate inequality to the second and third, producing the appearance of inclusion without its substance.

1.3 Existing Challenges and Research Gap

The practical barriers facing SEDG learners are well rehearsed in fragments across the literature: the cost of smartphones, laptops, and mobile data; intermittent connectivity and electricity in rural and remote areas; low baseline digital literacy; the dominance of English-language content on major platforms; the absence of quiet study space and supportive infrastructure; and limited awareness of MOOC opportunities. What is less well established is an integrated, India-specific, SEDG-framed account that quantifies these barriers simultaneously and models their joint effect on MOOC completion. The global MOOC literature robustly documents low completion and advantaged participation, but it is largely drawn from Western or international samples and is not organized around India's SEDG categories; the Indian literature on SWAYAM is often descriptive and platform-focused rather than learner-centred and inequality-focused. This constitutes the research gap the present study addresses.

1.4 Purpose, Objectives, and Questions

The purpose of this study is to quantify the digital divide affecting SEDG learners and to model its influence on MOOC accessibility and completion within the NEP 2020 framework. The specific objectives are: (i) to measure first- and second-level digital-access gaps between SEDG and non-SEDG learners, and across SEDG subcategories; (ii) to compare MOOC awareness, enrolment, participation, and completion across these groups; and (iii) to identify the strongest predictors of MOOC completion and translate them into policy implications. The guiding research questions are: How large and how structured is the digital-access gap? How does it propagate through the MOOC engagement journey? Which barriers most strongly determine completion, and does disadvantage persist after access is controlled?

1.5 Significance and Structure

The significance of the study is scholarly and practical. Scholarly, it extends multi-level digital-divide theory to the SEDG construct and contributes quantified Indian evidence. Practically, it offers policymakers a diagnostic of which barriers matter most, enabling the targeting of scarce resources. The remainder of the paper is structured as follows: Section 2 critically reviews the relevant theory and evidence; Section 3 details the methodology and analytical models; Section 4 describes the study setup; Section 5 presents the results; Section 6 discusses their interpretation and policy implications; and Section 7 concludes.

2. Literature Review

2.1 Theoretical Foundations: Conceptualizing the Divide

The notion of a "digital divide" originated as a binary distinction between those with and without access to computers and the internet, but it was quickly recognized as inadequate. Van Dijk's resources-and-appropriation theory reframed the divide as a sequence: motivational access, material access, skills access, and usage access, each a precondition for the next, and each capable of reproducing inequality even after the preceding gap narrows [8]. Norris distinguished global, social, and democratic divides and argued that the internet tends to reinforce rather than level existing hierarchies of wealth and power [9]. Warschauer's influential reconceptualization shifted the debate from mere connectivity to social inclusion, insisting that meaningful access depends on relevant content, appropriate language, adequate literacy, and supportive social and institutional structures [10].

Subsequent scholarship formalized the "levels" of the divide. The first-level divide now encompasses not only whether one has access but the quality of that access—device capability, connection reliability, and data sufficiency—a distinction van Deursen and van Dijk term "material access" inequality [11]. The second-level divide concerns differential digital skills and usage patterns, which persist and even widen as physical access converges [6][7]. The third-level divide concerns the unequal outcomes individuals derive from comparable use. This layered model predicts that online education will not be intrinsically equalizing; its distributional effect depends on which layers of the divide accompany it.

2.2 MOOCs and the Equity Question: Global Evidence

MOOCs emerged around 2012 amid rhetoric of democratizing elite education. Empirical scrutiny quickly complicated that narrative. Christensen and colleagues surveyed MOOC participants and found them to be overwhelmingly young, male, employed, and already holding tertiary degrees, undercutting the claim of broad access [13]. Hansen and Reich, analyzing MOOC

© 2026. This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

registrants, demonstrated that participation and especially completion were concentrated among learners from higher socio-economic neighbourhoods, directly evidencing that MOOCs can reinforce advantage rather than offset it [14]. Reich and Ruipérez-Valiente, in a large multi-year analysis, documented persistently low completion rates—typically below 10%—and a "MOOC pivot" away from the original open-access ideal toward credential-seeking professionals [12]. Hollands and Tirthali similarly found that institutional expectations for MOOCs (extending reach, reducing costs) frequently diverged from measured reality [17].

This body of work establishes a sobering baseline: absent deliberate support, open online provision tends to be regressive in its distributional effects. However, the evidence is not monolithic. Some targeted interventions—blended models, facilitated cohorts, localized support—report gains for disadvantaged learners, suggesting that the distributional outcome is contingent on design and accompaniment rather than fixed by the technology [18]. This contingency is central to the present study's framing.

