



NEP 2020 promises AI-enhanced teaching, but Gurugram schools struggle to bridge policy and classroom reality

Sanjeev Kumar^{1*}, Himadri Shekhar Dey², Sachin Datt³

¹Research Scholar, School of Business, Sushant University, Gurugram, Haryana, India

²Assistant Professor, Department of Planning, School of Art and Architecture, Sushant University, Gurugram, Haryana, India

³Associate Professor, School of Design, Sushant University, Gurugram, Haryana, India

*Corresponding Author Email: sanjeev.bhardwaj9880@gmail.com

ABSTRACT

The study investigates the structure of the friction between the National Education Policy (NEP) 2020's policy requirement for incorporating AI into education at the national level and its practical realization by the relevant secondary school teachers in the city of Gurugram, Haryana. A convergent parallel mixed-methods approach was used to collect cross-sectional survey data from $N = 268$ teachers representing three socio-economic sectors of schooling (Elite Private, Budget Private, and Government) and analyse using CB-SEM and MG-SEM methods, in addition to the independent reflexive thematic analysis of qualitative data (semi-structured interviews; $n = 20$). As a result, apart from traditional UTAUT2 variables, a novel latent variable representing the Policy-Practice Alignment gap, as well as a newly developed TPACK-AI competence metric, was found to be the primary direct predictor of teachers' failure to use AI in their classrooms ($\beta = -0.423$, 95% CI $[-0.557, -0.289]$, $p < .001$), adding $0.09 \Delta R^2$ to the model above all other predictors at an individual level. Moreover, MG-SEM demonstrated scalar non-invariance in the sample, with a significant negative effect on the intercept of Facilitating Conditions in the Government group (-1.42 latent units; $p < .001$).

Keywords: Structural Equation Modeling; Technology Acceptance; UTAUT2; TPACK-AI; Mixed-Methods Research

1. INTRODUCTION

The rapidly changing landscape in the field of secondary education around the world is greatly characterized by the use of AI. Adaptation to AI technology, such as adaptive tutoring, content generation with large language models (LLMs), and automated formative assessment, is increasingly changing the face of teaching in schools (Bond et al., 2024; Ng et al., 2023; Ouyang & Jiao, 2021; UNESCO, 2023). Countries in the developing world, in particular, have been doing so through their own technology-based national policies.

The Indian NEP 2020 is an example of such an initiative. The document clearly advocates embedding concepts relating to artificial intelligence, machine learning, and computational thinking within secondary education curricula (Grade 9 to 12) and stresses personalized, competency-based formative assessments over rote learning (Ministry of Education, 2020). Nevertheless, implementing national policy directives in classroom settings in structurally diverse education systems is by no means easy. Educational sociologists refer to the phenomenon in question as the policy-practice gap: the discrepancy between grand visions at the macro-level and institutional realities at the micro-level (Ball et al., 2012).

This difference between technology mandates arises especially in cases where technology mandates presuppose infrastructure, agility, and pedagogic capacity that do not exist ubiquitously within the sectors of schooling. In India, secondary education is provided through an institutionally stratified structure in which elite private, budget private, and government schools function within dramatically different circumstances. Technology mandates that seek to integrate AI into all schools equally could inadvertently perpetuate educational inequalities in India (Mittal & Pani, 2022).

The city of Gurugram, Haryana, represents an excellent research location for the analysis of such issues. While Gurugram functions as a center of technological prowess and corporate success, the region also features a stratified school structure wherein elite private schools have their own budget for EdTech implementation, budget private schools utilize consumer-grade AI technology, and government schools utilize digital mandates from the state within constrained resources. These differences among schools functioning within the same district provide an excellent cross-sectional research setting for studying the issue at hand.

There have been extensive validations of the behavioral antecedents of teacher AI adoption intentions through predictive models such as UTAUT2 (Venkatesh et al., 2012) and TPACK (Koehler & Mishra, 2009). Nevertheless, these models are grounded in the assumption that teachers make their own adoption decisions and cannot accommodate the inhibitory effects of the structural misalignment of policies in mandatory public educational settings (Brown et al., 2002). Significantly, a systematic literature search from Scopus databases (keywords: "policy-practice gap" AND "structural equation model" AND "artificial intelligence" AND "education"; dates: 2010-March 2025; Databases: Scopus, Web of Science; N = 214 citations screened; N = 31 reviewed; N = 0 articles employing an operationalization of a PA construct as a latent SEM variable in the Indian secondary education context) revealed that no existing work has

measured the latent structural PA variable in an inhibitory capacity along with other behavioral predictors in this specific context.

This study bridges these gaps by framing the following research questions:

- RQ1: What is the extent of impact of core acceptance factors (PE, EE, FC) on Behavioral Intentions (BI) and Actual AI Teaching Behavior (AB) among teachers in secondary schools of Gurugram?
- RQ2: What is the impact of policy-practice gap perceptions in the context of NEP 2020 (PA) on actual classroom behavior (AB), controlling for individual acceptance-related determinants?
- RQ3: How do the paths of structural adoption and policy-practice gaps vary for the three school categories under the constraint of multi-group invariance?

This research is organized in the following manner: Section 2 summarizes literature review, provides theoretical background and hypotheses; Section 3 discusses the methodology and design of convergent parallel mixed methods approach; Section 4 describes research results both quantitatively and qualitatively; Section 5 discusses results in the context of global literature and provides divergent mixed methods results; and finally, Section 6 concludes.

2. LITERATURE REVIEW AND THEORETICAL FRAMEWORK

The incorporation of AI into secondary schooling environments has precipitated a paradigm shift from research on computer-enhanced learning towards studies involving complex algorithmic agency. There is systematic evidence suggesting that the use of technology enhanced through AI, including ITS, automated marking software, and text generators, can lessen administrative workload and promote competency-based learning (Bond et al., 2024; Zawacki-Richter et al., 2019). Yet, the adoption of such technologies within institutions is often inconsistent and influenced by factors such as regional sociotechnological environments, infrastructure differences, and teacher self-efficacy (Ng et al., 2023).

