

## **GRADUATE SUPERVISOR DEVELOPMENT AS QUALITY IMPROVEMENT IN CHINESE PRIVATE HIGHER EDUCATION: A MIXED METHODS EVALUATION OF A FOUR-IN-ONE MODEL**

KAI YANG, XIAOFANG DONG\*, WEI WU, JINGCHAO LIU,  
LIANGCHEN REN, RUIXUAN ZHANG, JIARUI WANG, PENG MEI

School of Computer Science, Xijing University, Xi'an 710123, China

\*Corresponding Author: iifdong509@gmail.com

### **Abstract**

This study develops and evaluates a formative evaluation framework for graduate supervisor development in Chinese private higher education. The intervention evaluated in this study is a four-in-one model comprising differentiated selection, dual-path capability development, multidimensional process appraisal, and diagnostic feedback use. Using a mixed methods case study of an academic school at a Chinese private university, the study integrated item-level questionnaire evidence from supervisors, graduate students, and graduate programme administrators with semi-structured interviews, open-ended responses, and institutional documents through methodological triangulation. Baseline diagnosis identified four problems: weak differentiation in supervisor selection, limited capacity to supervise practice-based projects, unstable university-industry platforms, and appraisal criteria dominated by research outputs. First-round evidence indicated changes in supervisor composition, student-perceived supervision quality, and university-industry collaboration. These are perception, implementation, and organisational indicators rather than direct measures of graduate learning or supervisory behaviour. Because the evidence comes from a single case and an early implementation cycle, the model is interpreted as a context-bound quality-improvement framework rather than causal impact evidence.

Keywords: Graduate supervision, Higher education, Postgraduate education, Professionally oriented postgraduate education, Supervisor development.

## 1. Introduction

Graduate supervisors have a direct influence on the process and quality of graduate education. They undertake multiple responsibilities, including academic guidance, research training, career development support, academic integrity education, and management of the graduate education process. Existing research shows that effective supervision depends not only on academic qualifications but also on supervisors' pedagogical capability, feedback practices, relationship building, and responsiveness to students' development needs [1-4]. As graduate education shifts from scale expansion toward quality enhancement, supervisor development has become an important concern for internal quality assurance and educational evaluation.

In Chinese private higher education, this issue has a distinctive institutional context. The rapid development of professional-degree graduate education requires supervisors to possess engineering practice experience, capacity to respond to technological change, capability to guide industry projects, and ability to collaborate with enterprises. National policy has also emphasised the need to improve graduate education quality and strengthen supervisor teams [5]. Although previous Chinese studies have examined supervisor team development, competency structures for professional-degree supervisors, and supervisor responsibilities [6-8], empirical research remains limited on how private universities diagnose, implement, and evaluate supervisor development models.

The private university sector provides a particularly demanding site for examining this problem. In many such institutions, graduate education has expanded faster than the accumulation of supervisory experience, research infrastructure, and stable external partnerships. Supervisor development is therefore not a narrowly personnel-related issue, but part of a broader effort to build institutional capacity under conditions of growth. A system that relies mainly on conventional academic credentials may secure formal legitimacy, yet it may be less effective in preparing supervisors for practice-oriented degree programmes. Conversely, a system driven only by immediate industry demands may weaken academic coherence. The central challenge is to align academic standards, applied relevance, and evidence-informed improvement within one supervisory governance framework.

International scholarship provides a basis for understanding supervisor roles, supervisor development, and supervisory capability. It has focused on supervisors' professional learning, training for modern doctoral education, and power and interaction in supervisory relationships [9-13]. From an educational evaluation perspective, supervisor development involves problem diagnosis, personnel development, process monitoring, result feedback, and institutional improvement. Utilization-focused evaluation emphasizes that evaluation should respond to the improvement needs of specific users [14], while the Context, Input, Process, and Product (CIPP) model highlights the need to consider context, input, process, and product [15, 16]. Studies of educational development further suggest that evaluation should not stop at summative judgement but should examine development goals, process evidence, feedback use, and continuous improvement [17-20]. Research on boundary-crossing collaboration further indicates that teacher education and professional development often take place in collaborative fields involving universities, schools, practice institutions, and multiple stakeholders.

Evaluation in such contexts should therefore attend to knowledge integration, shared goals, trust relationships, and organisational support [21].

