SCIENTIFIC Publishing Limited

Journal of Qualitative Research in Education
https://ojs.ukscip.com/index.php/jqre

Article

Teachers' Perspectives on AI Use for Teaching Efficiency in Vocational Education in Shandong, China: Mechanisms and Enabling Conditions

Wenpang Liu [1,2], Goh Ying Soon [3,*] and Sujing Lin [1,4]

[1] School of Social Sciences, Humanities & Law, Teesside University, Middlesbrough TS1 3BX, UK

[2] Liaocheng Infant Normal School, Liaocheng 252600, China

[3] Faculty of Education, Nilai University, Nilai 71800, Malaysia

[4] School of Food and Pharmaceutical Engineering, Shenzhen Polytechnic University, Shenzhen 518055, China

[*] Correspondence: drysgoh@nilai.edu.my

Received: 6 February 2026; **Revised:** 20 March 2026; **Accepted:** 24 April 2026; **Published:** 20 May 2026

Abstract: Discussion of artificial intelligence in education often moves too quickly from access to tools to claims of improved efficiency. The interview evidence in this study does not support such a direct conclusion. This article explores how teachers and administrators in Shandong, China, described their use of AI in technical and vocational education and training (TVET). It focuses on the situations in which AI appeared to reduce routine work, and the situations in which it generated additional checking, explanation, or risk. Eight semi-structured interviews were conducted across secondary, higher, and adult vocational institutions. The interviews were recorded and transcribed in Chinese. English working versions were then prepared with AI assistance for analysis and reporting. During coding and manuscript preparation, the first author returned to the Chinese transcripts to verify key terms and selected quotations. The data were analysed in NVivo through iterative coding, memo writing, and comparison across teacher and administrator accounts. Findings are organised through a Constraints–Mechanisms–Enablers–Outcomes (CMEO) model. Participants did report some efficiency gains, but only in a limited sense. These gains were mainly associated with less time spent on first-draft writing, formatting, and routine screening. They also described new work. Teachers still had to verify technical accuracy, explain AI-supported feedback to students, and adapt outputs to local equipment, syllabi, and safety rules. In higher-risk settings, AI use was accepted only when teachers retained decision authority and could justify the result. For that reason, this paper treats AI-enabled teaching efficiency as a situated professional judgement rather than as a direct measure of productivity.

Keywords: Artificial Intelligence; Technical and Vocational Education and Training (TVET); Teachers' Perspectives; Perceived Efficiency; Thematic Analysis; Human-AI Co-Assessment; Data Governance

1. Introduction

AI now appears in routine teaching in many institutions. By itself, that tells us very little. In technical and vocational education and training (TVET), the practical issue is not whether AI can produce text quickly. The more relevant question is whether the output can withstand local equipment constraints, assessment criteria, and safety rules. If it cannot, speed has limited value [1–4].

Vocational teaching sits between school and workplace. Teachers have to turn process sheets, operating instructions, diagnostic routines, and quality requirements into tasks that students can actually use. This work is highly local. Machines differ across institutions. Enterprise data may be restricted. Some procedures carry obvious safety implications. Any claim about teaching efficiency in TVET therefore has to specify which tasks became shorter, what new work was added, and where new risks entered the workflow [5–7].

Since late 2022, generative AI and related large-language-model tools have made drafting easier. They can produce explanations, worksheet text, rubric language, and feedback comments within seconds. That convenience is real, but it is not enough. In technical fields, a polished answer may still be wrong, incomplete, or detached from local practice. These tools also blur source attribution and can invite premature trust when the first answer looks plausible [8–11].

Existing research helps to frame the issue, but it does not settle it. Formative feedback can shorten learning cycles [12]. Immersive tools may support skill development when they are tied to explicit pedagogy [13]. Learning analytics may assist decision-making when the data are interpretable and usable [14]. Even so, these studies do not tell us when teachers in vocational colleges judge AI worth the effort in daily work, or when later checking, correction, and explanation wipe out the apparent saving.

In this paper, teaching efficiency is treated as a reported change in work rather than as a verified output metric. Participants used the term when routine drafting took less time, when repetitive feedback work became lighter, or when they could redirect attention to explanation and problem-solving. That is a narrower claim than institutional efficiency. Interview data can support that narrower claim. They cannot establish strong causal effects on performance.

This study addresses that gap by examining how teachers and administrators in Shandong described four matters: the constraints that blocked efficient AI use, the practical responses they adopted, the conditions that made those responses workable, and the outcomes they attached to them. Two research questions guided the analysis: (RQ1) What constraints do TVET teachers and administrators face when integrating AI into efficiency-oriented teaching routines? (RQ2) What mechanisms and enabling conditions support ongoing AI-enabled teaching efficiency?

Two design choices follow from that framing. First, the article does not treat efficiency as a fixed quantity that can be inferred from output volume or production speed. In vocational settings, accuracy, traceability, industry fit, and safety all matter. Saving ten minutes is irrelevant if the material still has to be rebuilt before use. Second, teacher and administrator accounts are treated as evidence of work process and professional judgement, not as proof that AI improved performance.

That distinction matters because discussion of AI in education often compresses several questions into one. A tool may be accessible, widely used, and described in positive terms, yet still produce uneven results when it enters assessment-heavy or safety-sensitive work. This article stays with participants' accounts of routine teaching activity. It asks what changed, where they found AI useful, and where the extra checking was not worth the effort.

2. Method

This study uses a qualitative, interpretivist design based on semi-structured interviews. The aim is deliberately narrow. We are not testing whether AI improves teaching outcomes. We are examining how teachers and administrators described AI use in their work, where they found value, and where they encountered limits, trade-offs, or failure points. The study is concerned with situated professional judgement and changes in work processes. A qualitative design suits that purpose [15–17].