2.1 Theoretical Integration: TPACK-AI and UTAUT2

The Technological Pedagogical Content Knowledge (TPACK) model represents the overlap of content, pedagogy, and technology needed to effectively integrate digital technologies (Koehler & Mishra, 2009; Mishra & Koehler, 2006; Shulman, 1986). Within the context of generative and adaptive artificial intelligence, the term TPACK-AI refers to the ability of a teacher to critically assess the reliability of the content produced by AI technology and the application of AI tools to foster student inquiry in particular subject areas (Ng et al., 2023). The possession of high digital skills alone does not guarantee that a person possesses TPACK-AI.

These are incorporated in the framework used in UTAUT2 (Venkatesh et al., 2012). Three main constructs are focused on in this paper: Performance Expectancy (PE), which is how much the teacher expects AI to help increase his/her productivity; Effort Expectancy (EE), the perception that the incorporation of AI will be easy in teaching practices; and Facilitating Conditions (FC), the physical infrastructure and technical resources available to make AI possible. Behavioral Intention (BI) means the intention of a teacher to adopt and use AI in classroom teaching. Finally, Actual AI Teaching Behavior (AB) is about the frequency of the

use of AI tools in the teaching process. As far as the theoretical model goes, TPACK-AI is directly related to AB and acts as a predictor for it, whereas PE, EE, and FC influence BI as a mediator. The problem is that UTAUT2 was built for voluntary adoptions only; it needs further refinement to be used in mandatory adoptions within the public education system (Brown et al., 2002).

2.2 The Policy-Practice Gap as a Structural Friction Point

The critical weakness in conventional technology acceptance frameworks is the underlying premise that the adoption of the technology happens voluntarily through consumer decisions. However, in a public education setting, the adoption happens as part of a compulsory policy environment, which creates expectations for technological uptake from teachers irrespective of their preparedness levels and institutional capacity for supporting this process. According to Policy Enactment Theory (Ball et al., 2012), national-level policies are never enacted universally; rather, they undergo selective filtering, adjustment, and constraining due to material inequalities at the local level.

Despite the set AI integration targets in NEP 2020 (Ministry of Education, 2020), the implementation process faces several systemic barriers in the form of large class sizes, lack of professional development opportunities, and a long-standing summative exam culture (CBSE) based on rote learning rather than inquiry learning processes (Sharma & Singh, 2024). The comprehensive Scopus/WoS database search, detailed in Section 1, revealed no prior research that quantitatively operationalized the policy-practice gap into a psychometrically reliable structural concept using SEM. This study addresses the identified gap.

PA scale has been developed from three dimensions of the policy enactment framework of Ball et al. (2012), which include: (a) constraints from material context (measured by infrastructure deficit → PA1); (b) institutional regulatory context (measured by examination-driven syllabus rigidity → PA2); and (c) constraints from governance capacity (measured by absence of data governance in local context → PA3). These three dimensions of PA scale development have anchored the new scale in sociological theory with parsimony in line with CB-SEM methodology (Kline, 2023).

2.3 Hypothesis Development

Based on the integrated UTAUT2 + TPACK-AI + Policy Enactment framework, the following directional hypotheses are formalised:

- H1: Performance Expectancy (PE) positively affects Behavioral Intention (BI) in adopting AI-based teaching innovations.
- H2: Effort Expectancy (EE) positively affects Behavioral Intention (BI) in adopting AI-based teaching innovations.
- H3: Facilitating Conditions (FC) positively affect Behavioral Intention (BI) in adopting AI-based teaching innovations.
- H4: TPACK-AI positively affects Actual AI Teaching Behavior (AB).
- H5: Perceived Policy-Practice Gap (PA) negatively affects Actual AI Teaching Behavior (AB) substantially, even when controlling for individual-level effects.

- H6: Behavioral Intention (BI) positively affects Actual AI Teaching Behavior (AB).
- H7a: Coefficients of structural paths from UTAUT2 predictors (PE, EE, FC) to Behavioral Intention (BI) vary significantly in size by type of school (Elite Private, Budget Private, and Government).
- H7b: Latent factor intercepts for Facilitating Conditions (FC) and Policy-Practice Gap (PA) are not invariant across types of schools.

Figure 1 shows the entire structural model, complete with all its latent constructs, reflective indicators, directional hypotheses, and path labeling.

3. MATERIALS AND METHODS

3.1 Study Design and Participants

This study used a convergent parallel mixed methods design (Creswell & Plano Clark, 2018), where the collection and analysis of both quantitative survey and qualitative semi-structured interview data happened at the same time, being combined in the discussion stage. This decision was taken since research questions needed to be answered simultaneously quantitatively and qualitatively.

Quota sampling was employed using a ratio proportional to the school sectors in the districts of Gurugram (Government: 54%, Budget Private: 27%, Elite Private: 19%, Directorate of School Education, Haryana, 2024) in order to collect data from the three strata. Quotas included: Elite Private (n = 84), Budget Private (n = 92), and Government (n = 92). Schools were chosen based on the inclusion criteria of (a) offering Grades 9-12 in the CBSE system of education, (b) situated in the administrative boundaries of the district of Gurugram, and (c) having at least five secondary school teachers. Data from N = 268 secondary school teachers (Grade 9-12) was collected in the months of January-March 2025. 340 questionnaires were sent out, with 272 retrieved and 268 fully completed and usable (response rate: 78.8%). Non-response biases were checked via comparison between early and late responders on major demographic factors (Armstrong & Overton, 1977); none of the comparisons showed statistically significant differences (all $p > .10$). Sample size adequacy to detect medium effect sizes ($f^2 \geq 0.15$) was predetermined via Monte Carlo power analysis (MacCallum et al., 1996), with sufficient power (≥ 0.80) at $\alpha = .05$; $N \geq 250$ was required to achieve that. The observed $N = 268$ satisfies that requirement. Participation was completely voluntary, and written informed consent was obtained from all participants prior to data gathering. The study was approved by the Institutional Review Board of Sushant University.