Although international scholarship has generated a mature vocabulary for analysing supervision, its assumptions cannot be transferred uncritically to Chinese private higher education. Much of that literature is grounded in settings where doctoral cultures, supervisory norms, and quality assurance routines are already comparatively stable. The present case differs in that supervisory capacity, programme differentiation, industry collaboration, and internal evaluation are still being consolidated. Under these conditions, the relevant question is not simply whether supervisors receive training, but whether the institution can link appointment, support, appraisal, and feedback in ways that correspond to its own developmental stage and programme mission.

This study evaluates a four-in-one graduate supervisor development model in an academic school at a Chinese private university. Professionally oriented postgraduate programmes in technology-related disciplines depend especially on engineering practice, rapid technological updating, and industry-project guidance, making this setting useful for examining common challenges in supervisor development in private higher education. The study addresses three research questions: What major difficulties shape graduate supervisor development in this context? How does the four-in-one model respond to these difficulties? What first-round formative evidence can be observed after implementation? The contribution of the study is to reframe supervisor team development from an experiential reform topic into a developmental educational evaluation problem and to show how diagnostic findings, model design, and formative feedback can be connected. Unlike previous studies that often examine supervisor training, competency, or appraisal separately, this study connects these elements within a continuous improvement cycle.

The study is positioned against a practical limitation in existing reform accounts. Many institutional reports document new training arrangements or revised appraisal rules, but they rarely show how diagnostic evidence is converted into a coherent improvement model, or how early implementation evidence is used to refine that model. This article therefore treats evaluation not as an external judgement made after a reform has concluded, but as a constitutive part of the reform process. Baseline diagnosis defines the problem structure, model design translates that diagnosis into institutional arrangements, and formative monitoring examines whether those arrangements are beginning to work as intended.

The remainder of the article is organized as follows. Section 2 develops the analytical framework of the research on supervision, professional-degree education, and educational evaluation. Section 3 explains the case context, data sources, and analytical methods. Section 4 presents the diagnostic evaluation, the model design, and first-round formative evidence. Section 5 discusses the findings, limitations, and directions for future research.

## 2. Literature Review and Evaluation Framework

Supervision quality is multidimensional. Lee [2] conceptualized doctoral supervision through functions such as academic guidance, induction into academic culture, development of critical thinking, support for student autonomy, and construction of supervisory relationships. Pearson and Brew [4] located supervisor

development within research training systems and emphasised institutional arrangements and continuing professional learning. Halse [10] argued that becoming a supervisor is itself a process of professional learning, and Kiley [11] showed that supervisor training is linked to changes in doctoral education, accountability, and quality assurance. These studies indicate that supervisor quality cannot be judged only by static indicators such as degrees, academic titles, publications, or projects. It also requires attention to guidance capability, feedback capability, and developmental needs in the graduate education process. Evaluation studies of academic and teacher development similarly show that programme value should be judged through the relationship among development goals, implementation processes, evidence collection, and impact analysis, rather than through a single outcome indicator [17, 18, 20].

In this study, a graduate supervisor refers to an academic staff member who guides and oversees students undertaking master's or doctoral study. The case focuses on professionally oriented postgraduate programmes, in which supervisors are expected to maintain academic standards while also supporting practice-based learning, industry projects, and applied problem solving. Supervision quality in this context is therefore broader than the transmission of disciplinary knowledge alone. Supervisors may need to help students connect conceptual analysis with project constraints, professional judgement, and applied outputs. These responsibilities are only partly reflected in traditional indicators such as publications, grants, academic titles, or research projects. A highly productive researcher may still require support in industry-based project supervision, while an experienced practitioner may require support in thesis standards, research ethics, and academic mentoring. Supervisor evaluation should therefore recognise heterogeneous capability profiles rather than assuming that a single indicator set can represent all forms of supervisory quality.

Supervisory relationships and student perceptions provide important evidence for supervisor evaluation. Mainhard et al. [3] developed a questionnaire for doctoral students' perceptions of supervisor-student interaction and treated student perception as feedback for supervisory improvement. Vähämäki et al. [13] analysed doctoral supervision through leader-member relationships and power dynamics, revealing how critical interaction events, trust, and institutional context influence supervisory experience. Student perception can therefore be used as formative evidence of supervision quality, but it should not be reduced to a single summative judgement. Research on formative assessment also suggests that shared teacher and student perceptions of evaluation practice can help explain how evidence is used to improve teaching and guidance processes [22].