2.1. Participants and Setting

Eight participants were selected through purposive sampling to capture variation in role, experience, institution type, and subject area (see **Table 1**). The sample included frontline teachers and administrators from secondary vocational schools, higher vocational colleges, and adult higher vocational colleges. Sampling followed an information-rich logic [18, 19]. Analytic sufficiency was judged during late-stage coding rather than fixed in advance. By Interview 6, the four CMEO families were already established and no new top-level thematic family had emerged. Interviews 7 and 8 added examples, contrasts, and boundary cases, but they did not require new top-level categories or a restructuring of the codebook. For this study, that pattern was treated as adequate coverage of the research problem. The limit is also clear: participants were selected because they had recent experience with

AI-related teaching work, so the material is more likely to reflect informed users than complete non-users.

Table 1. Participant characteristics (n = 8).

ID	Role	Teaching Experience	Institution Type	Subject Taught/ Administrative Duty	Main Interview Focus Themes
T01	Teacher	11 years	Higher Vocational College	CNC Technology	Effectiveness of ITS, Technical Challenges
T02	Teacher	6 years	Secondary Vocational School	Nursing	AI Content Generation, Teacher Training Needs
T03	Teacher	15 years	Adult Higher Vocational College	Automotive Repair	Advantages of VR/AR, Assessment Challenges
T04	Teacher	8 years	Secondary Vocational School	Preschool Education	Automated Assessment, Data Privacy Concerns
T05	Teacher	10 years	Higher Vocational College	E-commerce	Learning Analytics, Data Interpretation
A01	Admin	20 years	Higher Vocational College	Dean of Academic Affairs	Policy Support, Resource Allocation, Infrastructure
A02	Admin	12 years	Adult Higher Vocational College	Director of Training	Teacher Training Effectiveness, Community Building
A03	Admin	18 years	Secondary Vocational School	Vice Principal	Strategic Planning, Change Management

2.2. Data Collection

Interviews were conducted face to face or online and lasted roughly 30 to 60 min. A semi-structured guide covered AI-supported teaching routines, implementation problems, and forms of institutional support. Illustrative interview questions are provided in **Appendix A**. All interviews were conducted in Chinese and audio-recorded with consent. Recordings were transcribed in Chinese and anonymised before analysis. English working transcripts were then produced with AI assistance for coding and reporting. The first author reviewed the translated material against the Chinese source when preparing the working corpus and again when selecting quotations for the manuscript. In addition, key expressions and theme-relevant wording were checked back against the original Chinese transcripts whenever translation choices could affect thematic interpretation. No back-translation was undertaken. Nor was there a second bilingual review or member checking of translated excerpts. Those absences narrow the strength of any cross-language claim and are treated here as methodological constraints, not as minor technicalities.

2.3. Data Analysis

The English working transcripts were analysed in NVivo 14. Coding moved from descriptive fragments to a more structured account of how constraints, mechanisms, enablers, and outcomes related to one another [20]. Braun and Clarke's thematic analysis and qualitative content-analysis procedures informed that process [21,22]. Early coding stayed close to participants' descriptions of tasks, frictions, and workarounds. Later rounds grouped those codes into the CMEO structure. Although the English working transcripts served as the coding corpus, key code definitions and selected quotations were checked back against the Chinese originals when thematic meaning was ambiguous, technically loaded, or conceptually sensitive. Translation was therefore handled cautiously rather than treated as a neutral preprocessing step. **Figure 1** shows the analytic workflow while **Table 2** shows the coding trail used to develop the model. The three-stage coding process (open, axial, and selective) with illustrative examples is detailed in **Appendix B** (**Table A1**).

Trustworthiness was supported through a documented codebook, an audit trail of coding decisions, excerpt IDs linked to a secure log, and reflexive memoing across the full coding process. We also compared accounts across roles and sites, retained negative cases, and revised provisional code definitions when a quotation did not fit well or when a category had become too broad. No independent peer coding, researcher triangulation, member checking, back-translation, or second bilingual review was conducted. To reduce the resulting risk of individual bias, coding decisions were revisited across the full dataset, discrepant cases were retained rather than forced into dominant categories, and key interpretive decisions involving translation-sensitive wording were checked against the Chinese source material.

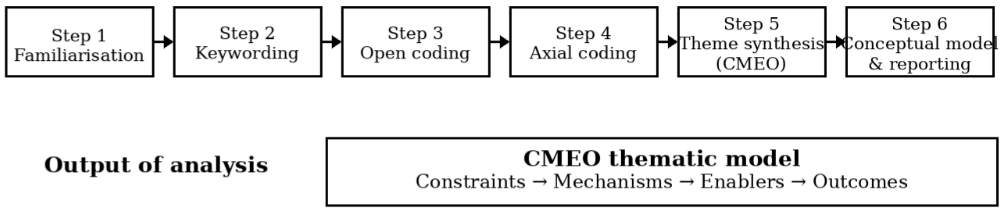

Figure 1. Inductive thematic analysis workflow (study-specific).

Table 2. Illustrative coding trail from quotation to CMEO theme.

Interview Excerpt	Keywords	Code	CMEO Family	Interpretive Memo/Concept
"Enterprise tooling and process data are often confidential." (T01)	data confidentiality	C1 Data access and confidentiality	Constraints	Limited authentic enterprise data reduces the realism of tasks and analytics.
"Only know how to use AI,not how to explain…" (T01)	use vs. explain parameter reasoning	C2 Use–Explain Gap	Constraints	Assessment validity requires visible reasoning,not output-only performance.
"Explanation precedes points." (T01)	explanation-first rubric	M1 Explainable rubrics	Mechanisms	Rubric redesign makes AI-assisted work assessable and defensible.
"No AI-only scoring without teacher review." (A02)	human review guardrails	E3 Governance guardrails	Enablers	Governance-by-design limits misuse and stabilises classroom adoption.