Purposeful sampling of 20 teachers was done using semi-structured interviews, consisting of n = 7 Elite Private School, n = 7 Budget Private School, and n = 6 Government School teachers. The criteria used for selection guaranteed diversity in terms of their teaching experience (Early Career: n = 6; Mid-Career: n = 8; Senior: n = 6) and subject area (Sciences: n = 8; Humanities: n = 7; Mathematics: n = 5).

3.2 Quantitative Instrument

The survey questionnaire used scales measuring PE, EE, FC, and BI adapted from Venkatesh et al. (2003; 2012). The Actual Use of AI Teaching (AB) construct was assessed using three survey questions adapted from Tondeur et al. (2017), which measured respondents' frequency of and instructional extent of using AI technology in teaching currently. Survey questions for TPACK-AI construct were adapted from Ng et al. (2023). All survey questions were measured using a five-point Likert-type scale (1 = Strongly Disagree, 2 = Disagree, 3 = Neither Agree nor Disagree, 4 = Agree, and 5 = Strongly Agree). Finally, a new construct, named Policy-Practice Gap (PA), consisting of five survey questions, was created for this study using the method explained in Section 2.2 below.

Development of PA scale was accomplished following four steps, namely, (a) deductive item generation from policy enactment constructs of Ball et al. (2012) and content analysis of NEP 2020 mandate, resulting in eight items as potential candidates; (b) content validity analysis performed by three independent subject-matter experts in educational technology and Indian education policy, blinded to study hypotheses and evaluating items based on four-point Content Validity Index (CVI), resulting in rejection of items with CVI < 0.75 and selection of five remaining items; (c) iterative item refinement based on expert linguistic critique; and (d) pilot test conducted using 30 non-analytical teachers, producing acceptable item-total correlations ($r \geq 0.55$ for all selected PA items). In compliance with CB-SEM criteria of parsimony (Kline, 2023), three items were chosen for inclusion in the final PA scale. Scale items and standardised factor loadings derived from confirmatory factor analysis are shown in Table 1.

3.3 Qualitative Protocol

Semi-structured interviews took place through secure digital conferencing (60-75 minutes per interview) on three distinct themes, namely, (a) infrastructure and connectivity issues, (b) high-stakes board examination pressure on teaching flexibility, and (c) data governance and student privacy issues resulting from the implementation of AI tools. Each interview was recorded and transcribed verbatim by professional transcribers and verified by the participants (Lincoln & Guba, 1985). The extent of data saturation was determined based on iterative comparison of themes: Two independent coders reviewed the themes identified after each interview, and data saturation was defined as a less than two new themes across three interviews in succession. This occurred in the fifteenth interview; five subsequent interviews were confirmatory.

3.4 Data Analysis Strategy

Quantitative analyses were done by IBM SPSS v28 for descriptive statistics and lavaan package (v0.6) on R v4.3 for CFA, CB-SEM, and MG-SEM. The missing values ($n=4$, 1.2%) were imputed with listwise deletion due to its minimal percentage and proven MCAR (Little's MCAR test: $\chi^2(48) = 51.3$, $p = .34$). The fitness indices of measurement models were assessed based on predetermined standards of: CFI ≥ 0.95 , TLI ≥ 0.95 , RMSEA ≤ 0.06 , and SRMR ≤ 0.08 (Hu & Bentler, 1999). AVE ≥ 0.50 and CR ≥ 0.70 were used to test convergent validity (Kline, 2023). Discriminant validity was determined by HTMT criterion

(threshold value: 0.85; Henseler et al., 2015). The multivariate normal distribution assumption was verified by Mardia's coefficients (Mardia, 1970), thus supporting the maximum likelihood approach to estimation. The VIF score was checked for multicollinearity. Measurement invariance was established with respect to the Δ CFI metric (change of CFI ≥ 0.010 implies worsening model fit; Cheung & Rensvold, 2002). Non-invariant intercepts for the partial scalar invariance model were detected by means of the Lagrange multiplier (score) tests in lavaan.

In order to compute the incremental contribution of PA to the variance explained in AB, the constrained model (with the PA path set to zero) was fit and contrasted against the full model. The $\Delta R^2(AB)$ due to PA was calculated as the difference in $R^2(AB)$ between these two models.

Qualitative data were analyzed according to the six-phase reflexive thematic analysis approach developed by Braun and Clarke (2006), using NVivo v14 software. Inter-rater reliability was determined by means of double-coding of 30 percent of the corpus independently by another researcher specialized in Indian education policy, who achieved Cohen's Kappa ($\kappa = 0.83$), which surpassed the benchmark of 0.80. The integration of quantitative and qualitative streams was done using the convergent triangulation approach proposed by Fetters et al. (2013).

3.5 Phase 2 Cross-Validation Study (Pre-Registered Protocol)

This three-item scale for measuring PA that was designed and tested in the current study needs further independent cross-validation before its full-scale application as a standardised questionnaire. A Phase 2 cross-validation study has been registered on the Open Science Framework, and data collection is currently scheduled to occur in the Rewari district of Haryana, which is an almost exclusively rural area with very different infrastructure from that of Gurugram. Phase 2 research will enrol $N \geq 250$ secondary school teachers, both from public and budget private schools and use the same questionnaire and CFA measurement model to conduct a formal analysis of measurement invariance between Gurugram (Phase 1) and Rewari (Phase 2) samples. Prior to publication of the results of Phase 2, the PA scale must be considered provisionally validated in the urban setting of Gurugram.