Professionally oriented postgraduate programmes place greater demands on practice-oriented supervision and boundary-crossing collaboration. Compared with academically oriented postgraduate programmes, they give greater priority to authentic problem solving, industry participation, project-based learning, and applied innovation. Supervisors in these programmes require academic knowledge, practical experience, industry understanding, project-guidance capability, and the capacity to collaborate with enterprise partners. In private universities, supervisor teams may include a relatively high proportion of early-career faculty, while research platforms and cooperation mechanisms are still developing. Supervisor development should therefore connect appointment, capability development, process appraisal, and feedback use rather than address isolated weaknesses.

Research on boundary-crossing collaboration further suggests that collaboration among practice institutions, research actors, and learners should be evaluated through knowledge integration, shared goals, trust, and organisational support [21].

University-enterprise collaboration is relevant to the present framework because it provides the practical settings in which supervisors can develop and demonstrate practice-oriented guidance capability. Collaboration may generate authentic project contexts, joint mentoring opportunities, access to professional problems, and feedback from practitioners. However, formal agreements alone do not guarantee sustained educational value. Collaboration may be constrained by different time horizons, uneven expectations, intellectual-property concerns, and discontinuity in project resources. These constraints help explain why practice-oriented supervision cannot be strengthened through individual motivation alone. It requires institutional platforms that make external cooperation usable for supervisor development, student project guidance, and process appraisal.

This study adopts a developmental educational evaluation framework. Utilization-focused evaluation argues that evaluation design should address specific users and improvement contexts [14]. The CIPP model requires evaluation to examine context, input, process, and product [15, 16]. Research on school self-evaluation and school improvement indicates that evaluation can support organisational development only when it goes beyond outcome judgement and attends to evaluation processes, evidence use, and enabling conditions for improvement [23, 24]. Studies of educational development and teacher evaluation also indicate that evaluation should promote professional growth, instructional improvement, and organisational learning, rather than function only as accountability or ranking [17, 19, 25-27]. Classroom observation and teaching quality evaluation further suggest that personnel evaluation should integrate general dimensions with context-specific practice evidence [28]. These studies justify the inclusion of supervision process records, student feedback, and university-industry collaboration in evaluating supervisor responsibilities.

Viewed through a developmental evaluation lens, the four components represent four linked institutional decisions. Differentiated selection concerns whether supervisors' qualifications and experience are appropriately matched to programme goals. Capability development concerns whether identified gaps are addressed through relevant support. Multidimensional appraisal concerns whether the institution collects evidence on supervision processes, student development, professional conduct, and practical contribution rather than relying only on terminal outputs. Diagnostic feedback concerns whether appraisal findings inform subsequent decisions about training, resources, incentives, recruitment quotas, and qualification review. Together, these components structure the analysis of problem diagnosis, model design, and early formative evidence.

The model also responds to two weaknesses often found in supervisory appraisal systems. The first is indicator proliferation, where additional criteria are added without a clear improvement logic. The second is weak use of results, where appraisal data are recorded but have limited consequences for support, resource allocation, or qualification review. The four-in-one model seeks to connect each evaluative component with a specific institutional function: selection with appointment, development with capability building, appraisal with process

visibility, and feedback with iterative adjustment. Its contribution lies less in any single component than in the linkage among them.

### 3. Methods

#### 3.1. Research design and case context

The study used a mixed methods case study design [29, 30]. The case institution is a Chinese private university, and the focal unit is an academic school offering technology-oriented postgraduate programmes with a strong professional and practice-oriented focus. To comply with double-anonymised review, the institution is referred to as the case university. In these programmes, supervisors are expected to support academic inquiry, engineering practice, technological updating, and industry-related project guidance. This setting is suitable for examining how supervisor development can be organized when academic standards and practice-oriented expectations must be addressed simultaneously.

A mixed methods case study was appropriate because the research problem concerns both patterns of perception and the institutional mechanisms that generate them. Questionnaire data were used to identify the relative salience of problems perceived by supervisors, students, administrators, and external partners. Interview and document evidence were then used to interpret those patterns in relation to actual rules, practices, and resource conditions. For instance, student concerns about project guidance can be understood more adequately when read alongside supervisors' career trajectories, industry-collaboration records, and appraisal arrangements. The design was intended to explain questionnaire patterns within their institutional context.