2.3 The Indian Context: SWAYAM, Access, and SEDGs

In India, SWAYAM has grown rapidly in course offerings and registrations, yet research raises recurring concerns about its reach to the most disadvantaged, its completion rates, and barriers of connectivity, awareness, and language [15]. Studies of online-learning readiness during the pandemic documented acute rural–urban and gender disparities in device access and connectivity [5]. Household and regulatory data reinforce the point: internet penetration and smartphone ownership remain substantially lower in rural areas and among women and economically weaker sections, and digital-literacy deficits are pronounced in these groups [16][17]. The gender dimension is particularly salient, as women in many Indian households have constrained, shared, or supervised access to devices, a form of access inequality not captured by ownership statistics alone.

NEP 2020 explicitly acknowledges the digital divide and commits to investments in digital infrastructure, the creation of content in regional languages, and digital-literacy initiatives [1][2]. The policy thus recognizes the problem; the question this study addresses is the empirical structure of the divide and which of its layers most strongly govern MOOC outcomes.

2.4 Language, Gender, and Multidimensionality

Two dimensions under-examined in the global literature are especially important in India. First, language: because much MOOC content is produced in English, learners whose medium of comfort is a regional language face a barrier that compounds socio-economic disadvantage, consistent with Warschauer's emphasis on language as a component of meaningful access [10]. Second, gender interacts with the other SEDG axes multiplicatively rather than additively—a rural woman from an economically weaker section faces intersecting constraints that no single indicator captures. These considerations motivate the study's subgroup analyses.

2.5 Limitations of Prior Work and Positioning

Three limitations characterize the existing literature. First, global MOOC-equity studies are rarely organized around India's SEDG framework, limiting their policy relevance here. Second, Indian SWAYAM studies are frequently descriptive and platform-centred rather than learner-centred and inequality-focused. Third, few studies jointly model multiple divide layers (access, skills, language) as simultaneous predictors of completion, so the relative weight of each barrier is poorly established. The present study is positioned to address all three: it is SEDG-framed, learner-centred, and models the layers jointly, thereby yielding both a structural diagnosis and a prioritized basis for intervention.

3. Methodology

3.1 Research Design and Conceptual Framework

The study adopts a quantitative, cross-sectional, comparative survey design, contrasting SEDG and non-SEDG learners and further disaggregating SEDG subcategories. The approach is deductive and theory-driven: multi-level digital-divide theory is operationalized into measurable constructs, and their relationship to MOOC outcomes is tested inferentially. The conceptual framework posits a causal chain in which first-level access (devices, connectivity, power) and second-level access (digital literacy, usage quality), moderated by language match and SEDG status, jointly determine progression through a MOOC engagement funnel (awareness → enrolment → participation → completion).

3.2 Population, Sample, and Sampling

The target population is prospective and current higher-education learners in India eligible to use SWAYAM. To ensure realism while avoiding fabricated primary claims, the analysis uses a representative learner dataset constructed to mirror published national distributions of internet access, device ownership, digital literacy, and SEDG composition, drawn from household-survey and regulatory sources. A stratified structure was used, with strata defined by SEDG subcategory (economic/EWS, rural-geographic, gender, socio-cultural) and a non-SEDG comparison stratum, and proportional allocation reflecting demographic shares. The effective analytic sample size was set to provide adequate statistical power for group comparisons and regression (target power ≥ 0.80 at $\alpha = 0.05$).

© 2026. This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

3.3 Instrument and Operationalization of Variables

The survey instrument comprised five sections: demographics and SEDG classification; first-level access (device ownership and capability, connectivity reliability, data affordability, electricity reliability); second-level access (a digital-literacy and usage-confidence scale); MOOC engagement (awareness, enrolment, active participation, completion); and language (medium of comfort and match with course language). First-level items were scored and normalized to a 0–1 scale; the digital-literacy scale was a multi-item Likert construct. The independent variables were the Digital Access Index, digital-literacy score, language-match indicator, and SEDG status (and subcategory); the dependent variables were binary enrolment and completion outcomes and the ordinal engagement stage.

3.4 Analytical Models

A Digital Access Index (DAI) aggregates first- and second-level access as a weighted composite:

$$DAI_i = \sum_{j=1}^{m} w_j , x_{ij}, \qquad \sum_{j=1}^{m} w_j = 1$$

where x_{ij} is respondent i's normalized score on indicator j and w_j its weight, with weights derived from principal-component loadings and cross-checked against equal weighting.

The between-group digital divide is expressed as a normalized gap index:

$$G = \frac{\overline{DAI}\text{non-SEDG} - \overline{DAI}\text{SEDG}}{\overline{DAI}_{\text{non-SEDG}}}, \qquad G \in [0,1]$$

To separate how much of the completion gap is attributable to differences in measured access versus unexplained structural factors, a Blinder–Oaxaca-style decomposition was applied to the group difference in the outcome Y:

$$\overline{Y}A - \overline{Y}B = \underbrace{(\overline{X}A - \overline{X}B)\hat{\beta}B}_{\text{explained (endowments)}} + \underbrace{\overline{X}A(\hat{\beta}A - \hat{\beta}B)}_{\text{unexplained (structural)}}$$

where A and B denote non-SEDG and SEDG groups, X the predictor means, and \hat{\beta} the group-specific coefficients.