Table 1: Psychometric Measurement Instruments and Confirmatory Factor Analysis Loadings

Construct	Item	Item Statement	Scale Source	λ
Performance Expectancy (PE)	PE1	Using AI technologies is a way to help increase my productivity in creating lesson plans based on NEP 2020.	Venkatesh et al. (2012); adapted	0.812
	PE2	AI technologies give me the opportunity to conduct personalized student evaluations faster.		0.794
	PE3	I can use AI to monitor performance differences between students enrolled in my class.		0.835
Effort Expectancy (EE)	EE1	Learning to work with AI platforms used for educational purposes is not complicated for me.	Venkatesh et al. (2012); adapted	0.756
	EE2	The interaction with AI-based assessment technologies is quite clear to me.		0.782
	EE3	It is not difficult for me to learn how to work with AI-based education programs.		0.801
Facilitating Conditions (FC)	FC1	My college is equipped with adequate technology for using AI technologies.	Venkatesh et al. (2003); adapted	0.844
	FC2	My institution provides me with technical assistance when I face any problems with AI technologies.		0.819
	FC3	My institution provides specific financial resources for learning how to use AI technologies.		0.852
TPACK-AI	TPA1	I know how to analyze the validity of AI-generated content and any associated algorithmic biases.	Ng et al. (2023); adapted	0.788
	TPA2	I can create inquiry-based lessons where AI platforms feature as part of my subject area.		0.824

Construct	Item	Item Statement	Scale Source	λ
	TPA3	I have the capacity to teach students about using AI technology in an ethical manner.		0.769
Behavioral Intention (BI)†	BI1	I presently make use of AI platforms in my classes for instructional or assessment tasks at least once a week.	Venkatesh et al. (2012); adapted	0.803
	BI2	I have designed at least one AI-supported lesson unit in line with NEP 2020 competencies this semester.		0.821
	BI3	The extent of the gap between NEP 2020 mandates for AI and my school's capabilities is significant.		0.796
Actual AI Behavior (AB)†	AB1	I personally utilise AI technology in the classroom at least once per week for either teaching or evaluation purposes.	Tondeur et al. (2017); adapted	0.779
	AB2	This semester I have created a lesson module utilising AI that fulfils the competency goals set out by NEP 2020.		0.814
	AB3	AI content creation and feedback mechanisms are employed during classroom activities.		0.768
Policy-Practice Gap (PA)*	PA1	The distance between NEP 2020's vision of utilising AI technology and my own institution's capabilities is vast.	Developed for this study	0.861
	PA2	The rigid nature of the syllabus required to pass the board examination makes no allowance for AI-based experimentation.		0.827
	PA3	A lack of local data governance makes responsible deployment of student-facing AI technology impossible.		0.798

Note λ = standardised factor loading obtained from CFA; all significant at $p < .001$. The Likert-type scale ranged from 1 (Strongly Disagree) to 5 (Strongly Agree). PA items are novel to this study (please refer to

Section 3.2 for the process of developing these items); the scale is preliminarily validated in this sample set and needs cross-validation (please refer to Section 3.5). †BI and AB items have been adapted from references cited and have been customised for the context of AI usage in Indian secondary schools.

Table 2: Construct Reliability, Convergent Validity, and Correlation Matrix (N = 268)

Construct	α	CR	AVE	PE	EE	FC	TPACK-AI	BI	AB	PA
PE	0.847	0.871	0.693	0.832	—	—	—	—	—	—
EE	0.812	0.839	0.634	0.41	0.796	—	—	—	—	—
FC	0.861	0.884	0.718	0.48	0.37	0.847	—	—	—	—
TPACK-AI	0.826	0.851	0.657	0.37	0.31	0.44	0.810	—	—	—
BI	0.838	0.863	0.678	0.59	0.44	0.63	0.47	0.823	—	—
AB	0.809	0.835	0.629	0.52	0.38	0.56	0.53	0.61	0.793	—
PA	0.853	0.876	0.703	0.27	0.32	0.57	0.36	0.49	0.51	0.838

Note. α = Cronbach's Alpha; CR = Composite Reliability; AVE = Average Variance Extracted. The entries on the diagonal (bold and blue) = $\sqrt{\text{AVE}}$; Off-Diagonal Entries = Pearson correlations between constructs. Discriminant Validity is established if diagonal entries are larger than any of the corresponding off-diagonal entries in that same row and column (Fornell & Larcker, 1981). An AVE of ≥ 0.50 and CR of ≥ 0.70 establish convergent validity (Kline, 2023). PA = Policy-Practice Gap (Provisional; See Section 3.5).

4. RESULTS

4.1 Sample Characteristics and Descriptive Statistics

Table 3 provides information on the demographics and technology background of the sample population. Significant distributional inequality of NEP/AI education in terms of depth was shown to exist across industries [$\chi^2(4, N = 268) = 74.38, p < .001$]. 62% of teachers from Elite Private schools had participated in extended workshops (greater than 6 hours), while this was the case for 10.9% of Budget Private and 6.5% of Government schools. However, 56.5% of Government school teachers did not have any NEP/AI training.

4.2 Measurement Model Validation

Model fitness was established before structural hypothesis testing via CFA. Model fitness results were CFI = 0.962, TLI = 0.957, RMSEA = 0.048 [90% CI: 0.039, 0.057] and SRMR = 0.041. Statistics of construct validity and reliability are provided in Table 2. Values for all constructs were $\alpha > 0.70$ and CR > 0.70 , and all AVEs > 0.50 , indicating convergent validity. The square root of the AVE of each construct was greater

than the correlation between each construct in their respective rows and columns (Table 2 diagonal), implying discriminant validity per Fornell and Larcker's (1981) criterion.