2.4. Trustworthiness and Ethics

Trustworthiness was supported through a documented codebook, an audit trail of coding decisions, excerpt IDs linked to a secure log, and reflexive memoing across the full coding process. We also compared accounts across roles and sites, retained negative cases, and revised provisional code definitions when a quotation did not fit well or when a category had become too broad. No independent peer coding, researcher triangulation, member checking, back-translation, or second bilingual review was conducted. To reduce the resulting risk of individual bias, coding decisions were revisited across the full dataset, discrepant cases were retained rather than forced into dominant categories, and key interpretive decisions involving translation-sensitive wording were checked against the Chinese source material.

The credibility strategy in this study is therefore limited rather than maximal. The paper should be read as a documented interpretation of participants' accounts, not as a triangulated verification exercise and not as evidence of exact equivalence between Chinese transcripts and English quotations. That matters for the scope of the claims. Interview data can show how participants described workflow change, risk, and acceptable use. They cannot, on their own, verify classroom effects or institutional productivity gains.

3. Findings

Findings are organised through the CMEO model. As shown in **Table 3**, the model is built from four linked dimensions: constraints, mechanisms, enablers, and outcomes. The framework is used analytically rather than descriptively. It traces how particular barriers led to particular redesign choices, how institutional conditions affected whether those choices held, and why reported outcomes varied across tasks and settings.

3.1. Constraints: Systemic and Cognitive Barriers

Participants described several constraints that limited efficient AI use in TVET. Some were structural. Restricted enterprise data reduced the realism of examples and weakened domain-specific modelling (C1). Others were concerned with reasoning and control. The use–explain gap meant that students could reproduce an answer without being able to justify a parameter or a procedural step (C2). Teachers also described overload when AI produced too many suggestions across long technical chains (C3). Feasibility problems appeared when AI output ignored machine conditions, fixtures, tool stock, or workshop sequence (C4). In safety-sensitive subjects, interviewees worried about hallucinated guidance and weak transfer from simulation to real judgement (C5). They also reported shortcut diagnosis and age-related pace problems (C6–C7). Taken together, these constraints show why AI could add work rather than remove it. Teachers often had to filter, ground, and repair the output before it became usable.

Table 3. The CMEO framework: themes, codes, definitions, and exemplar interview quotations.

Theme	Code	Definition	Exemplar Quotation (Speaker, Excerpt ID)
Constraints	C1 Data Access & Confidentiality	Enterprise process/tooling data are not easily available for teaching or model training.	"Three aspects: (1) Data scarcity—Enterprise tooling and process data are often confidential." (T01, Ex1) (2) Use–explain gap—Learners apply AI outputs without grasping the underlying reasoning. (3) "Cognitive overload—Lengthy multi-stage workflows combined with AI's high-volume suggestions exceed students' processing capacity." (T01, Ex1)
Constraints	C2 Use–Explain Gap	Students can reproduce AI-supported answers or procedures but cannot explain why a parameter, step, or recommendation is appropriate.	"Students tend to 'only know how to use AI, not how to explain' — they can run G-code but can't explain why parameters are set that way." (T01, Ex2)
Constraints	C3 Cognitive Overload	Too many AI recommendations across long task chains increase cognitive load and uncertainty.	"Excessive cognitive load — The chain from CAD/CAM to CMM inspection is long, and AI gives too many recommendations at once..." (T01, Ex3)
Constraints	C4 Real-World Constraints Gap	AI suggestions ignore shop-floor constraints (rigidity, fixtures, inventory), reducing feasibility.	"If AI doesn't account for real-world constraints like machine rigidity, fixtures, or tool inventory..." (T01, Ex4)
Constraints	C5 Safety/Ethics & Hallucinations	In safety-critical fields, hallucinated guidance or weak judgement transfer undermines trust.	"Potential for hallucinations in AI medical Q&A; students perform well in the simulation lab but show weak emergency judgment in the ward." (T02, Ex5)
Constraints	C6 Shortcut Diagnosis	Over-reliance on fault-code lookups discourages systematic troubleshooting.	"Students rely on 'input fault code → get answer,' neglecting systematic troubleshooting." (T05, Ex6)
Constraints	C7 Age-appropriate Pace Control	Need developmental safeguards and pace control to prevent overload and misuse.	"Age-appropriateness and ethical review are the hardest… Pace control is also key." (T04, Ex7)
Mechanisms	M1 Explainable Rubrics	Assessment is redesigned so explanations precede points, making AI outputs defensible.	"Treat AI as an optimizer within constraints… establish a scoring system where 'explanation precedes points.' " (T01, Ex8)
Mechanisms	M2 Human-AI Co-assessment	Dual-track verification: AI performs structured screening, teachers make final judgement.	"OSCE + AI co-assessment… Dual-track verification… After each simulation, submit an SBAR report." (T02, Ex9)
Mechanisms	M3 Evidence-chain Diagnostics	Reasoning is disciplined via required multi-layer evidence (signals, trends, tests).	"Each diagnosis must submit (a) Sensor waveform (b) Data stream trend (c) Substitution/Bypass test results… Final judgment requires a reasoning chain." (T05, Ex10)
Mechanisms	M4 Explainability via Analytics	Teachers use interpretable analytics to justify feedback and make learning bottlenecks visible.	"The learning analytics dashboard allows me to identify students' comprehension bottlenecks in advance, thereby reducing repetitive explanations." (T03, Ex11)
Enablers	E1 PD Matrix & Data Sandbox	Three-tier professional development plus a sandbox for safe experimentation and compliance.	"We built a three-tier PD matrix and a data sandbox system to let teachers experiment safely while ensuring compliance." (A01, Ex12)
Enablers	E2 Model Routing & Edge Inference	On-campus model routing and edge inference reduce latency and control data flows.	"For us, it's an on-campus model routing platform plus edge inference." (A02, Ex13)
Enablers	E3 Governance Guardrails	Technical and policy guardrails prevent unsafe use (e.g., AI-only scoring; sensitive uploads).	"We can technically enforce policies like 'no direct upload of sensitive personal data' and 'no AI-only scoring without teacher review'." (A02, Ex14)
Enablers	E4 Outcome-linked Accountability	Governance uses flagged cases as feedback to improve assignment design, not only punishment.	"Each flagged case requires the course team to submit a short note… we see assignment design gradually improving." (A01, Ex15)

