Available Online: 4 November 2026 DOI: 10.54254/2753-7102/2026.38601

Meritocracy and generative AI in higher education: institutional pathways and use–response alignment across four systems

Jingxuan Lin, Dongmei Cun*

Shenzhen Polytechnic University, Shenzhen, China

*Corresponding Author. Email: cundongmei@szpu.edu.cn

Abstract. This study examines how higher education systems respond to the rapid adoption of Generative Artificial Intelligence (GenAI) under different institutional conditions, asking how meritocratic competition is translated into assessment and credential pressure and how student GenAI use aligns with institutional responses. Using a comparative qualitative case-study design, the study analyzes secondary evidence from China, the United Kingdom, the United States, and Germany through two a priori frozen indicator frameworks. The first framework traces how meritocratic competition is institutionalized through admissions, grading, and credential differentiation; the second maps student GenAI use patterns against institutional responses in teaching, assessment, and support. Strong credential differentiation and education–labour-market linkage emerge in all four systems, but through distinct pathways: state-coordinated in China, marketized in the United Kingdom, decentralized and market-oriented in the United States, and state–professional in Germany. Partial use–response alignment is the cross-system norm; systems differ mainly in pedagogical integration, regulatory adaptation, and institutional support. AI literacy and support provision lag behind student adoption in all four cases, making institutional capacity the most persistent source of misalignment, while limited comparable measurement of equity alignment and of guarded and compensatory use shows that monitoring capacity is itself part of governance. GenAI governance therefore converges less on common rules than on common capability gaps. The findings suggest that governing GenAI in meritocratic systems depends less on uniform rules than on sustained investment in institutional monitoring and support capacity.

Keywords: comparative case study, generative AI, higher education systems, meritocracy, use–response alignment

1. Introduction

Meritocratic competition in higher education extends beyond admission. Once students enter university, competition is reproduced through grading, ranking, degree classification, institutional prestige, credential hierarchies, and the labour-market value attached to qualifications. These mechanisms make academic distinctions consequential because they structure later educational and occupational opportunities. Admissions criteria vary across institutional settings [1], and university grades can function as early labour-market signals

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

when employers have limited information about productivity [2]. The rapid diffusion of Generative Artificial Intelligence (GenAI) adds a further governance problem to this established field of competition: whether institutional responses correspond to how students actually use AI, the pressures they face, and the support they require.

Research on these processes has developed along two largely separate lines. Studies of meritocracy and credentialism examine admissions, grading, stratification, assessment, or labour-market signalling, whereas GenAI research more often begins with adoption, academic integrity, pedagogy, or institutional policy. Work in the first line shows that assessment practices shape students' understandings of legitimate achievement [3]; work in the second shows that student GenAI practices are heterogeneous rather than a simple use/non-use distinction [4]. What remains under-specified is the connecting mechanism: how system-level arrangements in admissions, assessment, stratification, and credential structures become concrete post-entry pressures; how those pressures condition students' GenAI-use strategies; and whether institutional responses align with, lag behind, or misrecognize those practices. Because the available evidence is concentrated in single-country, single-institution, and especially Anglophone settings, general mechanisms are also difficult to distinguish from system-specific effects.

This study addresses that gap through a framework combining credentialism, Institutional Logics Theory, and sociotechnical affordances. Credentialism explains why grades, classifications, institutional status, and qualifications acquire competitive value beyond their immediate educational content. Institutional Logics Theory explains why similar competitive pressures are translated differently through state, market, academic-professional, managerial, and equity-oriented arrangements; comparative research demonstrates its value for connecting societal structures to organizational practices [5]. Sociotechnical affordances explain why the same GenAI technology supports different forms of student action depending on task demands, capability, access, and institutional conditions, without requiring inferences about unobserved psychological motives. Differences in AI literacy and access already condition whether students can convert GenAI into educational advantage [6]. Figure 1 summarizes the resulting mechanism chain: institutional arrangements generate credential pressure, credential pressure shapes differentiated GenAI-use patterns, and GenAI-use patterns and institutional responses condition one another.

Two research questions organize the analysis. RQ1 asks how meritocratic competition and credential differentiation are institutionally translated across four higher education systems, focusing on the arrangements through which admissions, assessment, stratification, grading, credential differentiation, and labour-market linkage create post-entry competitive stakes. RQ2 asks how student GenAI-use patterns and institutional responses form use–response configurations, and where alignment or misalignment emerges; student practices and institutional responses are treated as analytically distinct before their relationship is examined. As operationalized a priori, use–response outcomes are coded as aligned, partially aligned, misaligned, or insufficiently measured rather than assumed in advance.

The study employs a comparative qualitative case design across China, the United Kingdom, the United States, and Germany. Cases were selected for maximum variation in institutional structures, assessment regimes, credential differentiation, and education–labour-market linkages rather than as predetermined levels of meritocracy. Two a priori frozen indicator frameworks are applied identically across cases: RQ1 is operationalized through institutional arrangements and credential pathways, and RQ2 through student-use profiles, institutional-response profiles, and use–response alignment. GenAI evidence covers 2023–2026 and is supplemented by the policy, institutional, statistical, and credential documents needed to characterize each system. South Korea and Finland are retained as shadow cases: they provide theoretically useful contrasts but

have weaker or less comparable student GenAI evidence. This design distinguishes recurring mechanisms from system-specific pathways instead of treating national cases as self-contained descriptions.