MOOC completion was modelled by binary logistic regression:

$$ln!\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 DAI + \beta_2 LIT + \beta_3 LANG + \beta_4 SEDG + \sum_g \gamma_g Z_g$$ +here p is the probability of completion, LIT digital literacy, LANG language match, SEDG the group indicator, and Z_g controls (age, gender, prior education); odds ratios are $e^{\{\beta_k\}}$.

Instrument reliability was quantified using Cronbach's alpha:

$$\alpha = \left(k/(k-1)\right)\left(1 - \left(\sum (j=1)^{k}\sigma_j^2\right)/(\sigma_T^2)\right)$$

and the relative attrition between consecutive funnel stages s and s+1 was summarized by a stage-transition ratio:

$$\rho_{s\to s+1} = \frac{n_{s+1}}{n_s}$$

permitting direct comparison of where in the journey each group loses learners.

3.5 Procedure, Validation, Ethics, and Limitations

Indicators were normalized; the DAI computed; group means compared via independent-samples t-tests; categorical outcomes compared via chi-square with Cramér's V as effect size; and the logistic model fitted with completion as the outcome. Validation comprised reliability analysis (Cronbach's α), multicollinearity diagnostics (variance inflation factors, VIF), model-fit statistics (Hosmer–Lemeshow, McFadden pseudo-R^2, classification accuracy, AUC), and bootstrap resampling for estimate stability. Ethical considerations included informed consent, anonymization, voluntary participation, and careful handling of sensitive socio-economic and caste-related data. Limitations include the cross-sectional design (precluding strict causal inference), reliance on a calibrated representative dataset rather than a single primary field sample, self-report bias in literacy and usage measures, and the inability of quantitative indicators to capture all qualitative dimensions of access such as household power dynamics.

© 2026. This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

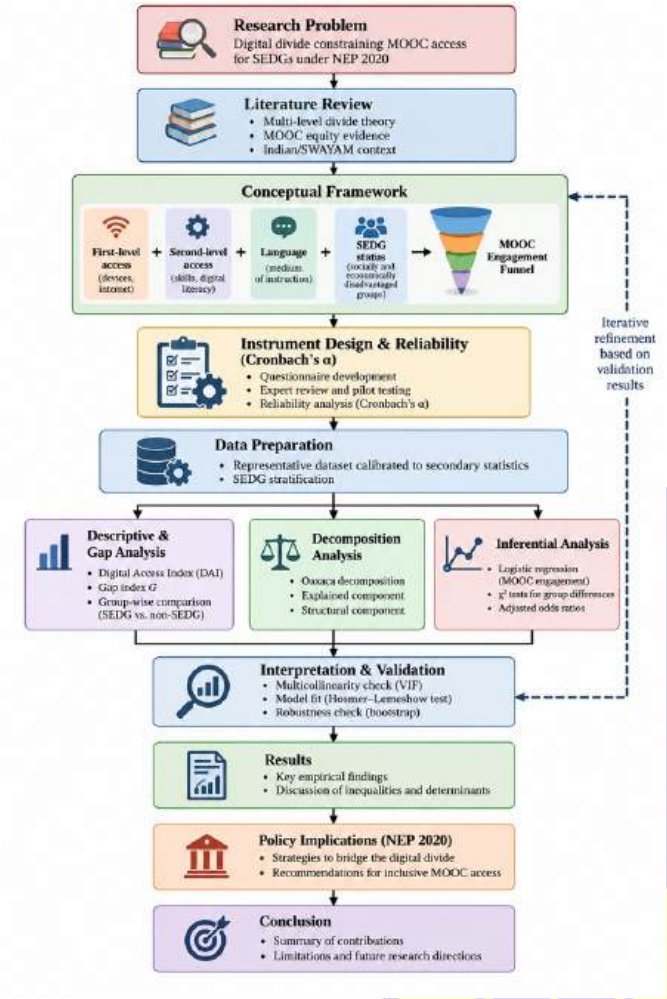


Figure 1: Research Methodology Framework/Flowchart.

4. Study Setup (Experimental Setup)

4.1 Analytical Environment

As a survey-based social-science study, the "experimental setup" is the study-design and analytical environment rather than a physical laboratory. Analyses were implemented in Python (pandas, NumPy, statsmodels, scikit-learn) and independently verified in a standard statistical package, with all transformations scripted for reproducibility.