The study followed an explanatory mixed methods case-study logic. The quantitative strand consisted of item-level questionnaire summaries that described the distribution of reported concerns among supervisors, students, and teaching administrators. The qualitative strand consisted of semi-structured interviews, open-ended questionnaire responses, and institutional documents. Integration occurred during interpretation: questionnaire findings were compared with interview accounts and documentary evidence for each diagnostic theme. For example, reported concerns about practical guidance were interpreted alongside supervisors' prior project experience, industry collaboration records, and appraisal arrangements. The purpose of this integration was to explain how reported concerns related to institutional conditions and implementation practices, rather than to make causal inferences from the questionnaire data alone.

Baseline diagnostic data were collected from March to May 2024. On the basis of this diagnosis, the focal school began an institutional pilot of the four-in-one model in June 2024. The pilot was subsequently incorporated into a formally approved educational reform project running from December 2024 to December 2026. This study reports the baseline diagnosis and early formative evidence from the initial implementation period; it is therefore an evaluation of model design and early implementation, not a causal impact evaluation.

The timing of the study shaped the scope of its claims. Since the reform project had not completed a full cycle, the analysis does not claim evidence of durable improvement in thesis quality, completion, employment outcomes, or long-term research productivity. Instead, it examines the coherence of the

diagnostic logic, the alignment between identified problems and model design, and the plausibility of early implementation signals. This distinction is methodologically important. Formative evaluation is not a preliminary substitute for outcome evaluation; it has a specific function in supporting learning and adjustment during implementation.

### 3.2. Data Sources and Participants

The study integrated three types of evidence: questionnaires, semi-structured interviews, and institutional document analysis. Questionnaire respondents included 45 graduate supervisors, 120 currently enrolled graduate students, and 8 teaching administrators. The supervisor group included supervisors from academically oriented and professionally oriented postgraduate programmes, while the student group covered different years of postgraduate study. The administrators worked in graduate education and school-level teaching management roles. Semi-structured interviews were conducted with 10 graduate supervisors and 5 industry collaboration leaders. These interviews were used to explain questionnaire patterns and to examine how supervision, enterprise collaboration, and appraisal arrangements operated in practice. Institutional texts included supervisor selection documents, training arrangements, university-industry cooperation materials, appraisal policies, and project implementation records. Table 1 summarises the participants, materials, and evaluation purposes.

**Table 1. Data sources and evaluation purposes.**

| <b>Data source</b>                | <b>Participants or materials</b>              | <b>Quantity</b>    | <b>Evaluation purpose</b>                                                                                      |
|-----------------------------------|-----------------------------------------------|--------------------|----------------------------------------------------------------------------------------------------------------|
| <b>Questionnaire</b>              | Graduate supervisors                          | 45                 | To understand supervisors' perceptions of selection, development, appraisal, and platform support              |
| <b>Questionnaire</b>              | Currently enrolled graduate students          | 120                | To understand students' evaluations of guidance frequency, feedback quality, and practice-oriented supervision |
| <b>Questionnaire</b>              | Graduate programme administrators             | 8                  | To understand implementation of supervisor management policies and operational barriers                        |
| <b>Semi-structured interviews</b> | Graduate supervisors                          | 10                 | To explain questionnaire findings and obtain supervisors' experiences of guidance and appraisal                |
| <b>Semi-structured interviews</b> | Industry-partnership managers                 | 5                  | To understand university-industry collaboration, industry mentor participation, and practical project support  |
| <b>Document analysis</b>          | Institutional documents and project materials | Multiple documents | To analyse supervisor selection, training, appraisal, and feedback system design                               |

### 3.3. Instruments and analysis

The questionnaire was developed internally as a formative diagnostic instrument rather than as a standardised psychometric scale. Separate questionnaires were administered to supervisors and students, while teaching administrators completed items concerning the implementation of supervisor-management policies. The questionnaires combined five-point Likert-type items with open-ended questions. Their content covered supervisor selection, professional development, supervision processes, university-industry support, appraisal, policy implementation, and improvement needs. The instrument was developed to support the institutional reform process and was reviewed by the project team for its relevance to the reform objectives. The quantitative findings reported in this article are item-level descriptive results. Unless otherwise stated, percentages represent the proportion of valid respondents selecting “agree” or “strongly agree” for a given item. Because the questionnaire included several heterogeneous diagnostic domains and was not treated as a single latent scale, no overall internal-consistency coefficient is reported. The questionnaire findings are therefore interpreted as formative diagnostic evidence and are considered alongside interview and documentary evidence rather than as validated scale scores.