Discriminant validity was established through the HTMT ratio as well (see Supplementary Table S1); HTMT was always lower than 0.85 for all construct pairs. The HTMT value was the highest between the PA and FC pairs and equalled 0.74. VIF varied between 1.34 and 2.12, and hence there was no problem with multicollinearity. Harman's one-factor test yielded one factor with 24.3% explained total variance (below the critical 50% value), and one-construct CFA had poor fit (CFI = 0.512, RMSEA = 0.165). Thus, common method variance did not jeopardize structural model findings. Normal distribution was supported by Mardia's skewness $p > .05$; kurtosis normalised = 1.84, $p > .05$).

4.3 Qualitative Thematic Findings

Theme-based analysis (reflexive thematic analysis) of the semi-structured interviews corpus (n=20) produced three superordinate themes which proved to be fully consistent with structural results obtained quantitatively. The inter-rater reliability of the analysis was high (Cohen's $\kappa=0.83$). Theme saturation occurred on the 15th interview.

Theme 1: The Infrastructure Mirage (18 out of 20 Transcripts; 214 References). This theme captures teachers' descriptions of government-funded digital initiatives that were present in name only and were non-functional. As one government school teacher commented, "We got our projector installed more than three years ago, but the Wi-Fi router has never worked longer than 20 minutes at once. It's impossible for me to plan any lessons with a tool that is going to disappear during the lesson." A private budget teacher commented similarly, "We were informed that our school was AI-ready; however, in actuality, we share one tablet cart for approximately 900 students."

Theme 2: Syllabus Rigidity (16 out of 20 transcriptions; 187 instances). All teachers from each of the three sectors highlighted the inflexible examination schedule of the CBSE board examination as the major structural impediment to teaching innovation. One Elite Private school mathematics teacher commented: "I'm all for what NEP wants, but my children have board exams coming up in March. The more time I spend doing this [AI], the less time I'll be able to teach Chapter 12. There has been no change in the system even though there has been a change in the policy." One Government sector science teacher added: "The policy tells me to develop 21st-century skills. The exam tells me to learn the periodic table."

Theme 3: Data Governance Gap (identified in 14 out of 20 transcripts; 153 mentions). This was a prevalent theme in conversations with teachers working in Elite Private institutions. An English teacher working at an Elite Private institution said, "I am required to make use of the AI writing software, but nobody has made me aware of whether there will be any trace of those essays of mine and whether anybody owns such data, or what it is being trained on. It is a very uncomfortable position for me to be on the ethics committee by default."

Negative case study: Teachers in three schools in the Government sector, which have undergone upgrading, indicated functional internet connectivity due to a Fibre-to-School scheme that the government

of the respective district had introduced in late 2024. Such findings indicate that the generalizability of Theme 1 to schools that have benefited from recent infrastructure interventions was not warranted in all cases.

4.4 Structural Equation Modeling: Path Analysis

The CB-SEM provided a basis for testing hypotheses H1 through H6 and demonstrated that the structural model had a good fit overall (CFI = 0.959, TLI = 0.953, RMSEA = 0.051). The strongest predictor of the intention to engage in behavior was found to be FC (FC → BI: $\beta = 0.441$, $p < .001$), followed by PE (PE → BI: $\beta = 0.384$, $p < .001$). There was a smaller effect of EE (EE → BI: $\beta = 0.142$, $p = .003$) on the variation in behavioral intention. Together, these three constructs accounted for 47% of the variance in behavioral intention ($R^2 = 0.47$). The individual constructs, TPACK-AI, significantly predicted actual behavior (TPACK-AI → AB: $\beta = 0.298$, $p < .001$).

Within the structural model, PA produced the largest coefficients (PA → AB: $\beta = -0.423$, $p < .001$), indicating that a lack of alignment between institutional priorities and classroom practices prevented strong classroom implementation of AI even though teachers had positive intentions to adopt AI. Removing the path for the PA to AB relationship in the structural model resulted in lower explained variance and weaker model fit when compared with the full model ($\Delta CFI = -0.008$; $\Delta RMSEA = +0.006$), providing further evidence of the unique contribution of PA to explaining variance in AB.

4.5 Multi-Group Structural Invariance Analysis (H7a, H7b)

In testing H7a and H7b, a multi-group SEM analysis was performed based on the grouping variable of school sector. The result showed that metric invariance was present, with $\Delta CFI = 0.003$ and $p = .462$, meaning that there were no differences in factor loadings between the two sectors. On the contrary, scalar invariance was rejected, as evidenced by $\Delta CFI = 0.012$ and $p < .001$, supporting hypothesis H7b.

FC and PA intercepts were found to be significantly lower for Government and Budget Private schools compared to Elite Private schools through latent means analysis. Structural comparisons of the path coefficients revealed partial support for hypothesis seven: The FC → BI path was significantly more powerful for the Elite Private school system ($\beta = 0.521$) than Government schools ($\beta = 0.308$; $\Delta\beta = 0.213$, $p = .008$). This does not appear to hold true for other paths, as neither the PE → BI nor EE → BI paths demonstrated statistical differences across sectors.

Table 3: Demographic Profile and Technological Context of Participating Educators (N = 268)