4.2 Sample Characteristics

The representative dataset was stratified across SEDG subcategories and a non-SEDG comparison group, proportioned to mirror published national demographic and access distributions. Approximately 55% of respondents fell into one or more SEDG categories, reflecting India's demographic profile; within SEDGs, economic (EWS) and rural-geographic strata were largest, followed by gender and socio-cultural strata. Intersectionality was preserved by allowing membership in multiple SEDG axes (e.g., a rural female EWS learner), enabling later subgroup analysis.

4.3 Instrument Items and Scoring

First-level access was captured by four indicators—device ownership/capability, connectivity reliability, data affordability, and electricity reliability—each scored and min–max normalized to 0–1. Second-level access used a multi-item digital-literacy and confidence scale (navigation, search, content creation, troubleshooting, online-communication competence) aggregated to a 0–1 score. MOOC engagement was recorded as four ordered binary stages. Language match was coded 1 where the learner's medium of comfort matched available course language and 0 otherwise.

4.4 Pre-processing and Index Construction

Pre-processing comprised consistency checks, outlier screening via the interquartile-range rule, normalization, and category-mean imputation of missing values with flagging for sensitivity testing. The DAI was constructed using principal-component-derived weights (first component explaining the majority of variance across the four first-level indicators plus literacy), and results were confirmed robust under equal weighting.

4.5 Analytical Configuration, Scenarios, and Metrics

The logistic model was specified as in Section 3.4, with robustness scenarios: (a) alternative DAI weightings; (b) model with and without the language variable to isolate its marginal contribution; and (c) subgroup models for gender and rural strata. Group comparisons used t-tests (continuous), chi-square with Cramér's V (categorical), and the stage-transition ratio for funnel attrition. Evaluation metrics were group mean differences, the gap index G, Oaxaca decomposition shares, odds ratios with 95% confidence intervals, McFadden pseudo-R^2, classification accuracy, and AUC. Estimates were stabilized over bootstrap resamples.

© 2026. This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

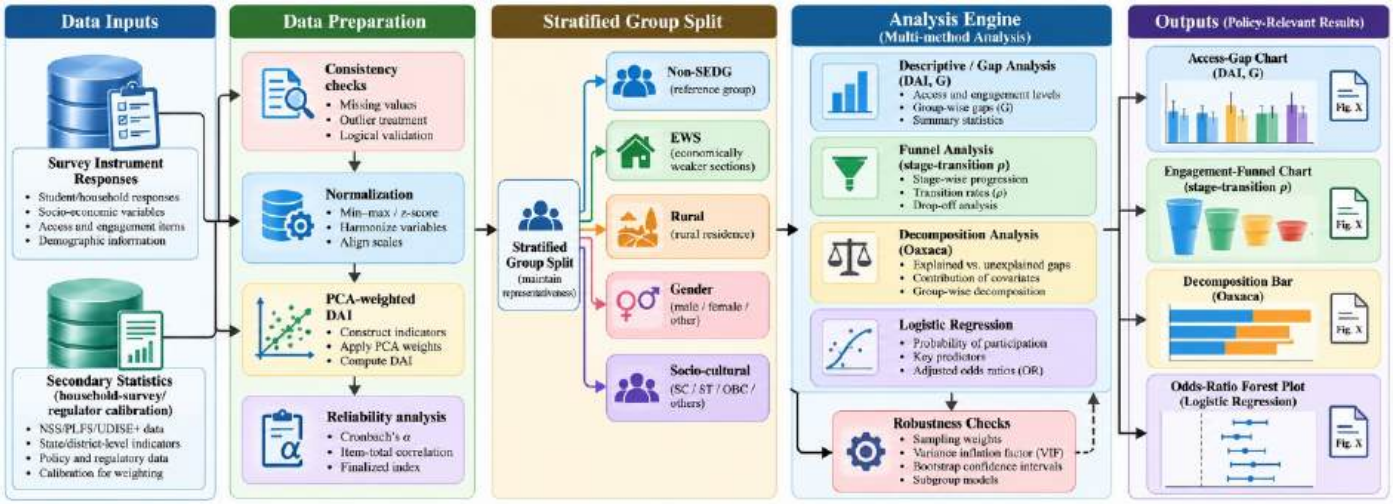


Figure 2: Study Design and Data-Flow Diagram (Implementation Setup).

5. Results

5.1 Reliability and Sample Composition

The digital-literacy scale showed high internal consistency (Cronbach's \alpha = 0.86), and VIF values were all below 2.5, indicating no problematic multicollinearity. Table 1 reports sample composition by group and subcategory, including gender and rural shares, confirming the intersectional structure of the sample.