Table 3. *Cont.*

Theme	Code	Definition	Exemplar Quotation (Speaker, Excerpt ID)
Enablers	E5 Industry Tri-alignment	Enterprise question bank - course project - certification standard alignment with data desensitization.	"Implement 'Enterprise Question Bank-Course Project-Certification Standard Tri-Alignment'... include 'data collection-annotation-desensitization' in special funding." (A03, Ex16)
Outcomes	O1 Efficiency as Time Reallocation	Teachers report shifting effort from routine checking to judgement, coaching, and remediation.	"AI only assesses structural completeness; I provide the decision-making comments." (T02, Ex17)
Outcomes	O2 Value-added Evaluation Orientation	Administrators emphasise moving from tool counts to value-added learning and employment readiness.	"We want to move from counting 'how many AI tools are used' to measuring what value is added." (A01, Ex18)

Note: Excerpt IDs (Ex1–Ex18) refer to an excerpt log kept by the research team that records transcript file names and time ranges, while exact timestamps are masked in the manuscript to reduce re-identification risk.

3.2. Mechanisms: Pedagogical and Design Responses

Teachers responded by redesigning assessment and feedback routines rather than by handing those routines over to the tool. Explanation-first rubrics (M1) required students to justify parameter choices before receiving credit. In higher-stakes assessment, human–AI co-assessment (M2) lets AI screen for structure while leaving final judgement to the teacher. In troubleshooting tasks, evidence-chain diagnostics (M3) required signals, trends, and test results so that students had to show how they reached a conclusion. Where analytics dashboards were available, teachers used them to flag bottlenecks and focus on later explanation (M4). These mechanisms did not make AI trustworthy on their own. They narrowed down where and how it could be used.

3.3. Enablers: Institutional and Infrastructural Conditions

Participants also stressed that these teacher-level responses depended on institutional conditions. Professional development and data sandboxes made cautious experimentation possible (E1). On-campus model routing and edge inference reduced latency and made data flow more controllable (E2). Governance guardrails limited unsafe practices, such as direct upload of sensitive material or AI-only scoring (E3). Outcome-linked accountability turned flagged cases into feedback for assignment redesign rather than punishment alone (E4). Industry tri-alignment connected enterprise materials, course projects, and certification standards (E5). These conditions mattered because they stabilised the extra verification work described by teachers [23,24].

3.4. Perceived Shifts in Work and Time Use

Participants did not use efficiency to mean "doing more in less time." Their accounts were narrower. Some tasks became shorter, especially first-pass drafting, formatting, and the initial sorting of routine answers. Other tasks became longer. Teachers still had to check language, correct examples, adapt material to local machines or syllabi, and decide when an AI-supported answer could be trusted. The work did not simply shrink. It shifted.

The interviews also suggest that the workflow change was uneven rather than settled. Participants were not describing a clean move from pre-AI work to post-AI work. They were describing a period of adjustment. Some steps became faster. Other responsibilities became more visible, especially the need to explain acceptable use, verify outputs, and define where human judgement still had to sit. In that sense, AI redistributed work across drafting, checking, revision, and explanation.

Time use was social as well as technical. A teacher might save time preparing worksheet material but the whole process is time-consuming if the output produced later disputes, distrust, or extra checking during marking. Administrators made a similar point from another angle. They cared less about whether one task became shorter than about whether AI-supported work could be defended to students, parents, enterprise partners, and internal quality teams. In this dataset, legitimacy affected efficiency almost as much as speed.

3.5. Descriptive Indicators for Analytic Transparency

As shown in **Table 4**, the descriptive indicators are included for analytic transparency, not for statistical inference. They show the size and spread of the sample, the balance between teacher and administrator accounts, and the final shape of the codebook that fed into the CMEO model. They do not demonstrate prevalence or effect size. Their value is procedural. They show that the model was built from a bounded but varied dataset and that the analytic frame remained controlled rather than diffuse. The indicators also support the saturation claim in a limited way: the final framework remained fixed at four families and 18 codes while the last interviews added clarification rather than new top-level categories.

Table 4. Descriptive indicators of the qualitative sample and CMEO coding output.

Indicator	Value	Notes
Number of interviews	8	Semi-structured interviews with teachers and administrators.
Participant roles	Teachers: 5; Administrators: 3	Purposive sampling across roles.
Institution types	Higher Vocational College: 3; Secondary Vocational School: 3; Adult Higher Vocational College: 2	Secondary, higher, and adult vocational settings.
Teaching experience (years)	Range: 6–20; Median: 11.5	Self-reported years in role/sector.
Codebook size (CMEO)	18	Codes summarised in **Table 3** (Constraints/Mechanisms/Enablers/Outcomes).
Codes by CMEO category	Constraints: 7; Mechanisms: 4; Enablers: 5; Outcomes: 2	Counts indicate analytical coverage, not effect size.

4. Discussion

Our interviews do not support a simple efficiency claim. Participants did say that some tasks became quicker. They also described the checking, correction, and explanation that followed. In TVET, speed has little value if the output does not fit local equipment, assessment requirements, or safety procedures [1,8,10,25]. For that reason, the paper treats efficiency as a bounded professional judgement rather than as a neutral productivity metric. The evidence here concerns reported practice. It does not establish objective gains in institutional efficiency or student outcomes.

The CMEO model helps explain why. Constraints were not background noise. They shaped the redesign work itself. The pattern in the data was not random: use–explain problems and assessment defensibility concerns more often pushed teachers toward explanation-first rubrics and teacher-controlled co-assessment, whereas feasibility gaps between generic output and workshop conditions more often led to evidence-chain diagnostics and manual adjustment. Enablers then affected whether those responses could be sustained. Professional development, sandboxed experimentation, model routing, and explicit governance rules lowered uncertainty and made verification easier to carry out. When those conditions were absent, the apparent gain often collapsed into more correction work.