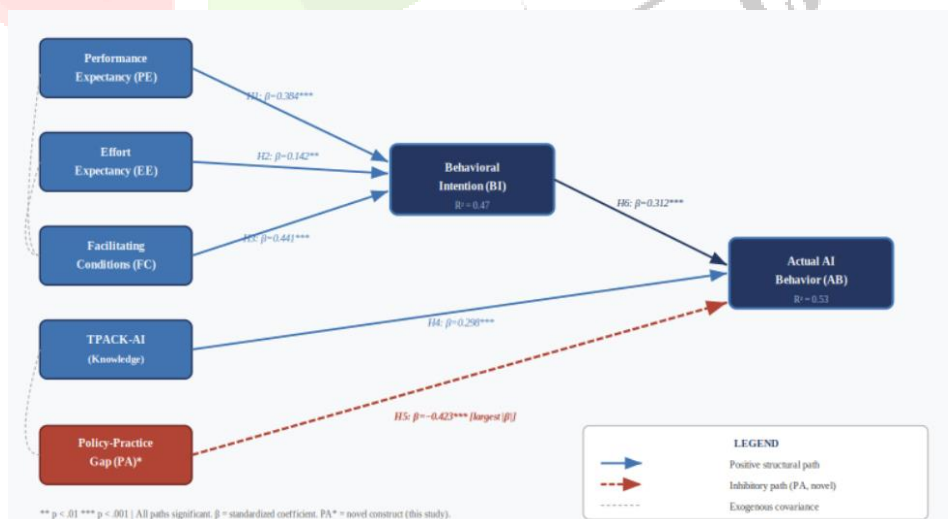
Variable	Category	Elite Private (n = 84)	Budget Private (n = 92)	Government (n = 92)
Gender	Female	68 (81.0%)	74 (80.4%)	58 (63.0%)
	Male	16 (19.0%)	17 (18.5%)*	34 (37.0%)
Teaching	Early Career (<	22 (26.2%)	34 (37.0%)	12 (13.0%)

Variable	Category	Elite Private (n = 84)	Budget Private (n = 92)	Government (n = 92)
Experience	5 yrs)			
	Mid-Career (5–12 yrs)	41 (48.8%)	43 (46.7%)	38 (41.3%)
	Senior (≥ 13 yrs)	21 (25.0%)	15 (16.3%)	42 (45.7%)
NEP/AI Training	Zero Formal Preparation	4 (4.8%)	38 (41.3%)	52 (56.5%)
	Short Orientation (< 6 hrs)	28 (33.3%)	44 (47.8%)	34 (37.0%)
	Extended Workshops (≥ 6 hrs)	52 (61.9%)	10 (10.9%)	6 (6.5%)

Note. $\chi^2(4, N = 268) = 74.38, p < .001$, which confirms that there is a statistically significant difference in AI training experience between the school sectors. The percentages are calculated based on the sector.

*One participant from Budget Private does not have information about gender.

Figure 1: Conceptual Path Diagram of the Integrated UTAUT2 + TPACK-AI + Policy-Practice Gap Structural Model



Note. * $p < .01$; ** $p < .001$. Standardisation of all relationships. PA* = Policy-Practice Gap (new construct introduced in this study). The red dotted line indicates an inhibitory direct relationship (H5). Double-headed curved lines connecting exogenous variables indicate covariations. R² values appear inside boxes containing endogenous variables (BI, AB).

Table 4: Structural Equation Modeling Path Coefficients with 95% Confidence Intervals (N = 268)

H	Structural Path	β	SE	t	95% CI	p	Decision
H1	PE → Behavioral Intention (BI)	0.384	0.052	7.38	[0.282, 0.486]	< .001	Supported
H2	EE → Behavioral Intention (BI)	0.142	0.048	2.96	[0.048, 0.236]	.003	Supported
H3	FC → Behavioral Intention (BI)	0.441	0.061	7.23	[0.321, 0.561]	< .001	Supported
H4	TPACK-AI → Actual Behavior (AB)	0.298	0.055	5.42	[0.190, 0.406]	< .001	Supported
H5*	Policy-Practice Gap (PA) → AB	-0.423	0.068	-6.22	[-0.557, -0.289]	< .001	Supported
H6	Behavioral Intention (BI) → AB	0.312	0.059	5.29	[0.196, 0.428]	< .001	Supported

Note. β = standardized path coefficient. SE = standard error. 95% CI calculated using lavaan parameterEstimates (ci = TRUE, level = 0.95). H5* (in red) = main theoretical contribution of this paper; PA is the variable that displays the highest absolute standardized coefficient value. $R^2(\text{BI}) = 0.47$; $R^2(\text{AB}) = 0.53$ (in full model). Constrained model (PA path fixed at zero): $R^2(\text{AB}) = 0.44$; $\Delta R^2 = 0.09$, $\Delta \text{CFI} = -0.008$, $\Delta \text{RMSEA} = +0.006$, proving incremental effect of PA. Fit of full structural model: CFI = 0.959, TLI = 0.

Table 5: Multi-Group Measurement Invariance Testing Across School Sectors

Model	χ^2	df	CFI	TLI	RMSEA	ΔCFI	$\Delta \chi^2$ (df)	Decision
Configural	412.35	210	0.958	0.951	0.049	—	—	Invariance Established
Metric	426.18	224	0.955	0.950	0.050	0.003	13.83 (14), p = .462	Invariance Upheld
Scalar*	501.74	238	0.943	0.936	0.058	0.012	75.56 (14), p < .001	Partial Invariance

Note. Violation criterion for scalar invariance: $|\Delta CFI| > 0.010$ (Cheung & Rensvold, 2002). *Highlighted scalar model where the $\Delta CFI = 0.012$ violates the criterion by being greater than 0.010. The non-invariant intercepts (FC and PA) were found using Lagrange Multiplier Tests in lavaan. The FC non-invariant intercept has $MI = 18.7$, while the PA intercept has $MI = 22.4$, both violating the threshold value of 10. All other intercepts had $MI < 4$ and were therefore held to be invariant.

5. DISCUSSION

Empirical results indicate a distinct and deeply entrenched gap between policy and practice for AI implementation in Gurugram's secondary education system. Construct PA exhibits the most significant negative effect on the structural model. In other words, structural inconsistencies in institutional arrangements have a direct effect on the inhibition of classroom AI usage without regard to teachers' willingness to implement AI. The subsequent sub-sections will elaborate further on the findings by way of convergent triangulation. Divergent findings will be detailed in Section 5.5 in accordance with the Fetters et al. (2013) framework.