The questionnaire content was organized around the four model components. Selection content concerned programme differentiation, practice background, topic quality, and supervisory responsibility. Development content concerned training needs, enterprise practice, and collaborative mentoring. Appraisal content concerned supervision processes, student development, professional ethics, and practical contribution. Feedback-use content concerned the use of evaluation results in training, resources, recruitment quotas, and qualification review. Open-ended responses provided contextual evidence on difficulties not fully captured by the structured items. Because the instrument combined heterogeneous diagnostic domains and was not analysed as a unidimensional or latent scale, an overall internal-consistency coefficient would not provide an appropriate summary of reliability. No claims of construct validity or scale-score reliability are made. Its evidentiary role is limited to item-level diagnosis and interpretation alongside interview and documentary evidence. Table 2 therefore presents translated illustrations of the content addressed rather than a complete standardised scale or verified verbatim questionnaire.

**Table 2. Questionnaire dimensions and illustrative item content.**

| <b>Dimension</b>              | <b>Illustrative item content</b>                                                                                |
|-------------------------------|-----------------------------------------------------------------------------------------------------------------|
| <b>Supervisor selection</b>   | The supervisor selection criteria reflect the characteristics of different graduate programmes.                 |
| <b>Capability development</b> | Training activities improve supervisors' ability to guide practice-oriented graduate projects.                  |
| <b>Appraisal mechanism</b>    | The appraisal system considers supervision processes and student development rather than only research outputs. |
| <b>Feedback use</b>           | Evaluation results are used to improve supervisor development and resource allocation.                          |

Semi-structured interviews focused on the alignment between supervisor selection and programme type, practice capacity among early-career supervisors, stability of university-industry collaboration, effectiveness of appraisal, and use of evaluation results. The interview materials were reviewed and coded to identify recurring issues. Related codes were subsequently grouped into four themes concerning supervisor selection, practice-oriented capability, university-industry collaboration, and appraisal. This process was informed by thematic-analysis principles [31], but the available project records do not support a claim of independent double coding or intercoder-reliability testing. Open-ended questionnaire responses were treated as qualitative questionnaire evidence rather than as student interviews. Institutional documents were examined through content analysis [32], with attention to appointment requirements, training arrangements, appraisal indicators, use of results, and periodic adjustment mechanisms.

The qualitative analysis paid particular attention to the standpoint of different actors. Supervisors described workload, resource conditions, and ambiguity in expectations. Students' open-ended questionnaire responses emphasised feedback timeliness, guidance frequency, and support for practical projects. Administrators focused on policy feasibility and the balance between accountability and development. Enterprise partners were more concerned with project relevance, communication costs, and continuity. Rather than treating these accounts as interchangeable, the analysis used differences in standpoint to locate coordination problems within the supervisory development system.

### **3.4. Data collection, ethics, and evaluation logic**

Baseline diagnostic data were collected from March to May 2024 and used for needs analysis before the institutional pilot began in June 2024. First-round formative evidence came from early implementation records, including supervisor selection records, training and enterprise practice records, student satisfaction monitoring, university-industry cooperation platform materials, and revised policy documents. Questionnaire data were analysed through descriptive statistics, including item frequencies, proportions, and overall distributions. Interview and document findings were used to interpret the common problems shown in the questionnaires and to compare them with institutional arrangements for supervisor appointment, training, appraisal, and feedback.

The study was reviewed and approved retrospectively by the Xijing University Biomedical Research Ethics Committee (Reference No. 2026-015) on 28 September 2026. The ethical review and approval took place after the relevant data collection and initial implementation activities had commenced; it was therefore a retrospective ethical review and approval; not prospective approval obtained before data collection. Participation was voluntary, and responses were anonymised before analysis. No information identifying individual participants is reported. Relevant materials were used only for this research analysis and for improving supervisor development. The conclusions are contextual because the study is based on a single school case. The first-round outcomes are formative monitoring evidence and do not include a comparison group, independent external audit, or long-term student outcome data. They therefore cannot support direct causal claims about the model's impact.

Trustworthiness was supported through source triangulation and interpretive restraint. Questionnaire percentages showed the distribution of perceived problems or

early changes; open-ended responses and interviews helped explain how participants experienced those problems; and institutional documents and implementation records indicated whether relevant rules, activities, or organisational arrangements were present. Findings were interpreted more cautiously when they rested primarily on self-report or internal monitoring. Identifiable information about individuals, departments, and partner enterprises was excluded, consistent with the improvement-oriented purpose of the project and blind-review requirements.