This also clarifies the role of institutional enablers. They matter because they reduce the cost of cautious use. When teachers know which tools are allowed, what data may be entered, and when teacher review is mandatory, routine use becomes more stable. When those rules are missing, the apparent gain often disappears into hesitation, rework, and informal workaround labour. Recent work on generative AI in education points in the same direction. UNESCO treats human oversight and context-sensitive governance as basic conditions rather than optional add-ons [2]. Kasneci et al. likewise show that large language models generate new demands for verification and scrutiny [8]. Our data do not simply repeat those points. They show how those concerns become sharper in vocational settings, where error costs, local equipment fit, assessment defensibility, and safety demands directly shape whether AI-assisted routines remain workable.

Participants were careful with the term "efficiency." They used it for ordinary teaching work: repeating instructions, drafting worksheet variants, screening common errors, or rephrasing feedback so that more time could be given to cases needing judgement. They were not claiming general improvements in institutional performance. They were also clear that new work entered the process. Verification, tailoring, and explanation took time. A more accurate description is that AI reallocated effort rather than removed it.

The model also shows why some tasks were easier to support with AI than others. AI worked best when the input was stable and the expected output was tightly specified: a known rubric, a standard lesson structure, a familiar administrative format. It was less dependable when the work depended on tacit knowledge, real-time diagnosis, or safety-sensitive judgement. This task-by-task variation matters. General adoption models tend to foreground usefulness and ease of use. They say much less about auditability, error cost, and the local consequences of being wrong. That is where this study adds something specific to current research on AI in education [2,8,23,25].

Legitimacy also mattered in practical terms. When students did not trust the basis of feedback, teachers spent more time explaining, defending, and repairing. When explanation was built into the workflow, there was less downstream friction. In that narrow sense, explainability was not just an ethical principle. It was one of the conditions under which participants judged AI-supported routines to be workable. The institutional context changed the meaning of use as well. Higher-risk workshops allowed little tolerance for uncertain output. Lower-risk tasks, such as language polishing or lesson structuring, allowed more experimentation. A single policy for all tasks would miss that difference.

This leads to a direct criticism of current institutional uptake. Some schools still treat tool access as the main implementation problem. The interviews do not support that view. Access matters. Access without rules, role clarity, and task differentiation simply shifts risk onto individual teachers. They are then left to decide alone which data may be entered, which outputs require checking, and how to respond when student AI use blurs authorship or competence claims. That is a weak implementation model.

A related problem is the separation of policy from classroom design. Participants' accounts show that the two are connected at each step. Decisions about allowable use shaped lesson planning, assessment design, evidence requirements, and teacher workload. The stronger cases in this study were not the ones with the broadest policy language. They were the ones where practical expectations were tied to ordinary work and where governance rules could be applied to specific tasks [26,27].

5. Conclusion and Implications

This study provides a qualitative account of how teachers and administrators in Shandong described AI-related changes in teaching work. The data do not support a claim that AI objectively increased productivity. A narrower claim is more defensible. Participants were more likely to describe their work as efficient when AI shortened routine drafting or first-pass screening without weakening explanation, safety, or the legitimacy of the assessment. The CMEO model is therefore not a recipe for productivity. It is an explanatory account of when AI-supported routines became workable in TVET and when they did not. At a practical level, the findings point to four recurring tasks: identify boundary constraints, redesign assessment and feedback around explainable routines, invest in enabling infrastructure and staff capacity, and keep teacher judgement in the loop.

At the institutional level, governance by design needs to be concrete. Schools can assign responsibility for AI-supported tasks, keep short records of use, and build review cycles that fit ISO/IEC 42001:2023. They can also identify risk points across the workflow in line with ISO/IEC 23894:2023 and use the NIST GenAI Profile as a working tool for procurement checks, incident disclosure, and record-keeping [28–30]. This is not an argument for heavy bureaucracy. It is an argument against treating governance as an afterthought. Because regulation is moving quickly, schools also need to watch legal changes, including the EU Artificial Intelligence Act, when they purchase tools or work with external suppliers [31].

5.1. Summary of Contributions

The main contribution of the study is explanatory. It shows how participants linked changes in teaching work to constraints, redesign practices, enabling conditions, and perceived outcomes. That is different from claiming that AI improves teaching in general. The contribution is smaller than that. It is also more defensible. The paper explains why some uses were experienced as worth the effort while others shifted work elsewhere or created new verification burdens. At the level of practice, participants described several routines as workable: explanation-first rubrics, evidence-chain diagnostics, human-controlled co-assessment, and bounded prompt use. These routines reduced friction only when teachers retained responsibility for fit, judgement, and risk.

5.2. Theoretical Implications

These findings extend current discussions of teaching efficiency by treating it as a bounded sociotechnical judgement. In the accounts analysed here, efficiency did not arise from speed alone. It depended on verification, explanation, traceability, and alignment with standards. This helps clarify a tension already visible in the literature. UNESCO's 2023 guidance argues that generative AI in education requires human oversight and context-sensitive governance [2]. Kasneci et al. make a related point when discussing inaccuracy, over-reliance, and the need for scrutiny [8]. Our study does not simply repeat those arguments in general form. It shows more specifically how those concerns are refracted through vocational education, where the cost of error is shaped by local equipment, assessment rules, and safety demands. That is the paper's main theoretical contribution: it recasts AI-enabled teaching efficiency not as a generic gain from tool adoption but as a conditional judgement embedded in task type, verification burden, and institutional safeguards.