The study makes three contributions. First, it connects system-level arrangements to student-level practice through institutional mechanisms without assuming direct individual-level causation. Second, it shows how structurally different systems can generate similar credential consequences and converge on comparable governance problems—a pattern consistent with policy research finding that universities combine restriction, conditional permission, pedagogical integration, and enabling conditions rather than following a single regulatory model [7, 8]. Third, it identifies institutional capacity, particularly AI literacy and support provision, as a source of misalignment distinct from the formal question of whether GenAI is permitted. Policy implications are developed with particular reference to China and the ASEAN region, where existing qualification-transparency infrastructure, notably the ASEAN Qualifications Reference Framework (Association of Southeast Asian Nations [ASEAN]) [9], provides a basis for regional coordination. The article proceeds as follows. Section 2 reviews the evidence on credential competition, assessment, student GenAI use, and institutional governance. Subsequent sections examine how institutional arrangements translate into credential pressure (RQ1) and how this pressure conditions GenAI-use patterns and institutional responses (RQ2), before identifying which configurations are common across systems and which remain path-dependent.

2. Literature review

Research on meritocracy, admissions, grading, and credentialism increasingly treats competition in higher education as a multi-stage institutional process rather than an event confined to entry. Admissions criteria, grading systems, credential hierarchies, and labour-market signalling jointly define legitimate achievement and allocate educational and occupational opportunities. Formally neutral admission mechanisms can have unequal effects because the meaning of "merit" depends on institutional context [1, 10], and competitive environments can generate markets for supplementary advantage whose legitimacy students actively negotiate. Differentiation continues after admission: grading standards vary across institutions and faculties, grade inflation alters the signalling value of marks, and university GPA can influence early-career earnings because employers use grades as indicators of otherwise unobservable productivity [2, 11]. Interpretations nevertheless differ. Some studies emphasize inequality reproduction through ostensibly meritocratic procedures, while others show that standardized selection, grading, and credential signals retain allocative functions. System context also matters: UK research emphasizes marketization [12], whereas work on Chinese elite universities highlights institutional cultures in which academic performance interacts with political, organizational, and socioeconomic criteria [13]. Read as a whole, this literature treats meritocratic competition as an institutional pathway linking admission selectivity, grading and ranking, credential differentiation, and education–labour-market linkages rather than as a fixed national trait.

A second stream identifies assessment as the mechanism through which institutional competition becomes consequential for student behaviour. Assessment does more than measure learning: it defines performance stakes and shapes understandings of legitimate effort, deservingness, efficiency, and academic risk. Students may interpret academic standards through individualized notions of responsibility and competence that obscure unequal starting conditions [3]. GenAI-related assessment research similarly indicates that different assessment environments produce different student strategies: mastery-oriented engagement may support knowledge construction, while procedural use may enable efficient task completion without comparable learning gains. The behavioural effects of competitive assessment remain contested, however. Opportunity

structures, perceived usefulness, and moral disengagement may encourage strategic or prohibited AI use, including among high-achieving students [14], whereas strong achievement identities can foster guarded adoption intended to preserve independence, epistemic agency, and authentic achievement [15]. GenAI may also function as a compensatory resource when students perceive gaps between their capabilities and assessment demands [16]. These findings complicate any simple "pressure leads to cheating" account and support treating student strategies as instrumental, guarded, compensatory, or mastery-oriented responses rather than reducing them to misconduct versus legitimate use. Most of these studies remain institution-, course-, or sample-specific, however, and rarely connect observed assessment behaviour to the wider organization of credential competition.

A third stream shows that student GenAI use is heterogeneous and institutional governance multidimensional. Prevalence alone cannot capture student practice: research distinguishes learning support from cognitive substitution, surface from deep engagement, and different levels of AI literacy and agency [4, 6, 17]. Unequal access, knowledge, and confidence may further shape students' capacity to convert GenAI into educational advantage. On the institutional side, responses range from academic-integrity restrictions and disclosure requirements to pedagogical integration, assessment redesign, AI-literacy provision, and equitable-access support [7, 8]. The interpretive disagreement is whether these responses primarily protect academic standards or indicate regulatory lag when restrictive rules fail to reflect normalized student practices. Methodologically, student studies frequently rely on single institutions, non-equivalent national surveys, or Anglophone samples, while governance studies often analyse policy without examining actual student behaviour. Even cross-national policy studies rarely pair institutional rules with comparable student-use evidence from the same systems. This study therefore treats student-use patterns and institutional responses as analytically distinct before examining their relationship as a use–response configuration, so that alignment, regulatory lag, over-restriction, pedagogical adaptation, and equity gaps can emerge from the evidence rather than being assumed in advance.

These three streams converge on a specific mechanism gap. Existing research establishes that higher education systems institutionalize merit through admission, assessment, credentials, and labour-market signalling; that assessment environments shape students' academic strategies; and that GenAI use and university governance vary considerably. What remains insufficiently explained is how these levels connect within the same institutional pathway: how system-level arrangements become concrete assessment and credential pressures, how those pressures shape the strategic and legitimate uses students attach to GenAI, and whether universities subsequently respond in ways that align with or misrecognize those practices. The unresolved issue is therefore not simply whether meritocracy, assessment pressure, and GenAI use are related, but which institutional arrangements generate which student-use patterns and which use–response configurations. This limitation is reinforced by the concentration of existing evidence in single-country, institution-specific, or Anglophone contexts, which makes general mechanisms difficult to distinguish from effects produced by particular admissions regimes, grading traditions, credential hierarchies, or labour-market structures. RQ1 and RQ2, as stated in Section 1, are designed to close this gap within the same four systems: RQ1 traces the institutional translation of meritocratic competition and credential differentiation, and RQ2 examines the use–response configurations through which student practices and institutional responses meet.

3. Theoretical framework

Credentialism anchors the framework: it explains why admission outcomes, grades, degree classifications, institutional prestige, and qualifications acquire competitive value beyond their immediate educational

content. Higher education allocates scarce positional goods through formally certified achievements, and competition therefore persists after entry. Test-based admission can generate differential outcomes under standardized merit-based rules [10], while institutional variation in admissions confirms that indicators of "merit" are context dependent [1]. Differentiation continues after entry: grade inflation can weaken the scarcity of marks [11], while grades retain labour-market value because employers use them as signals of otherwise imperfectly observable productivity [2]. Critiques of meritocracy clarify the legitimating logic of these arrangements: competitive allocation may reproduce privilege while being represented as deserved achievement [18], and meritocratic discourse can naturalize institutionally produced inequalities as individual merit [3].