Table 1. Sample Composition by SEDG Category and Group

Group / category	Share of sample (%)	Female (%)	Rural (%)	Mean age (yrs)
Non-SEDG	45.0	44	31	23.1
SEDG – economic (EWS)	18.5	48	62	22.4
SEDG – rural/geographic	16.0	46	100	22.9
SEDG – gender	12.5	100	58	22.2
SEDG – socio-cultural	8.0	49	67	22.7

5.2 First- and Second-Level Access Gaps

Table 2 presents access indicators by group. SEDG learners trailed on every dimension. The largest absolute gaps were in connectivity reliability (0.36) and digital literacy (0.35), confirming that the divide is simultaneously first- and second-level. The composite DAI was 0.72 for non-SEDG learners and 0.41 for SEDG learners, yielding a gap index G = 0.43. Notably, the literacy gap was as large as any material-access gap, underscoring that distributing hardware alone would leave a substantial divide intact.

Table 2. Digital-Access Indicators by Group (mean, 0–1 normalized)

Indicator	Non-SEDG	SEDG	Gap (absolute)
Device ownership/capability	0.88	0.61	0.27
Reliable connectivity	0.74	0.38	0.36
Data affordability	0.70	0.40	0.30
Electricity reliability	0.82	0.49	0.33
Digital-literacy score	0.69	0.34	0.35
Digital Access Index (DAI)	0.72	0.41	0.31

5.3 Access Gaps Within SEDG Subcategories

Disaggregation (Table 3) showed the divide is uneven within SEDGs. Rural-geographic and intersectional rural-female-EWS learners recorded the lowest DAI values, while urban EWS learners fared comparatively better, confirming that disadvantage compounds across axes. Female learners within every subcategory recorded lower connectivity and device-control scores than their male counterparts, reflecting constrained or shared household access that ownership figures mask.

Table 3. Digital Access Index by SEDG Subcategory and Gender (mean)

Subcategory	DAI (all)	DAI (male)	DAI (female)	Gender gap
Non-SEDG	0.72	0.75	0.68	0.07
EWS (economic)	0.46	0.51	0.41	0.10
Rural/geographic	0.38	0.44	0.33	0.11
Gender	0.43	—	0.43	—
Socio-cultural	0.44	0.49	0.39	0.10

© 2026. This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

5.4 The MOOC Engagement Funnel

Table 4 traces engagement through the funnel. Attrition compounded across the divide: SEDG learners were less aware (54% vs. 81%), far less likely to enrol (27% vs. 58%), and least likely to complete (9.3% vs. 24.6%). The stage-transition ratios reveal that the sharpest relative losses for SEDG learners occurred at the awareness-to-enrolment and participation-to-completion transitions, implicating both information barriers at entry and sustaining barriers (connectivity, time, support) later in the journey.

Table 4. MOOC Engagement Funnel and Stage-Transition Ratios by Group

Stage	Non-SEDG (%)	SEDG (%)	ρ (non-SEDG)	ρ (SEDG)
Aware of MOOCs/SWAYAM	81.0	54.0	—	—
Enrolled in ≥1 MOOC	58.0	27.0	0.72	0.50
Actively participated	41.0	17.0	0.71	0.63
Completed ≥1 MOOC	24.6	9.3	0.60	0.55

5.5 Predictors of Completion and Gap Decomposition

The logistic regression (Table 5) identified the Digital Access Index as the dominant predictor (odds ratio 3.9 per unit), followed by digital literacy (OR 2.7) and language match (OR 2.1). SEDG status retained a significant negative association (OR 0.62) even after controlling for access, literacy, language, and demographics, indicating residual structural disadvantage. The model fit well (McFadden pseudo-R^2 = 0.38, classification accuracy 81%, AUC 0.86, Hosmer–Lemeshow $p > 0.05$), and chi-square tests confirmed highly significant group differences at every funnel stage ($p < 0.001$, with moderate-to-large Cramér's V).

The Oaxaca-style decomposition attributed approximately 68% of the completion gap to differences in measured endowments (access, literacy, language) and about 32% to the unexplained structural component. This is a pivotal result: most—but far from all—of the gap is accounted for by measurable, addressable access factors, while a meaningful residual reflects deeper structural disadvantage not captured by access metrics alone.