The paper also treats efficiency as a value-laden judgement rather than a neutral technical outcome. Participants repeatedly described trade-offs among speed, fairness, personalisation, and professional credibility. Time saved in one stage of the workflow could be lost later through more checking or more explanation. Any account of AI-enabled efficiency that ignores those trade-offs will misread the setting.

5.3. Practical Implications for Vocational Schools

The practical implications are fairly plain. Schools should start with lower-risk, more formalised tasks where the cost of failure is low: lesson-plan scaffolds, worksheets, rubric drafting, post-formatting, and language polishing. Higher-risk uses, such as feedback generation for assessed work, competence mapping, or troubleshooting guidance, need a slower introduction and tighter checking. At the school level, a minimum support package includes an approved tool list, a controlled access route, short rules on safe input and storage, a template bank for recurring prompts and feedback formats, and fast peer support for sharing working examples and common failures. A summary of AI-related activities by teaching phase, together with expected outputs and associated governance checks, is provided in **Appendix C** (**Table A2**).

Professional development should be organised around workflows, not around tools in the abstract. Teachers in the study learned best when training stayed close to familiar vocational tasks: turning a workshop procedure into a micro-lesson, drafting feedback for recurring procedural errors, or checking whether AI-generated wording matched local practice. Training also needs an explicit checking stage. A prompt without a checking routine is incomplete. Evaluation should fit the character of TVET work. Asking only whether time was saved is too blunt. Schools should also look at feedback turnaround time, consistency of rubric use, student rework caused by unclear instructions, and safety-related near misses linked to procedural misunderstanding.

5.4. Ethical Governance and Equity

Another lesson from the study is that governance is not a compliance add-on. Where confidentiality, intellectual property, or safety rules were unclear, participants either avoided AI or used it in improvised ways that increased the checking burden. Practical guidance about data handling—what may be entered, how student work should be anonymised, and what outputs may be stored—made the use of data less uncertain. Equity concerns were just as concrete. AI can widen disparities if reliable access, templates, and support are available only to digitally confident teachers or well-resourced departments. Schools should not allow AI competence to remain a private asset of a few enthusiasts.

5.5. Transferability and Boundary Conditions

Although the study is based in Shandong Province, the mechanism-and-conditions model is offered as a context-bounded explanation rather than as a universal template. Transfer is more plausible in settings that share some of the same features: competency-based curricula, close ties to industry standards, and strong accountability for safety and assessment quality. Even then, caution is needed. Perceived efficiency is likely to vary where access is unstable, curricula are less codified, or error costs are lower.

For the same reason, schools should avoid one-size-fits-all AI policies. Governance needs to be segmented by task and subject. Language polishing or lesson structure may be appropriate for broad use. Technical procedures, troubleshooting advice, or safety content may need approved knowledge bases and tighter review. A risk-tier approach fits vocational education better than blanket permission or blanket prohibition.

5.6. A Practical "Minimum Viable Package" for Scaling

For schools that want to move beyond isolated experimentation, the data point to a small operational package that can be introduced within a semester. Step one is standardisation: a shared prompt-and-template library for recurrent vocational tasks, plus brief verification checklists. Step two is a light quality-assurance loop. A school can sample a small number of AI-assisted artefacts each month - one rubric, one set of feedback comments, one workshop instruction sheet - and review them against agreed criteria such as technical accuracy, local fit, clarity for students, and traceability of sources. This does not need to be punitive. It does need to happen.

Data governance should also be treated as part of basic implementation infrastructure. When teachers do not know what may be pasted into a tool, they either waste time reformatting material or avoid AI in the tasks where it might actually help. A simple data-classification scheme—public curriculum material, internal teaching material, de-identified student material, identifiable student data—reduces that friction and makes later audit easier. Monitoring should stay modest but relevant to TVET. Time saved is not enough. Schools also need indicators that quality has held: fewer procedural confusions, less rework, faster feedback loops, more consistent assessment, and steadier safety performance.

These implications are intentionally limited. The evidence does not suggest that every school needs a large AI strategy or a dedicated AI office. Early implementation is more plausible when institutions start with a narrow set of high-frequency tasks, define acceptable use with some precision, and build short feedback loops so that poor practice is spotted quickly. Large claims without operational detail do very little for teachers.

The same caution applies to professional development. One-off sessions on tool features are unlikely to change practice in a durable way. Teachers need examples tied to their own subject areas, discussion of likely failure points, and opportunities to compare where AI assistance is useful and where manual judgement must remain. That work is slower. It is also more defensible.

6. Limitations and Future Research

This study has several clear limits. First, the sample is small and bounded. Eight interviews can support focused qualitative explanation, but they cannot represent the full range of vocational subjects, institutional conditions, or AI practices in Shandong. Some constraints may therefore be under-represented, especially those tied to subject areas not included in the sample. The findings should be read as a bounded account of selected cases, not as a map of the full sector.

Second, the evidence base is narrow. The study combines teacher and administrator accounts and relies on interviews rather than observation, document analysis, platform logs, or student perspectives. That broadens the view of implementation, but it also leaves blind spots. Interview data can show how participants understood their work. They cannot verify how often reported routines occurred, whether they were carried out consistently, or how students experienced them. Some claims about workflow change therefore remain perception-based, and some profession-specific constraints may remain less visible than they would in a multi-source design.

This study has several clear limits. First, the sample is small and bounded. Eight interviews can support focused qualitative explanation, but they cannot represent the full range of vocational subjects, institutional conditions, or AI practices in Shandong. Some constraints may therefore be under-represented, especially those tied to subject areas not included in the sample. The findings should be read as a bounded account of selected cases, not as a map of the full sector.

Second, the evidence base is narrow. The study combines teacher and administrator accounts and relies on interviews rather than observation, document analysis, platform logs, or student perspectives. That broadens the view of implementation, but it also leaves blind spots. Interview data can show how participants understood their work. They cannot verify how often reported routines occurred, whether they were carried out consistently, or how students experienced them. Some claims about workflow change therefore remain perception-based, and some profession-specific constraints may remain less visible than they would in a multi-source design.