Institutional Logics Theory addresses cross-system variation: similar competitive pressures and technologies are translated differently across higher education systems. Universities operate through plural state, market, academic-professional, managerial, and equity/community logics rather than through a single organizing principle. Institutional-logics research shows how these competing value systems connect societal structures with organizational practices and become dominant or hybridized in different settings [5]. In the present study, the theory frames how systems organize admissions, assessment, credentialing, labour-market linkage, and GenAI governance. Cross-national evidence likewise shows that universities balance usefulness, risk, enabling conditions, and institutional capacity differently when governing GenAI [7].

At the student level, sociotechnical affordances explain why the same GenAI technology does not generate a uniform mode of use. Affordances are relational possibilities for action that emerge through interactions among technological capabilities, student resources, task demands, and institutional conditions. Students report both educational benefits and concerns about accuracy, integrity, privacy, and skill development [4, 6]. Problematic reliance reflects interactions among environmental stress, capability gaps, technological convenience, and institutional support rather than technology alone [14], and outcomes depend on pedagogical framing and reflective use [19]. The four patterns distinguished in this study—instrumental, guarded, compensatory, and mastery-oriented—are observable affordance actualizations, not inferred psychological types.

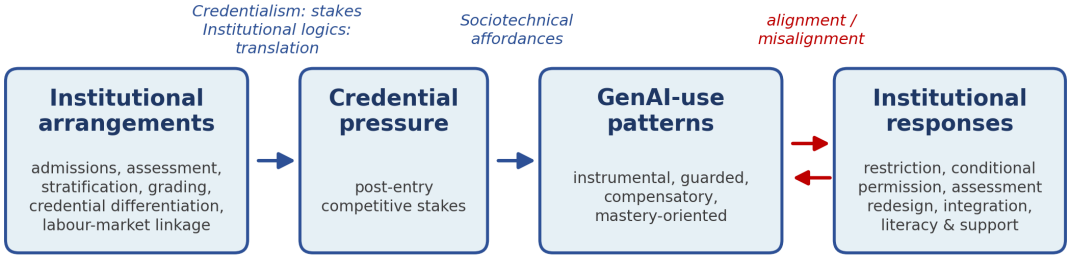


Figure 1. Mechanism chain linking institutional arrangements, credential pressure, differentiated GenAI-use patterns, and institutional responses; alignment or misalignment emerges from their interaction

Combined, the three theories specify a single mechanism: credentialism supplies the stakes (educational distinctions as consequential positional goods), institutional logics translate those stakes into different system-level arrangements, and affordances shape how students actualize GenAI within those arrangements. Figure 1 summarizes the mechanism, running from institutional arrangements through credential pressure to differentiated GenAI-use patterns and institutional responses; alignment or misalignment emerges from their

interaction. RQ1 focuses on the translation of institutional arrangements into assessment and credential pressure, while RQ2 examines how student practices and university responses form system-specific use–response configurations. The framework identifies system-level configurations and theoretically plausible institutional mechanisms; it does not infer individual-level causal effects or internal motivations from aggregate evidence.

4. Methodology

This study employed a comparative qualitative multiple-case design to examine how meritocratic competition is institutionally translated into assessment and credential pressure and how student generative-AI (GenAI) use is configured in relation to university responses across higher education systems. Comparative case studies are appropriate for identifying mechanisms, configurations, similarities, and differences across bounded contexts rather than seeking statistical representation; multiple-case designs are suited to examining whether similar outcomes arise through different pathways [20, 21]. Comparative higher education research also requires a common analytical framework that preserves contextual differences while maintaining comparability.

Cases were selected through maximum-variation purposive sampling rather than predetermined levels of meritocratic competition. This strategy deliberately selects heterogeneous cases to identify both contrasting pathways and common patterns. The four core cases—China, the United Kingdom, the United States, and Germany—maximize variation in admission structures, institutional stratification, assessment regimes, credential differentiation, and education–labour-market linkages. These characteristics were coded empirically rather than treated as country labels. South Korea and Finland were retained as shadow cases because they provide theoretically useful contrasts but had weaker or less comparable 2023–2026 student GenAI evidence. They informed transferability rather than receiving equal evidentiary weight—a conceptually driven rather than numerically representative approach to case selection [21].

All empirical GenAI evidence covered 2023–2026, supplemented by current statistics and policy materials relevant to RQ1. Sources comprised: (1) government and ministry documents; (2) university policies on GenAI, assessment, academic integrity, and support; (3) national, multi-institution, or sector-level student surveys; and (4) OECD/UNESCO comparative reports and statistics. Documentary sources were analysed systematically with attention to provenance, purpose, context, and relevance to the research questions.

Because GenAI surveys differed across systems, a comparability audit recorded, for each survey: source, year, sampling frame, N, UG/PG coverage, representativeness, GenAI definition, prevalence item, motive items, mode/use items, assessment items, raw-data availability, and link. Such metadata were retained because cross-national estimates may differ through sampling, population definitions, wording, response categories, fieldwork timing, and harmonization procedures [22], and credible secondary comparison requires explicit documentation of harmonization decisions [23]. Raw prevalence percentages were not treated as directly comparable unless sampling frame, population, construct, and wording were sufficiently equivalent. Each estimate retained its year, population, sample size, wording, and sampling frame. Surveys therefore served primarily as within-case pattern evidence, while cross-case analysis compared standardized qualitative codes.