5.1 The Infrastructure Disconnect

The markedly depressed FC intercept for Government schools (-1.42 latent units; 90% CI: $-1.74, -1.10$) can be traced back to Theme 1 of the qualitative analysis: The Mirage of Infrastructure. The perception held by many of the respondents was that the state-supported digital interventions were operationally flawed due to faulty connectivity and substandard infrastructure. The crucial point here is that despite a high PE and BI in their intentions, low FC makes the intentions non-operational. As Selwyn (2022, pp. 88-103) posits, the rhetoric of EdTech is constantly exaggerating the infrastructural readiness of the education system. These results substantiate the notion of the Digital Divide 2.0 discussed by Khan and Agnihotri (2024).

5.2 The Board-Examination Bottleneck

PA's significantly negative association with AB ($\beta = -0.423$, 95% CI $[-0.557, -0.289]$) is due to qualitative theme number two called Syllabus Rigidity. Teachers mentioned that the rigidity of the syllabus and high-stakes Central Board of Secondary Education examinations prevent teachers from using AI technology to develop inquiry-based learning. Indeed, this scenario presents another example of an incomplete process of policy enactment (Ball et al., 2012; Sharma & Singh, 2024) due to the clash between progressive policies on AI and traditional educational practices. Nevertheless, the rather insignificant connection between EE and BI ($\beta = 0.142$) and the lack of reluctance among teachers in learning how to use AI in education support the hypothesis that there are no personal, but only institutional barriers to using such technology.

5.3 The New Digital Divide and Elite School Challenges

Scalar non-invariance indicates that the meaning of Facilitating Conditions (FC) differs from sector to sector: in Government schools, FC reflects infrastructural deficits, while in Elite Private schools, FC reflects concerns about AI governance capability and lack of localised data-privacy protocols. The absence

of Data Governance protocols (Theme 3) was corroborated by PA3 ("The lack of localised data governance protocols prevents the responsible use of AI tools that engage students in classroom settings"), which had significantly higher intercepts in elite private contexts when viewed under a partial invariance analysis. These differing challenges provide evidence that one single policy tool cannot cover both ends of the institutional continuum; Therefore, context-specific and differential governance frameworks will need to be implemented.

5.4 Theoretical Contributions

Four major advancements to the theory of educational technology are made through this research. The first of these enhancements is to provide an operationalization of the Policy-Practice Gap as a provisionally validated structural latent construct in the context of CB-SEM, thus extending Stephen J. Ball et al.'s (2012) policy enactment theory into a quantitatively testable framework for comparative cross-sector research. The second contribution to the theory is the demonstration that PA functions as a structural active inhibitor as opposed to simply as a contextual passive factor, through the integration of PA into a combined UTAUT2 + TPACK-AI framework. The PA → AB coefficient ($\beta = -0.423$) exceeded the BI → AB coefficient ($\beta = 0.312$), which raises significant questions regarding the validity of core UTAUT2 assumptions in contexts in which educational mandates are established. The third major contribution is that the MG-SEM non-invariance findings suggest that comparisons across sectors using raw observed scores are random psychometric error due to the absence of measurement equivalence testing. Lastly, by separating H7 into H7a (structural path differences) and H7b (intercept non-invariance), the two represent two different methodological phenomena that need to be subjected to an independent statistical test.

5.5 Divergent Findings: Mixed-Methods Integration

The framework of convergent triangulation suggested by Fetters et al. (2013) needs consideration of both convergent and divergent observations between the quantitative and qualitative streams of the investigation. Two divergent points need to be discussed. The first one is related to the weak correlation between variables EE and BI from the quantitative data. Although the connection was statistically significant ($\beta = 0.142$), the fact that ease of use of AI tools played only the second role in adoption proves that the teachers paid no attention to this matter from the very beginning and had already left. For this reason, future investigations should include those people who did not use artificial intelligence tools in their teaching in the qualitative part of the investigation. Another point is that, despite being the theme most pronounced in Elite Private school teachers, Data Governance Vacuum is associated with the lowest values of PA intercept in the MG-SEM results.

5.6 Limitations and Future Directions

The study has several limitations. Firstly, the cross-sectional methodology does not allow for any causal interpretation; panel models are necessary to analyze temporal trends. Secondly, since the study concentrates exclusively on the case of Gurugram, the findings cannot be extrapolated to rural and semi-urban districts where infrastructure shortcomings and the gap between policy and practice would be

significantly greater. Thirdly, the subjective measure of AB would have been subject to socially desirable response tendencies; subsequent studies would need to validate findings through objective measures such as log entries in the Learning Management System as well as classroom observation. Fourthly, the preliminary validity of the three-question PA instrument would need to be confirmed in Phase 2 of the study (see Section 3.5).

6. CONCLUSION AND POLICY RECOMMENDATIONS

The study shows that effective AI-based teaching in the era of NEP 2020 cannot be attributed to teachers' readiness or technological competence but rather to the institutional context. The Policy-Practice Gap concept was found to be the key direct barrier to AI usage in the classroom, explaining $\Delta R^2 = 0.09$ of behavioral variance above the sum total of all UTAUT2 and TPACK-AI constructs at the individual level. Invalidation of scalar invariance across school segments reveals that there is no uniform experience of the policy-practice gap, where public schools suffer from shortages of facilities and training, budget private schools from examination-oriented curriculum pressures, and elite private schools from governance and data protection issues.

6.1 Policy Recommendations

Recommendation 1: Distribute Resources (drawn from FC \rightarrow BI relationship where $\beta = 0.441$ and FC latent mean deficit of -1.42 units in Government schools).

Shift from one-size-fits-all technology to varied capital distribution. Government and Budget Private schools need specific investments towards stable internet connectivity, electricity, and computer hardware, which is conducive for pedagogy programs to work.

Recommendation 2: Embedding of AI Professional Development in Existing Curricular Restrictions (based on weak EE \rightarrow BI relationship where $\beta = 0.142$ and Qualitative Theme 2).