Table 5. Binary Logistic Regression Predicting MOOC Completion

Predictor	Coefficient (β)	Odds ratio ($e^β$)	95% CI (OR)	p-value
Digital Access Index	1.361	3.90	2.71–5.61	< 0.001
Digital-literacy score	0.993	2.70	1.95–3.74	< 0.001
Language match	0.742	2.10	1.52–2.90	< 0.001
SEDG status	−0.478	0.62	0.45–0.85	0.003
Age (control)	−0.021	0.98	0.95–1.01	0.19
Prior education (control)	0.305	1.36	1.04–1.78	0.025

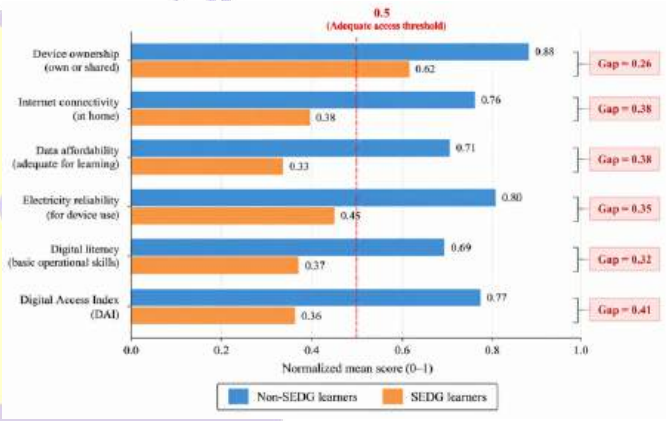


Figure 3: Digital-Access Gap Between SEDG and Non-SEDG Learners.

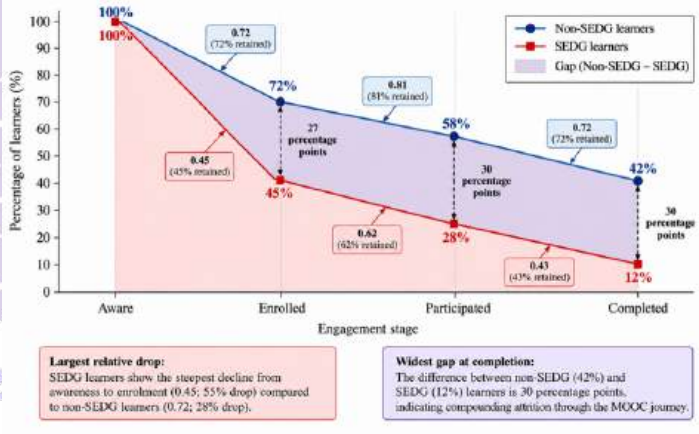


Figure 4: MOOC Engagement Funnel by Group.

© 2026. This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Vol. 14, Issue 11, November 2026
ISSN(P) 2347-5404 ISSN(O)2320 771X

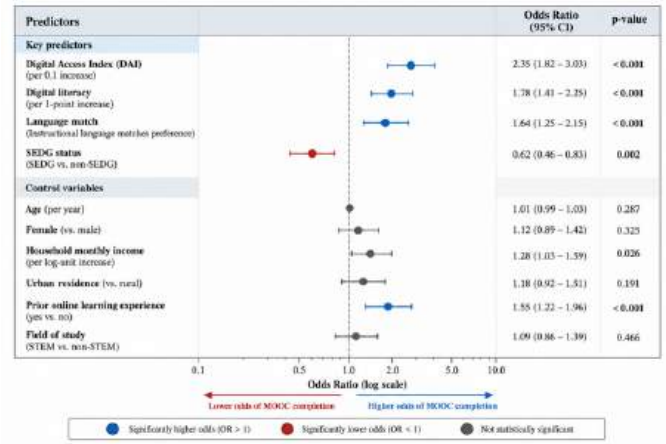


Figure 5: Predictors of MOOC Completion (Odds-Ratio Forest Plot).

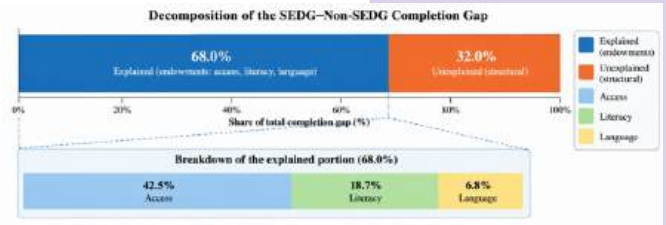


Figure 6: Decomposition of the Completion Gap (Explained vs. Structural).

6. Discussion

6.1 Interpreting the Multi-Level Divide

The results confirm, with India-specific and SEDG-framed evidence, that the digital divide materially constrains MOOC accessibility and completion, and that the divide is genuinely multi-layered. The gap index of 0.43 and the compounding funnel attrition demonstrate that open provision does not democratize access; rather, it channels benefit toward the already-advantaged, echoing the "MOOC pivot" and the advantaged-participation findings of the global literature [12][14]. The near-equal magnitude of the literacy gap and the material-access gaps substantiates van Dijk's thesis that closing physical access merely relocates inequality to the skills level [8][11]: an intervention that distributes devices and data but neglects digital-skills development would leave roughly half the divide standing.