Author Contributions

Conceptualization: W.L. and G.Y.S.; Methodology: W.L.; Investigation, transcription, anonymisation, and preparation of the English working transcripts for analysis: W.L.; Formal analysis and theme development: W.L.; Writing—

original draft: W.L.; Writing—review and editing: W.L. and G.Y.S.; Syntactic modifications: S.L.; Supervision: G.Y.S. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Ethical approval was obtained from Teesside University (Approval Reference: 2025 Jun 30138 LIU). The study was conducted in accordance with Teesside University's research ethics policies and applicable data-protection requirements. Interview audio files and transcripts were stored on password-protected devices and access was limited to the research team. Identifiers were removed during transcription and participants were assigned pseudonyms (T01–T05; A01–A03).

Informed Consent Statement

Participants received a study information sheet detailing the purpose, procedures, confidentiality protections, and intended use of de-identified quotations. Written informed consent was obtained before data collection began, including consent for audio recording. Participation was voluntary. Participants were free to decline any question and could withdraw without penalty. The options for withdrawal and data removal were explained during the consent process.

Data Availability Statement

The full interview transcripts are not publicly available because they contain potentially identifiable institutional experiences and are covered by the conditions of the ethics approval. The interview guide, selected de-identified excerpts used to support the findings, and the analytic codebook may be made available by the corresponding author upon reasonable request. Any such request will be considered subject to an appropriate data-sharing agreement and a review of re-identification risk. Requests should be directed to the corresponding author.

Acknowledgments

The authors thank the vocational teachers and administrators who participated in the study.

Conflicts of Interest

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

AI Use Statement

During the preparation of this manuscript, the authors used ChatGPT solely for language refinement. No AI tools were used for data analysis, interpretation, or generation of scientific content. All outputs were critically reviewed and edited by the authors. The authors take full responsibility for the integrity and accuracy of the work.

Appendix A

Figure A1 shows where participants located AI-related tasks across the teaching cycle. It is not intended as a prescribed workflow. Rather, it summarises the points at which AI entered planning, delivery, assessment, and administrative work in the cases examined in this study.

Illustrative interview questions (selected examples): (1) Please describe a lesson or assessment task in which you used an AI tool. What problem were you trying to address, and how did it change your workflow? (2) Please recall a case in which AI-generated output was difficult to explain, unsafe, or impractical in the workshop context. How did you check or correct it? (3) What rules, technical controls, or training arrangements shaped what you could and could not do with AI, such as restrictions on data upload or teacher-in-the-loop requirements?

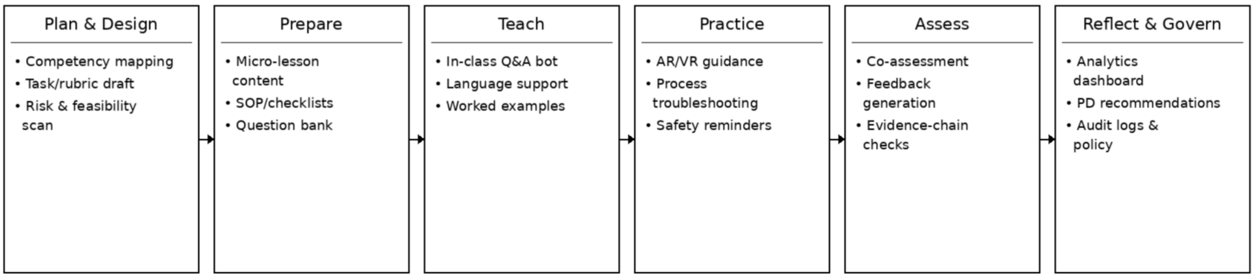

Figure A1. AI-enabled teaching activity map.

Note: The figure summarises the AI-related activities described by participants across the teaching cycle. It is provided as an organising aid rather than as a formal process model.

Appendix B

Table A1. Three-stage coding process (open, axial, and selective) with illustrative examples.

Coding Level	Analytic Focus	Illustrative Example from This Study
Open coding	Identify meaning units; label actions, constraints, and practices.	"students can use AI but cannot explain parameters" → use–explain gap (C2).
Axial coding	Relate codes to conditions, actions/interactions, and consequences.	Link 'cognitive overload' (C3) to 'information layering' mechanism (M4).
Selective coding	Integrate into a higher-order storyline and framework.	Synthesize into CMEO: constraints → mechanisms → enablers → perceived outcomes.

Appendix C

Table A2. AI-related activities, expected outputs, and associated governance checks.

Teaching Phase	AI Activity	Tool Category	Output Artifact	Human Check/Governance
Pre-class	Draft lesson outline and examples for differentiated instruction	AI content generation/NLP	Lesson plan draft; Examples; Prompts	Teacher reviews for factual accuracy and alignment with curriculum outcomes.
Pre-class	Generate task sheets and safety checklists for workshop sessions	AI content generation	Task cards; Checklists	Instructor validates safety-critical steps against institutional/industry SOPs.
In-class	Real-time Q&A and language support during demonstrations	NLP tools	Explanations; Translated terms	Teacher mediates; students must provide 'why' justification where required.
In-class	Adaptive practice prompts for common misconceptions	Intelligent tutoring/ITS	Hint sequences; Targeted questions	The teacher monitors for over-scaffolding and adjusts difficulty.
Practice	Simulated practice in virtual/augmented environments	VR/AR	Simulation scenarios; Performance traces	Instructor confirms fidelity and links actions to workplace constraints.
Assessment	First-pass feedback on assignments and reports	Automated assessment & feedback	Feedback draft; Rubric suggestions	No AI-only scoring; teacher finalizes grades and comments.
Assessment	Evidence-chain diagnostics for troubleshooting tasks	Learning analytics/diagnostic support	Required evidence set (e.g., waveform, trend, test result)	The teacher verifies evidence completeness and coaches reasoning.
Reflection	Risk monitoring and governance logging (e.g., sensitive data, hallucination)	Governance-by-design	Usage logs; Flagged cases	The institution enforces policy; teachers use flagged cases to refine tasks.