Data collection and analysis followed two a priori frozen indicator frameworks. RQ1 was operationalized through nine indicators: meritocratic legitimation; admission selectivity; institutional stratification; assessment stakes; grading and ranking practices; credential differentiation; education–labour market linkage; equity and compensatory arrangements; and institutional logic configuration. RQ2 was operationalized through three blocks: Block A covered GenAI prevalence, motives, instrumental, guarded, compensatory, and mastery-oriented affordance actualization, and AI access/literacy conditions; Block B covered overall response

orientation, restrictive controls, conditional permission and disclosure, assessment redesign, pedagogical integration, AI literacy and support provision, and equity and access provision; Block C covered normative use–policy alignment, assessment alignment, pedagogical alignment, capability-support alignment, equity alignment, and overall use–response configuration. For RQ1, common thresholds were applied across cases. Coding distinguished system-wide from localized selection, institutional consequences from psychological pressure, and policy provision from effectiveness. Elite institutions were not treated as nationally representative; institutional logics were coded separately; mixed or insufficient evidence was coded conservatively rather than inferred.

The framework was applied identically across China, the UK, the US, and Germany. Indicator names, order, wording, and categories were fixed before comparison. Missing evidence was coded as Not found, Not measured, or Insufficient evidence; absence was never inferred from missing data; non-equivalent survey percentages were not directly compared; and institutional documents and student surveys remained analytically distinct. Cross-case interpretation relied on configurations rather than a composite "meritocracy score." Credentialism directed attention to admission, grading, credentials, and labour-market linkage; Institutional Logics structured the comparison of state, market, academic-professional, managerial, and equity/community logics; and Sociotechnical Affordances framed reported GenAI practices as observable actualizations rather than inferred psychological traits.

Analysis proceeded in two stages. First, within-case configurations were developed separately for China, the UK, the US, and Germany. For RQ1, the nine indicators were interpreted jointly to identify the dominant institutional pathway. For RQ2, Blocks A and B were analysed independently before Block C paired student-practice and institutional-response evidence to determine alignment, partial alignment, misalignment, or insufficient evidence. This sequencing preserved contextual coherence before comparison [20]. Second, cross-case analysis compared the same frozen indicators across all four systems. RQ1 examined whether similar credential stakes arose through different combinations of admission selectivity, institutional stratification, assessment stakes, grading/ranking, credential differentiation, labour-market linkage, equity arrangements, and institutional logics. RQ2 distinguished variant from invariant indicators, treating repeated invariant codes as findings and testing equifinality—whether different RQ1 pathways produced the same RQ2 configuration. Pattern matching, explanation building, and comparison through predefined dimensions support this transition from within-case interpretation to cross-case explanation [20]. No numerical composite index was calculated.

Trustworthiness was strengthened through source triangulation, a common frozen codebook, conservative coding, and retention of survey metadata and source links. Triangulation across government documents, university policies, sector reports, student surveys, and OECD/UNESCO evidence helps identify convergence, contradiction, and source-specific limitations [24]. An explicit audit trail connected source evidence to indicator codes, within-case configurations, and cross-case findings.

The inferential boundary remained system-level: the study identifies configurations and theoretically plausible institutional mechanisms but does not estimate individual-level causal effects of meritocratic or assessment pressure on GenAI use. Policy documents establish rules and priorities, not student behaviour; aggregate surveys establish reported patterns, not individual causal mechanisms. Not measured and Insufficient evidence were retained as analytical outcomes: the former indicated that a construct was absent from the survey instrument, the latter that available sources did not meet the frozen coding threshold. Their systematic distribution—including limited comparable measurement of guarded or compensatory GenAI use and insufficient evidence for equity alignment across all four core cases—was interpreted as evidence of limits in institutional monitoring and comparative data capacity, not as behavioural absence.

5. Findings

5.1. Institutional pathways to credential differentiation (RQ1)

Strong credential differentiation and education–labour market linkage emerge as universal features across China, the United Kingdom, the United States, and Germany, yet they operate through structurally distinct institutional pathways (Table 1). This equifinality challenges the assumption that admission selectivity alone determines competitive pressure and shifts analytical attention from who gets in to what happens after entry.

A second finding qualifies the first: admission selectivity does not predict the intensity of post-entry credential pressure. Germany combines low system-wide entry restriction with strong credential differentiation and labour-market linkage, while the United States pairs moderate admission selectivity with equally strong institutional stratification and credential differentiation. Post-entry distinctions carry independent market value: university grades affect early-career earnings because employers use them as productivity signals [2], and grading standards themselves are institutionally malleable in ways that condition progression and differentiation [11]. Competitive pressure is therefore reproduced primarily through post-entry assessment, qualification differentiation, institutional positioning, and labour-market signalling rather than through selection at the point of entry.

Four institutional configurations organize these outcomes. China follows a state-coordinated pathway, combining Moderate admission selectivity with Strong institutional stratification, Strong grading differentiation, and Strong credential linkage to labour markets. State and Managerial logics dominate, transmitting competitive pressure through institutional hierarchy, differentiated performance expectations, and organizational management; Chinese universities are known to mediate state priorities through institutional evaluation and stratified career structures rather than through direct policy transmission to individuals [25].

The United Kingdom follows a marketized pathway. Competitive pressure is organized through degree classification, assessment performance, institutional positioning, and regulatory oversight, with Market, Managerial, and Equity/community logics strongly present—a configuration consistent with the marketized structure of UK higher education [12].

The United States exemplifies decentralized market stratification. Despite Moderate admission selectivity, Strong institutional stratification arises from institutional autonomy combined with a highly differentiated organizational field: institutions apply divergent admissions criteria and distribute opportunities unevenly, reinforcing the positional value of institutional location and credentials [1]. Market, Academic-professional, and Equity/community logics drive this mechanism.

Germany operates through qualification-differentiation with comparatively open entry. Low admission selectivity coexists with Strong credential differentiation organized through Bachelor/Master qualification hierarchies, field differentiation, grading, and labour-market linkage. State and Academic-professional logics predominate, and even standardized merit-based procedures operate differently across institutional contexts [10].