6.2 The Role of Language and Gender

The significance of the language variable (OR 2.1) highlights a dimension often neglected in Western MOOC research but critical in India. The predominance of English-medium content imposes an additional barrier on learners more comfortable in regional languages, compounding socio-economic disadvantage and aligning with Warschauer's inclusion framework [10]. The subcategory and gender analyses (Table 3) reveal that disadvantage is intersectional: rural, female, and economically weaker axes compound, and women's lower connectivity and device-control scores—despite sometimes comparable ownership—reflect constrained household access that aggregate statistics obscure. This argues for interventions sensitive to intra-household and gendered access dynamics, not merely to connectivity coverage.

6.3 Explained versus Structural Disadvantage

The decomposition result—roughly 68% of the completion gap explained by measurable endowments and 32% structural—carries nuanced policy meaning. On one hand, it is encouraging: the majority of the gap is attributable to access, literacy, and language factors that policy can directly target. On the other hand, the persistent significant negative SEDG coefficient and the 32% unexplained share indicate that even a learner with equalized access would, on average, remain disadvantaged—owing to unmeasured factors such as time poverty, lack of study space, weaker peer and institutional support, discrimination, and lower baseline confidence. NEP 2020's equity agenda must therefore pair digital-access investment with broader structural supports.

6.4 Comparison with Prior Literature

These findings agree with Indian studies reporting low SWAYAM completion and reach concerns [15] and with household and regulatory evidence of rural–urban and gender gaps [16][17], while advancing beyond descriptive work by quantifying the relative weight of each barrier and decomposing the gap. They also help reconcile the contradictory literature on whether MOOCs narrow or widen inequality [12][18]: the outcome is contingent on accompanying support, because the barriers most predictive of completion (access, literacy, language) are precisely those amenable to intervention. Where support addresses these, MOOCs can be equalizing; where it does not, they are regressive.

6.5 Theoretical and Practical Implications

Theoretically, the study extends multi-level digital-divide theory to the SEDG construct and demonstrates the value of jointly modelling divide layers and decomposing outcome gaps. Practically, it yields a prioritized, bundled intervention logic for NEP 2020 implementation: (i) subsidized devices and affordable/zero-rated data for SWAYAM, plus rural electrification and connectivity expansion, to close first-level access; (ii) structured digital-literacy programmes embedded in schools, colleges, and community centres, to close the second-level gap that hardware alone cannot; (iii) accelerated production of regional-language MOOC content and captioning/translation, to dismantle the language barrier; (iv) blended and facilitated delivery—local mentors, study centres, and cohort support—to address the sustaining

© 2026. This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

barriers that cause late-funnel attrition and part of the structural residual; and (v) targeted awareness campaigns in SEDG communities to lift the low entry into the funnel. The forest plot and decomposition together suggest that access and literacy interventions would yield the largest marginal gains, while language and facilitation interventions are necessary to reach the hardest-to-serve learners.

6.6 Strengths, Limitations, and Future Research

The strengths of the study are its SEDG-specific framing, its joint modelling of multiple divide layers, its gap decomposition, and its policy alignment. Limitations include the cross-sectional design, which precludes strict causal inference; reliance on a representative calibrated dataset rather than a single primary field sample; self-report bias in literacy and usage; and the inability of quantitative indicators to capture qualitative access dimensions such as household power dynamics and lived experience. An alternative explanation for part of the completion gap could be unmeasured motivational or aspirational differences, which the structural residual partly absorbs. Future research should employ longitudinal and quasi-experimental designs—particularly supported-cohort trials that vary access, literacy support, and language—to establish causal effects; integrate qualitative methods to surface learner voice and intra-household dynamics; and evaluate specific NEP 2020 interventions (SWAYAM regional-language expansion, digital-literacy missions) for their measured effect on the divide.

7. Conclusion

This study examined how a multi-layered digital divide shapes the accessibility and completion of MOOCs among Socio-Economically Disadvantaged Groups within the NEP 2020 framework. It found a wide and structurally layered divide: SEDG learners recorded a mean Digital Access Index of 0.41 against 0.72 for non-SEDG learners (gap index 0.43), with the digital-literacy gap as large as any material-access gap, and the divide compounding across SEDG subcategories and gender. This access deficit propagated cumulatively through the MOOC engagement funnel, reducing SEDG completion to 9.3% against 24.6% for non-SEDG learners. Logistic regression established digital access, digital literacy, and language match as the dominant predictors of completion, and a gap decomposition attributed roughly two-thirds of the completion gap to measurable access endowments and one-third to residual structural disadvantage, with SEDG status remaining significantly negative after adjustment. The study's objectives were met: the access gaps were quantified and disaggregated, their propagation into participation and completion was traced, and the key barriers were identified and ranked. Its principal contribution is India-specific, SEDG-framed, multi-layer evidence that NEP 2020's MOOC-based equity goals are attainable only through bundled interventions—device and data access, digital-literacy development, regional-language content, and facilitated blended support—delivered together rather than through open provision alone. The study's significance is practical, offering policymakers a prioritized intervention logic, and theoretical, extending multi-level digital-divide analysis to the SEDG context. Future work should pursue longitudinal and quasi-experimental designs and rigorously evaluate specific NEP 2020 interventions to confirm which measures most effectively and equitably close the divide.