The evaluation logic is shown in Fig. 1. Evidence from questionnaires, open-ended responses, interviews, and institutional documents supported the baseline diagnosis of problems in selection criteria, supervisor capability, platform support, and appraisal mechanisms. The diagnostic findings then informed the four linked components of the model: differentiated selection, dual-path capability development, multidimensional process appraisal, and diagnostic feedback use. University-enterprise-research collaboration functioned as a supporting mechanism within capability development and appraisal rather than as a separate fifth component. Early implementation records and first-round formative evidence were subsequently integrated to examine initial operation and support adjustment. In line with formative evaluation research, these findings are interpreted as improvement-oriented evidence rather than as evidence of long-term causal impact [17, 18, 20].

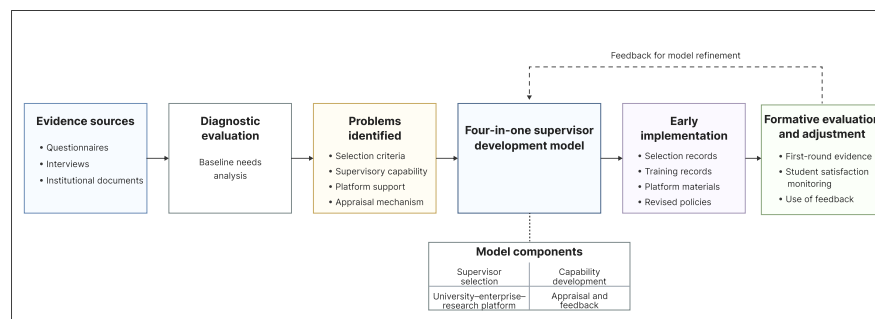


**Fig. 1. Evaluation logic of the four-in-one supervisor development model.**

## 4. Findings

### 4.1. Baseline diagnostic evaluation

The diagnostic evaluation identified four main problems in graduate supervisor development at the focal school.

The diagnostic findings are presented below as four related areas of concern: supervisor selection, practice-oriented guidance capability, university-industry support, and appraisal mechanisms. The questionnaire findings describe stakeholder perceptions, while interview and documentary evidence provide contextual explanations. The results do not establish a causal sequence among these concerns; possible relationships are considered later in the Discussion section.

First, supervisor selection standards were not consistently aligned with differentiated programme goals. Supervisors and graduate programme administrators reported that the existing selection system did not clearly distinguish the capability profiles required for academically oriented and professionally oriented postgraduate programmes. The process still emphasised academic

indicators such as publications, competitive research grants, and academic titles, while giving less attention to industry-project experience, engineering practice capability, applied project work, and practical problem-solving capacity. This created a risk of mismatch between supervisors' experience and the requirements of professionally oriented postgraduate education.

The issue of selection illustrates a tension between two forms of legitimacy. Academic credentials, publications, and research projects are visible, comparable, and administratively convenient. They are also important safeguards for graduate education quality. However, professional-degree supervision requires additional evidence of applied research capacity, project experience, and the ability to guide students in authentic professional contexts. The diagnostic results suggest that the existing system had not sufficiently balanced these requirements. Differentiated selection is therefore best understood as a refinement of academic standards, not as a relaxation of them.

Second, capacity to supervise practice-based projects was uneven. Early-career faculty accounted for 28 of the 45 supervisors (62.2%). Many early-career supervisors reported limited experience of large industry projects, and 81 of 120 graduate students (67.5%) reported insufficient guidance in engineering project management, technology selection, and system-architecture design. Interview records similarly indicated that some supervisors had strong research backgrounds but limited experience in industry-project supervision, engineering management, and practice-based problem solving.

The guidance-capacity problem should be interpreted developmentally. Many supervisors, especially those at earlier career stages, had been socialised primarily into disciplinary research, publication practices, and academic evaluation. Professional-degree supervision required additional competencies in project diagnosis, stakeholder communication, applied problem framing, and evaluation of practice-based outputs. These competencies are difficult to cultivate through formal lectures alone. The findings therefore support a more situated approach to supervisor development, in which supervisors learn through engagement with industry projects and collaborative mentoring arrangements.