The four pathways illustrate how different institutional-logic combinations—China's State–Managerial, the UK's Market–Managerial–Equity, the US's Market–Academic-professional–Equity, and Germany's State–Academic-professional—generate comparable credential stakes through institutionally distinct mechanisms [5, 26]. What unites the cases is not selectivity but strong credential differentiation coupled with strong education–labour market linkage, which supports credentialism as the primary explanatory frame.

Table 1. RQ1 nine-indicator coding matrix across the four higher education systems

RQ1 Indicator	China	United Kingdom	United States	Germany
1. Meritocratic legitimation	Explicit	Implicit	Implicit	Implicit
2. Admission selectivity	Moderate	Moderate	Moderate	Low
3. Institutional stratification	Strong	Moderate	Strong	Moderate
4. Assessment stakes	Moderate	High	Moderate	Moderate
5. Grading and ranking practices	Strong differentiation	Strong differentiation	Moderate differentiation	Moderate differentiation
6. Credential differentiation	Strong	Strong	Strong	Strong
7. Education–labour market linkage	Strong	Strong	Strong	Strong
8. Equity and compensatory arrangements	Partial provision	Extensive provision	Extensive provision	Partial provision
9. Institutional logic configuration	State = Strongly present; Market = Present; Academic-professional = Present; Managerial = Strongly present; Equity/community = Present	State = Present; Market = Strongly present; Academic-professional = Present; Managerial = Strongly present; Equity/community = Strongly present	State = Present; Market = Strongly present; Academic-professional = Strongly present; Managerial = Present; Equity/community = Strongly present	State = Strongly present; Market = Present; Academic-professional = Strongly present; Managerial = Present; Equity/community = Present

5.2. GenAI use–response configurations and alignment patterns (RQ2)

Table 2. Use–response configurations across the four higher education systems

Case	Overall configuration (C6)	Organizational form of partial alignment	Capability support (C4)	Equity alignment (C5)
China	Partially aligned	State-led pedagogical adaptation within targeted integrity controls	Partially aligned	Insufficient evidence
United Kingdom	Partially aligned	Rapid normalization with regulatory catch-up (conditional permission and disclosure)	Partially aligned	Insufficient evidence
United States	Partially aligned	Decentralized adaptive response across institutions	Partially aligned	Insufficient evidence
Germany	Partially aligned	Decentralized adaptive response localized at faculty/discipline/examination level	Partially aligned	Insufficient evidence

Note.Partial alignment is the cross-system norm (C6); capability support is the most persistent misalignment (C4 Partially aligned in all cases); equity alignment is unmeasured in every system (C5), constituting a monitoring-capacity gap. The United States and Germany illustrate equifinality: different institutional pathways converge on the same use–response configuration.

Partial use–response alignment is the dominant pattern across all four systems (Table 2, C6), though its organizational form varies substantially. Systems do not converge on a single regulatory model; they combine risk control, conditional authorization, pedagogical integration, and enabling conditions in different configurations [7].

Student-side variation reflects both behaviour and measurement infrastructure. GenAI prevalence (A1) ranges from insufficient comparable evidence in China to near-universal adoption in the UK and widespread adoption in the US and Germany. This variation cannot be read as purely behavioural, because national survey infrastructures measure different phenomena and possess unequal capacity to monitor student use. Prevalence is in any case an incomplete indicator of adoption: educational benefits depend on students' access, AI literacy, verification skills, and capacity to convert technological access into educational advantage [6, 17].

Across the remaining student-use indicators, instrumental affordance actualization (task support, drafting) is strongest in the UK and moderate elsewhere. Guarded actualization—verification, restraint, and preservation of independent competence—is strong in China and moderate in the UK but insufficiently documented in the US and Germany. Compensatory use remains underdocumented in all systems except the UK, an evidence gap rather than behavioural absence. Mastery-oriented use is consistently moderate across the four cases. A7 (access, literacy, and capability conditions) is uneven in every case, making institutional support the clearest common student-side constraint.

Institutional responses converge more than student-use measurement does. All four systems are coded Mixed on overall orientation (B1): governance neither purely restricts nor fully integrates GenAI. Restrictive controls are uniformly Targeted (B2), applying to particular assessment contexts or use types rather than constituting blanket bans. Control and pedagogical adaptation thus coexist within the same governance packages, contradicting the assumption that restriction and integration are opposing strategies.

The organizational form of this convergence differs by case. China combines Targeted integrity controls with Extensive pedagogical integration (B5) and Emerging assessment redesign (B4), a state-led model in which pedagogical adaptation occurs within defined integrity boundaries. The UK pairs Extensive conditional permission and disclosure (B3) with Emerging pedagogical integration, formalizing acceptable use through authorization and assessment-specific rules after rapid student normalization. The US implements Partial conditional permission with Emerging assessment redesign and pedagogical integration, decentralized across institutions. Germany shows the same coded profile—Partial B3, Emerging B4 and B5—but adaptation is localized in faculty-, discipline-, and examination-level governance [7].

The persistent misalignment concerns capacity rather than permission. B6 (AI literacy and support provision) is only Partial in all four systems, as is B7 (equity and access provision); correspondingly, C4 (capability-support alignment) is Partially aligned everywhere and C5 (equity alignment) is insufficiently measured in every case. The cross-system problem is not that students use AI while universities prohibit it, but that widespread or rapidly expanding use coexists with incomplete assessment redesign, partial literacy provision, and incomplete equity monitoring. This is consistent with evidence that institutional benefits depend on systematic support and enabling conditions [6] and with institutional studies identifying uneven training, resources, and guidance as persistent governance constraints [27]. Capability support is the most persistent cross-system misalignment mechanism.