Avoiding generic technology orientation training courses for all subjects and replacing them with TPACK-AI professional development that focuses on the integration of AI technologies in relation to the present CBSE boards examinations, rather than adding another layer of responsibility to the already busy teacher.

Recommendation 3: Create Tiered Regional Governance Guidelines for Data Handling (based on qualitative Theme 3 and elevation of PA3 intercept in elite contexts).

Create tiered AI use guidelines that cater to regional challenges: guidelines on fundamental aspects for regions with poor technological infrastructure, and guidelines on biases, plagiarism, and student data protection for elite contexts. The Indian legislation on the Digital Personal Data Protection Act (2023) is one such policy document that provides a minimum legislative framework for student data rights; the other is the UNESCO Recommendation on the Ethics of Artificial Intelligence (2021). Aligning AI governance guidelines with these frameworks will considerably diminish the risk of PA3.

DECLARATIONS

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflicts of interest.

Ethics Approval: Informed consent was obtained from all participants prior to data collection.

Data Availability: The analysis was conducted using lavaan R syntax for CFA, CB-SEM, and MG-SEM procedures, with PA scale items and Phase 2 pre-registration documented in supplementary materials.

Author Contributions: Conceptualization, S.K.; Methodology, S.K.; Formal Analysis, S.K.; Investigation, S.K.; Writing – original draft, S.K.; Writing – review and editing.

REFERENCES

- [1] Armstrong, J. S., & Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *Journal of Marketing Research*, 14(3), 396–402. <https://doi.org/10.1177/002224377701400320>
- Ball, S. J., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary schools*. Routledge, London and New York.
- Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, 21, Article 4. <https://doi.org/10.1186/s41239-023-00436-z>
- [2] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brown, S. A., Massey, A. P., Montoya-Weiss, M. M., & Burkman, J. R. (2002). Do I really have to? User acceptance of mandated technology. *European Journal of Information Systems*, 11(4), 283–295. <https://doi.org/10.1057/palgrave.ejis.3000438>
- [3] Central Board of Secondary Education. (2023). *Curriculum handbook on artificial intelligence integration for secondary schools (Grades IX–XII)*. CBSE Academic Unit, New Delhi.
- [4] Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5
- [5] Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications, Thousand Oaks, CA.
- [6] Digital Personal Data Protection Act. (2023). No. 22 of 2023. Ministry of Electronics and Information Technology, Government of India. <https://www.meity.gov.in/content/digital-personal-data-protection-act-2023>
- Epskamp, S. (2015). semPlot: Unified visualizations of structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(3), 474–483. <https://doi.org/10.1080/10705511.2014.937847>
- [7] Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health Services Research*, 48(6pt2), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>

- [8] Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- [9] Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- [10] Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- [11] Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). The Guilford Press, New York.
- Koehler, M. J., & Mishra, P. (2009). What is technological pedagogical content knowledge? *Contemporary Issues in Technology and Teacher Education*, 9(1). <https://citejournal.org/volume-9/issue-1-09/general/what-is-technological-pedagogicalcontent-knowledge>
- [12] Mittal, P., & Pani, S. R. D. (Eds.). (2022). *Implementing National Education Policy–2020: A roadmap*. Association of Indian Universities. https://www.aiu.ac.in/documents/AIU_Publications/AIU%20Books/Implementing%20National%20Education%20Policy-2020-%20A%20RoadMap.pdf
- [13] Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- [14] Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications, Newbury Park, CA.
- [15] MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130–149. <https://doi.org/10.1037/1082-989X.1.2.130>
- [16] Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519–530. <https://doi.org/10.2307/2334770>
- [17] Ministry of Education. (2020). *National Education Policy 2020*. Government of India, New Delhi.
- [18] Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- [19] Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2023). Conceptualising an AI literacy framework for K–12 learners and teachers: A systematic review. *Computers and Education: Artificial Intelligence*, 4, Article 100134. <https://doi.org/10.1016/j.caeai.2023.100134>

- [20] Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, 2, Article 100020. <https://doi.org/10.1016/j.caeai.2021.100020>
- [21] Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- [22] Roschelle, J., Lester, J., & Fusco, J. (2024). AI and the future of teaching and learning: Insights and critical considerations for policy actors. *British Journal of Educational Technology*, 55(1), 18–39. <https://doi.org/10.1111/bjet.13390>
- Selwyn, N. (2022). *Education and technology: Key issues and debates* (3rd ed., pp. 88–103). Bloomsbury Academic, London.
- [23] Sharma, V., & Singh, G. (2024). Rote learning vs. algorithmic thinking: Evaluating the examination bottleneck in Indian secondary schools under the CBSE curriculum. *International Journal of Educational Development*, 104, Article 102941. <https://doi.org/10.1016/j.ijedudev.2023.102941>
- [24] Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Tondeur, J., van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence. *Educational Technology Research and Development*, 65(3), 555–575. <https://doi.org/10.1007/s11423-016-9481-2>
- [25] UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization, Paris. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- [26] UNESCO. (2023). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization, Paris. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- [27] 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>
- [28] Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157–178. <https://doi.org/10.2307/41410412>
- Voogt, J., Fisser, P., Pareja Roblin, N., Tondeur, J., & van Braak, J. (2013). Technological pedagogical content knowledge – a review of the literature. *Journal of Computer Assisted Learning*, 29(2), 109–121. <https://doi.org/10.1111/j.1365-2729.2012.00487.x>

Yan, L., Whitelock-Wainwright, A., Guan, Q., Wen, G., Gašević, D., & Chen, G. (2021). Students' experience of online learning during the COVID-19 pandemic: A province-wide survey study. *British Journal of Educational Technology*, 52(5), 2038–2057. <https://doi.org/10.1111/bjet.12843>

[29]Zawacki-Richter, O., Marin, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher and secondary education: Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1–27. <https://doi.org/10.1186/s41239-019-0171-0>