References

1. Selwyn,N. *Distrusting Educational Technology: Critical Questions for Changing Times*; Routledge: Abingdon, UK, 2014.
2. UNESCO. *Guidance for Generative AI in Education and Research*; UNESCO: Paris, France, 2023. Available online: https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research
3. Holmes,W.; Bialik,M.; Fadel,C. *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*; Center for Curriculum Redesign: Boston, MA, USA, 2019.
4. Luckin,R.; Holmes,W.; Griffiths,M.; et al. *Intelligence Unleashed: An Argument for AI in Education*; Pearson: London, UK, 2016.
5. Øgård,M. Boundary objects as a starting point for reflective learning in vocational education and training classrooms. *Nord. J. Vocat. Educ. Train.* **2024**, *14*, 424–442. [CrossRef]
6. Mårtensson,Å. Creating continuity between school and workplace: VET teachers' in-school work to overcome boundaries. *J. Vocat. Educ. Train.* **2022**, *74*, 682–700. [CrossRef]
7. Yusop,S.R.M.; Rasul,M.S.; MohammadYasin,R.; et al. Identifying and Validating Vocational Skills Domains and Indicators in Classroom Assessment Practices in TVET. *Sustainability* **2023**, *15*, 5195. [CrossRef]
8. Kasneci,E.; Sessler,K.; Küchemann,S.; et al. ChatGPT for good? On opportunities and challenges of large language models for education. *Learn. Individ. Differ.* **2023**, *103*, 102274. [CrossRef]
9. Ogunleye,B.; Zakariyyah,K.I.; Ajao,O.; et al. A Systematic Review of Generative AI for Teaching and Learning Practice. *Educ. Sci.* **2024**, *14*, 636. [CrossRef]
10. Wang,X.; Zainuddin,Z.; Hain,H.L. Generative artificial intelligence in pedagogical practices: A systematic review of empirical studies (2022–2024). *Cogent Educ.* **2025**, *12*, 2485499. [CrossRef]
11. European Commission: Directorate-General for Education, Youth, Sport and Culture. *Ethical Guidelines on the Use of Artificial Intelligence (AI) and Data in Teaching and Learning for Educators*; Publications Office of the European Union: Luxembourg, Luxembourg, 2022. [CrossRef]
12. Shute,V.J. Focus on formative feedback. *Rev. Educ. Res.* **2008**, *78*, 153–189. [CrossRef]
13. Radianti,J.; Majchrzak,T.A.; Fromm,J.; et al. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Comput. Educ.* **2020**, *147*, 103778. [CrossRef]
14. Long,P.; Siemens,G. Penetrating the fog: Analytics in learning and education. *EDUCAUSE Rev.* **2011**, *46*, 31–40. Available online: https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education
15. Creswell,J.W.; Poth,C.N. *Qualitative Inquiry and Research Design: Choosing among Five Approaches*, 4th ed.; Sage: Thousand Oaks, CA, USA, **2018**.
16. Tong,A.; Sainsbury,P.; Craig,J. Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *Int. J. Nurs. Stud.* **2007**, *19*, 349–357.
17. O'Brien,B.C.; Harris,I.B.; Beckman,T.J.; et al. Standards for reporting qualitative research: A synthesis of recommendations. *Acad. Med.* **2014**, *89*, 1245–1251. [CrossRef]
18. Guest,G.; Bunce,A.; Johnson,L. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* **2006**, *18*, 59–82. [CrossRef]
19. Malterud,K.; Siersma,V.D.; Guassora,A.D. Sample size in qualitative interview studies: Guided by information power. *Qual. Health Res.* **2016**, *26*, 1753–1760. [CrossRef]
20. Strauss,A.; Corbin,J. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 2nd ed.; Sage: Thousand Oaks, CA, USA, 1998.
21. Braun,V.; Clarke,V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
22. Hsieh,H.-F.; Shannon,S.E. Three approaches to qualitative content analysis. *Qual. Health Res.* **2005**, *15*, 1277–1288. [CrossRef]
23. Miao,F.; Cukurova,M. *AI Competency Framework for Teachers*; UNESCO: Paris, France, 2024. Available online: https://www.unesco.org/en/articles/ai-competency-framework-teachers
24. Schmitt,C.; Brutzer,A. Generative artificial intelligence in vocational education and training: A framework for sustainable teacher competence development. *Vocat. Technol. Educ.* **2025**. [CrossRef]
25. Garzón,J.; Patiño,E.; Marulanda,C. Systematic Review of Artificial Intelligence in Education: Trends, Benefits, and Challenges. *Multimodal Technol. Interact.* **2025**, *9*, 84. [CrossRef]
26. Papagiannidis,E.; Mikalef,P.; Conboy,K. Responsible artificial intelligence governance: A review and research framework. *J. Strateg. Inf. Syst.* **2025**, *34*, 101885. [CrossRef]
27. Alfiras,M.I.I.; Emran,A.Q.; Mohamed,A.M. Ethics and governance of generative AI in education: A systematic

review on responsible adoption. *Discover Educ.* **2025**, *5*, 37. [CrossRef]
28. ISO/IEC. *Information Technology—Artificial Intelligence—Management System (ISO/IEC 42001:2023)*; International Organization for Standardization: Geneva, Switzerland, 2023. Available online: https://www.iso.org/standard/81230.html
29. ISO/IEC. *Information Technology—Artificial Intelligence—Guidance on Risk Management (ISO/IEC 23894:2023)*; International Organization for Standardization: Geneva, Switzerland, 2023. Available online: https://www.iso.org/standard/77304.html
30. NIST. *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1)*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2024. [CrossRef]
31. European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). Available online: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (accessed on 27 February 2026).