Missing evidence is itself analytically distributed. China lacks comparable national prevalence data (A1), the US and Germany lack adequate measurement of guarded and compensatory use (A4–A5), and all four cases lack equity outcome evidence (C5). These gaps expose uneven secondary-data infrastructures and institutional blind spots in monitoring student needs and in evaluating whether governance responses reach different groups equitably.

6. Discussion

The "Insufficient evidence" codes that recur across the RQ2 results—A1 in China, A4 and A5 in the United States and Germany, and C5 in all four systems—are better read as properties of measurement infrastructure than of student behaviour. They indicate that higher education systems possess far better instruments for establishing whether students use GenAI than for establishing who benefits, who is disadvantaged, how students regulate their own use, or whether institutional support reaches target populations.

This interpretation follows from how the relevant literature has developed. Beckman et al. [6] show that the GenAI divide operates through access, AI literacy, and the capacity to convert technological resources into educational advantage rather than through access alone. Zhang et al. [17] demonstrate that socioeconomic status predicts AI literacy more strongly than conventional digital literacy, so differentiated survey instruments reveal inequalities that aggregate use measures obscure. Hou et al. [19] reach a parallel methodological conclusion in validating a reliance scale that separates reflective, cautious, thoughtless, and collaborative uses, on the grounds that conventional measures do not capture how students actually depend on GenAI during problem solving. Against this background, the uneven prevalence coding across the four cases is informative about instruments rather than behaviour: guarded and compensatory forms of use require purpose-built measurement, and where such instruments are absent, comparative coding cannot proceed. The equity dimension is sharper still. Systematic evidence mapping finds only partial visibility into how access, skills, and outcome inequalities intersect with GenAI adoption [28], and research on students with disabilities shows that GenAI can compensate for barriers in academic writing while simultaneously creating new inequalities through subscription costs, accuracy risks, and uneven institutional support [29]. Equity effects therefore cannot be inferred from aggregate adoption rates. Jiang [27] documents the institutional side of the same problem: students and faculty frequently face unclear guidance, uneven resources, and insufficient training, so governance capacity depends on whether institutions can identify where support needs are concentrated.

Uneven survey infrastructure should accordingly be treated as part of institutional response capacity itself. That C5 (equity alignment) is coded Insufficient evidence in all four systems is a governance finding as much as a literature gap: systems cannot align support, equity interventions, or assessment redesign with practices they do not systematically measure. Measurement capacity, on this account, is a component of governance capacity.

The second point concerns the joint shape of the RQ1 and RQ2 findings. Structurally different systems produce comparable credential differentiation and comparable partial GenAI alignment through distinct institutional pathways—an equifinality pattern at odds with single-system accounts of meritocratic competition and with the expectation that a shared technology drives systems toward a common governance model. On the RQ1 side, credential differentiation and education–labour market linkage are Strong in all four cases, while admission selectivity varies from Low in Germany to Moderate elsewhere without predicting post-entry credential pressure. The four pathways condense into distinct logic configurations: a state-coordinated pathway in China, where Strongly present State and Managerial logics transmit pressure through institutional hierarchy and evaluation [13, 25]; a marketized pathway in the UK, organized through Market, Managerial, and Equity/community logics [12]; decentralized market stratification in the United States, where institutional autonomy within a highly stratified field makes institutional location the decisive positional mechanism [1]; and qualification-differentiation with comparatively open entry in Germany, where State and Academic-professional logics organize pressure through qualification level and professional pathways [10]. RQ2 displays a parallel structure: partial alignment is universal, but it is organized as state-led pedagogical adaptation in China, rapid normalization with regulatory catch-up in the UK, and a shared decentralized adaptive response in the United States and Germany (Table 2). These results extend institutional-logics

research by showing that comparable pressures and comparable governance problems arise from different combinations of state, market, managerial, professional, and equity logics [5, 26]. What converges across systems is the problem structure—credential stakes, partial alignment, capacity gaps—rather than the institutional arrangements themselves.

Equifinality has two implications for comparative governance research. Because each governance arrangement is embedded in a national logic configuration, models cannot be transferred across systems without adjustment to the configuration in which they are implemented. At the same time, the comparison identifies problems that can be expected in any system regardless of pathway: credential pressure amid incomplete assessment redesign, uneven capacity support, and persistent measurement blind spots.

The third point returns to the mechanism behind partial alignment. The invariant codes across the RQ2 findings—Mixed response orientation (B1), Targeted restriction (B2), Emerging or Extensive pedagogical integration (B5), Partial literacy and support provision (B6), and Insufficient evidence on equity alignment (C5)—locate the misalignment in institutional capacity rather than in permission. The recurring configuration is expanding student use combined with incomplete assessment redesign, partial literacy provision, and unmonitored equity effects. Beckman et al. [6] supply the normative benchmark, arguing that institutions must monitor access, capability development, and students' ability to leverage GenAI for educational advantage; the present findings identify capability support as the most persistent cross-system misalignment mechanism. Jiang [27] documents the barriers that sustain this condition, and Li et al. [7] show that institutional capacity and enabling conditions matter alongside formal regulation. The finding also has a temporal dimension: in all four systems, permission and conditional-use rules have developed faster than the capability-building and equity-monitoring infrastructure that would make those rules effective.

These findings reframe the central question for GenAI governance research. The issue is not whether to restrict or integrate GenAI but whether institutions can measure where students need support and provide that support equitably. Governance failures, on this account, stem less from defective rules than from underdeveloped measurement and support systems, and policy–practice alignment depends more on investment in monitoring and capability infrastructure than on further refinement of formal policies.