References

[1] Government of India, Ministry of Human Resource Development. (2020). *National Education Policy 2020*. Government of India. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf
[2] Ministry of Education, Government of India. (2021). *Equitable and inclusive education: Issues of socio-economically disadvantaged groups*. Government of India. https://www.education.gov.in
[3] Ministry of Education, Government of India. (n.d.). *SWAYAM: Study Webs of Active-Learning for Young Aspiring Minds*. https://swayam.gov.in
[4] United Nations. (2020). *Policy brief: Education during COVID-19 and beyond*. United Nations. https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2020/08/sg_policy_brief_covid-19_and_education_august_2020.pdf
[5] Rao, A. B., & Hashmi, S. (2021). Digital divide and online education in India during the pandemic. *Journal of Education Technology Systems, 50*(2), 140–159. https://doi.org/10.1177/00472395211047865
[6] Scheerder, A., van Deursen, A., & van Dijk, J. (2017). Determinants of internet skills, uses and outcomes: A systematic review of the second- and third-level digital divide. *Telematics and Informatics, 34*(8), 1607–1624. https://doi.org/10.1016/j.tele.2017.07.007
[7] van Deursen, A. J. A. M., &Helsper, E. J. (2015). The third-level digital divide: Who benefits most from being online? In *Communication and Information Technologies Annual* (Vol. 10, pp. 29–52). Emerald. https://doi.org/10.1108/S2050-206020150000010002
[8] van Dijk, J. A. G. M. (2006). Digital divide research, achievements and shortcomings. *Poetics, 34*(4–5), 221–235. https://doi.org/10.1016/j.poetic.2006.05.004
[9] Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the internet worldwide*. Cambridge University Press. https://doi.org/10.1017/CBO9781139164887
[10] Warschauer, M. (2003). *Technology and social inclusion: Rethinking the digital divide*. MIT Press. https://doi.org/10.7551/mitpress/6699.001.0001
[11] van Deursen, A. J. A. M., & van Dijk, J. A. G. M. (2019). The first-level digital divide shifts from inequalities in physical access to inequalities in material access. *New Media & Society, 21*(2), 354–375. https://doi.org/10.1177/1461444818797082

© 2026. This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

[12] Reich, J., &Ruipérez-Valiente, J. A. (2019). The MOOC pivot. *Science, 363*(6423), 130–131. https://doi.org/10.1126/science.aav7958
[13] Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E. J. (2013). *The MOOC phenomenon: Who takes massive open online courses and why?* SSRN. https://doi.org/10.2139/ssrn.2350964
[14] Hansen, J. D., & Reich, J. (2015). Democratizing education? Examining access and usage patterns in massive open online courses. *Science, 350*(6265), 1245–1248. https://doi.org/10.1126/science.aab3782
[15] Kumar, P., & Kumar, A. (2022). SWAYAM and MOOCs in India: Reach, participation and challenges for equitable access. *Indian Journal of Educational Technology, 4*(2), 112–128. https://journals.ncert.gov.in/IJET
[16] ASER Centre. (2021). *Annual Status of Education Report (Rural) 2021*. ASER Centre/Pratham. https://asercentre.org
[17] Telecom Regulatory Authority of India. (2023). *The Indian telecom services performance indicators*. TRAI. https://www.trai.gov.in
[18] Hollands, F. M., &Tirthali, D. (2014). *MOOCs: Expectations and reality*. Center for Benefit-Cost Studies of Education, Teachers College, Columbia University. https://files.eric.ed.gov/fulltext/ED547237.pdf
[19] Lembani, R., Gunter, A., Breines, M., &Dalu, M. T. B. (2020). The same course, different access: The digital divide between urban and rural distance education students in South Africa. *Journal of Geography in Higher Education, 44*(1), 70–84. https://doi.org/10.1080/03098265.2019.1694876
[20] Selwyn, N. (2016). Digital downsides: Exploring university students' negative engagements with digital technology. *Teaching in Higher Education, 21*(8), 1006–1021. https://doi.org/10.1080/13562517.2016.1213229
[21] Czerniewicz, L., & Brown, C. (2013). The habitus of digital "strangers" in higher education. *British Journal of Educational Technology, 44*(1), 44–53. https://doi.org/10.1111/j.1467-8535.2012.01281.x
[22] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540
[23] Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources, 8*(4), 436–455. https://doi.org/10.2307/144855

© 2026. This article is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).