7. Policy implications

First, governance should intervene in post-entry credential and assessment pathways, not only at admission. The four-case findings show that credential pressure is reproduced through grading, classification, institutional stratification, qualification structures, and labour-market linkage even where admission selectivity is relatively limited. For China, reform should therefore not focus solely on gaokao competition or entry access; universities also need to review how GPA, ranking, scholarships, postgraduate selection, honours, and employer-facing credentials intensify positional competition after enrolment. Assessment reform should reduce unnecessary high-stakes differentiation while preserving meaningful standards, for example through more authentic, process-based, and competence-oriented assessment—a need sharpened by evidence that traditional assessment designs are increasingly vulnerable to GenAI substitution and may encourage strategic rather than educationally productive use [8]. For ASEAN, the same principle suggests linking the ASEAN Qualifications Reference Framework (AQRF) more closely to transparent learning outcomes and skills rather than allowing qualification differentiation to become an opaque hierarchy of institutional prestige. The AQRF already provides a regional mechanism for comparing qualifications while respecting national diversity, offering a basis for strengthening learning-outcome transparency and mobility without imposing uniformity [9].

Second, universities should prioritize institutional capacity over increasingly detailed permission rules. Across all four cases, targeted restrictions already coexist with pedagogical integration; the more persistent gap is uneven AI literacy, tool provision, assessment redesign, and staff and student support. China should therefore complement its strong policy-level promotion of AI-enabled education with institution-level investment in staff development, student AI literacy, secure approved tools, and assessment redesign. For ASEAN, where digital infrastructure and educational capacity remain uneven across member states, common minimum expectations for AI literacy, faculty development, and accessible institutional tools would be more consequential than attempts to harmonize detailed GenAI prohibitions. ASEAN's digital-literacy policy agenda likewise links educational transformation to talent and infrastructure development. Equitable access, GenAI literacy, institutional infrastructure, professional development, and continuous policy review are consistently identified as foundational components of effective higher-education AI governance [6, 7].

Third, comparable student-use survey infrastructure should be treated as a governance function. China and ASEAN should develop recurring, multi-institution instruments measuring not only prevalence but also instrumental, guarded, compensatory, and mastery-oriented use; AI literacy; access conditions; assessment-related use; and equity differences. Performance-based AI-literacy measures should supplement self-reports where feasible, because they predict actual GenAI task performance better than self-assessment alone [30]. Without comparable monitoring, institutions cannot determine whether policy, support, and equity interventions match student practice. The same holds for other systems: mature GenAI governance requires measurement, capability-building, and assessment reform alongside regulation, rather than permission rules alone.

8. Conclusion

This study shows that strong credential differentiation and education–labour market linkage can emerge through markedly different post-entry institutional pathways across China, the United Kingdom, the United States, and Germany. Admission selectivity alone does not explain the intensity or location of credential pressure: China combines state-coordinated stratification and strong grading differentiation; the UK concentrates pressure more visibly in degree classification and market-regulatory outcomes; the US relies on decentralized institutional stratification and locally governed assessment; and Germany combines relatively open system-wide entry with qualification differentiation and strong field- and credential-based labour-market linkage. These divergent pathways nonetheless generate similar positional stakes, illustrating equifinality in how higher-education systems translate institutional arrangements into credential value. A parallel pattern appears in GenAI governance. Partial use–response alignment is the cross-system norm: blanket prohibition is absent, targeted restrictions coexist with pedagogical integration, and the most persistent gap concerns uneven AI literacy, support provision, assessment redesign, and institutional capacity rather than permission alone. The US and Germany are especially instructive because they occupy different RQ1 institutional positions but share the same coded RQ2 configuration, demonstrating that distinct institutional logics can produce equivalent use–response outcomes. The systematically distributed "Not measured" and "Insufficient evidence" codes also constitute findings in their own right: the absence of comparable evidence on equity alignment, guarded use, and compensatory use reveals uneven monitoring infrastructures and identifies measurement capacity as part of institutional governance rather than merely a methodological issue.

These conclusions should be interpreted within several limitations. First, the study relies entirely on secondary data, so it identifies system-level configurations and plausible institutional mechanisms rather than individual-level causal effects. Second, the student surveys are not fully equivalent across countries in

sampling frame, timing, wording, population coverage, or measurement of GenAI use; raw prevalence estimates therefore cannot be treated as directly comparable national rates. Third, the inferential boundary remains at the level of higher-education systems, institutional arrangements, and aggregate student-use patterns, and the analysis cannot determine whether particular forms of credential pressure cause particular students to adopt specific GenAI practices. Fourth, the China coding for A3–A6 relies heavily on one large multi-institution study, which improves internal consistency but narrows source triangulation for those affordance patterns. Fifth, all four cases were coded Partially aligned at C6, limiting outcome variation and making it difficult to identify conditions associated with clearly aligned or clearly misaligned configurations. Future research should therefore combine harmonized cross-national student surveys with institution-level policy data, assessment designs, and longitudinal evidence; expand measurement of guarded, compensatory, and equity-related GenAI use; include cases exhibiting stronger variation in C6 outcomes; and, where possible, add primary interviews or panel data to test whether the institutional pathways identified here systematically shape student practices over time.

As GenAI becomes embedded in everyday study, the governance of meritocratic competition shifts from admission gates and permission rules to the post-entry institutions of assessment, support, and measurement. Whether GenAI use becomes educationally productive will depend less on the rules systems write than on the capacities they build—a conclusion that extends beyond the four cases examined here.

References

[1] Taylor, B. J., Rosinger, K., & Ford, K. S. (2024). The shape of the sieve: Which components of the admissions application matter most in particular institutional contexts? *Sociology of Education, 97*(3), 233–251.

[2] Hansen, A. T., Hvidman, U., & Sievertsen, H. H. (2024). Grades and employer learning. *Journal of Labor Economics, 42*(3), 659–682.

[3] Martinussen, M., Singh, N., & Rangarajan, S. (2025). 'The standard is the standard': Constructing meritocratic student selves through sense-making on academic writing. *British Journal of Sociology of Education, 46*(3), 358–374.

[4] Yang, Y., Luo, J., Yang, M., Yang, R., & Chen, J. (2024). From surface to deep learning approaches with generative AI in higher education: An analytical framework of student agency. *Studies in Higher Education, 49*(5), 817–830. https://doi.org/10.1080/03075079.2024.2327003

[5] Cai, Y., & Mountford, N. (2022). Institutional logics analysis in higher education research. *Studies in Higher Education, 47*(8), 1627–1651.

[6] Beckman, K., Apps, T., Howard, S. K., Rogerson, C., Rogerson, A., & Tondeur, J. (2025). The GenAI divide among university students: A call for action. *The Internet and Higher Education, 67*, Article 101036. https://doi.org/10.1016/j.iheduc.2025.101036

[7] Li, M., Xie, Q., Enkhtur, A., Meng, S., Chen, L., Yamamoto, B. A., Cheng, F., & Murakami, M. (2026). A framework for developing university policies on generative AI governance: A cross-national comparative study. *Studies in Higher Education*. Advance online publication. https://doi.org/10.1080/03075079.2026.2696496

[8] Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A call to reconsider the "originality" of students' work. *Assessment & Evaluation in Higher Education, 49*(5), 651–664.

[9] Association of Southeast Asian Nations. (2016). *The ASEAN Qualifications Reference Framework*. ASEAN. https://aanzfta.asean.org/uploads/docs/AQRF/ASEAN_Qualifications_Reference_Framework_January_2016.pdf

[10] Finger, C., & Solga, H. (2023). Test participation or test performance: Why do men benefit from test-based admission to higher education? *Sociology of Education, 96*(4), 344–366.

[11] François, M., & De Witte, K. (2025). A decade of grade inflation boosted by the COVID-19 pandemic—An empirical analysis of a top European university. *British Educational Research Journal, 51*(5), 2271–2339.
[12] Martini, M., & Robertson, S. L. (2022). UK higher education, neoliberal meritocracy, and the culture of the new capitalism: A computational-linguistics analysis. *Sociology Compass, 16*(12), Article e13020.
[13] Zheng, Y., & Postiglione, G. A. (2025). Untangling the dominant culture in China's elite universities. *Higher Education, 90*, 407–427.
[14] Theoharakis, V., Mylonopoulos, N., & Papadopoulou, K. (2026). AI's learning paradox: How business students' engagement with AI amplifies moral disengagement-driven misconduct. *Studies in Higher Education, 51*(7), 1590–1607.
[15] Zagami, J. (2026). Guarded adoption of generative AI in higher education: High-achieving students, successful-student identity, and epistemic agency in a single-university mixed-methods survey. *International Journal of Educational Technology in Higher Education, 23*, Article 47.
[16] Hooley, Z., Williams, S., & Hancock-Smith, B. (2026). Impact of text-generative artificial intelligence tools on students' approach to assessment: A case study of a UK institution. *Studies in Higher Education*. Advance online publication.
[17] Zhang, C. X., Rice, R. E., & Wang, L. H. (2026). College students' literacy, ChatGPT activities, educational outcomes, and trust from a digital divide perspective. *New Media & Society, 28*(2), 673–695.
[18] Sandel, M. J. (2020). *The tyranny of merit: What's become of the common good*. Farrar, Straus and Giroux.
[19] Hou, C., Zhu, G., Sudarshan, V., Lim, F. S., & Ong, Y. S. (2025). Measuring undergraduate students' reliance on generative AI during problem-solving: Scale development and validation. *Computers & Education, 234*, Article 105329. https://doi.org/10.1016/j.compedu.2025.105329
[20] Bouncken, R. B., Qiu, Y., Sinkovics, N., & Kürsten, W. (2021). Qualitative research: Extending the range with flexible pattern matching. *Review of Managerial Science, 15*(2), 251–273. https://doi.org/10.1007/s11846-021-00451-2
[21] Collins, C., Neely, M. T., & Khan, S. (2024). "Which cases do I need?" Constructing cases and observations in qualitative research. *Annual Review of Sociology, 50*, 21–40.
[22] Ortmanns, V. (2020). Explaining inconsistencies in the education distributions of ten cross-national surveys—The role of methodological survey characteristics. *Journal of Official Statistics, 36*(2), 379–409.
[23] Organisation for Economic Co-operation and Development. (2023). *Education at a glance 2023: Sources, methodologies and technical notes*. OECD Publishing.
[24] Aguilar Solano, M. (2020). Triangulation and trustworthiness—Advancing research on public service interpreting through qualitative case study methodologies. *FITISPos International Journal, 7*(1), 31–52.
[25] Zhao, L., & Shih, T. (2026). Institutional logics and academic work in China: Career structures, governance mechanisms, and researcher responses. *Higher Education Policy*. Advance online publication.
[26] Pietilä, M., & Pinheiro, R. (2021). Reaching for different ends through tenure track—Institutional logics in university career systems. *Higher Education, 81*(6), 1197–1213. https://doi.org/10.1007/s10734-020-00606-2
[27] Jiang, S. (2025). Exploring the effectiveness of institutional policies and regulations for generative AI usage in higher education. *Higher Education Quarterly, 79*(4), Article e70054.
[28] Jamal Eddine, R., Gide, E., & Al-Sabbagh, A. (2026). Systematised evidence mapping of generative artificial intelligence (GenAI) and digital divide phenomena in higher education. *Discover Computing, 29*, Article 157.
[29] Zhao, X., Cox, A., & Chen, X. (2025). The use of generative AI by students with disabilities in higher education. *The Internet and Higher Education, 66*, Article 101014.
[30] Jin, Y., Martinez-Maldonado, R., Gašević, D., & Yan, L. (2025). GLAT: The generative AI literacy assessment test. *Computers and Education: Artificial Intelligence, 9*, Article 100436. https://doi.org/10.1016/j.caeai.2025.100436