Third, university-industry platform support was unstable. Although the school had established collaboration with several enterprises, activity was concentrated in student internships, short-term project matching, and general exchanges. Stable joint supervision arrangements had not yet been fully formed. Supervisors and interview participants described the absence of a sustained mechanism for collaboration. One industry representative explained: "The pace of university research does not always match the rapid iteration of industry, and many proposed research directions are not sufficiently advanced to encourage our deeper participation." This account indicates that formal collaboration does not necessarily ensure sustained involvement in authentic research and development tasks. In technology-related professionally oriented postgraduate education, weak collaboration platforms limit practical resources and weaken the connection between research-proposal development and authentic industry problems. Research on boundary-crossing evaluation likewise shows that collaboration quality depends on shared goals, knowledge integration, trust, and organisational support [21].

Fourth, supervisor appraisal overemphasised research outputs. Forty-one of 45 supervisors (91.1%) reported that existing appraisal indicators placed excessive weight on research funding and publications, while giving insufficient attention to

the supervision process, graduate student development, professional ethics and conduct, practical contribution, and collaborative education outcomes. Appraisal results were only weakly connected to the number of postgraduate places allocated to each supervisor, performance rewards, resource allocation, and review of supervisor qualifications. As a result, the link between diagnostic feedback and subsequent improvement was underdeveloped. Teacher-evaluation research similarly warns that excessive reliance on quantifiable results can weaken the developmental and instructional-improvement functions of evaluation [25-27]. Table 3 summarises the diagnostic findings and their implications for model design.

**Table 3. Main diagnostic problems before reform.**

| <b>Evaluation dimension</b>  | <b>Main evidence</b>                                                                                                                                                                          | <b>Problem interpretation</b>                                                                                                        | <b>Direction for model design</b>                                                                        |
|------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|
| <b>Supervisor selection</b>  | Supervisors and graduate programme administrators reported weak differentiation in selection criteria                                                                                         | Supervisor appointment standards were insufficiently aligned with differentiated programme goals                                     | Establish differentiated selection and prior research-proposal review mechanisms                         |
| <b>Supervisor capability</b> | 28 of 45 supervisors were early-career faculty; 81 of 120 students reported insufficient practice-based project guidance                                                                      | Capacity to supervise practice-based projects did not fully meet the requirements of professionally oriented postgraduate programmes | Strengthen industry-based practice, industrial workshops, and project-based training                     |
| <b>Platform support</b>      | Supervisors and interview participants reported that university-industry collaboration lacked a sustained mechanism; some proposed research directions were distant from enterprise R&D tasks | University-industry support was insufficient, and research-proposal development was not closely connected to industry needs          | Build stable joint supervision platforms and increase industry participation in research-proposal review |
| <b>Appraisal</b>             | 41 of 45 supervisors reported that appraisal overemphasised funding and publications                                                                                                          | The evaluation system gave insufficient attention to supervision processes and developmental feedback                                | Build a diversified, process-oriented, developmental appraisal mechanism                                 |

#### **4.2. Formation of the four-in-one supervisor development model**

Based on the diagnostic results, the focal school constructed a four-in-one supervisor development model. The model connects problem diagnosis, reform measures, and the use of evaluation results. Selection responds to the mismatch between supervisor appointment standards and programme type. Development responds to insufficient practice-oriented guidance capacity among early-career supervisors. Appraisal responds to the overemphasis on research outputs. Feedback use responds to the weak

use of evaluation results. This logic is consistent with educational development evaluation research, which emphasizes connecting development goals, implementation processes, evidence collection, and improvement feedback rather than limiting evaluation to summative judgement [17, 18, 20].

The four components are interdependent. Appraisal findings inform feedback, diagnosed needs guide development activities, and post-appointment support complements differentiated selection. Selection establishes the initial fit between a supervisor's profile and programme goals. Development addresses identified capability gaps after appointment. Appraisal collects evidence on how supervisory responsibilities are enacted during graduate education, and feedback connects that evidence to training, resources, and qualification decisions.

In the selection component, the focal school adjusted supervisor appointment standards through differentiated selection and prior topic review. Academic-degree supervisors were evaluated mainly on academic innovation capability, research foundation, and capacity for theoretical research guidance. Professional-degree supervisors were evaluated mainly on industry project experience, engineering practice capability, industry-funded project foundation, patent commercialisation, industry influence, and practical problem-solving capacity. Topic directions were reviewed before supervisor appointment by academic committee members and industry experts, with a focus on research foundation, feasibility, application value, and alignment with programme type.

Prior topic review played a strategic role because professional-degree supervision begins before the thesis process is formally visible. A topic that lacks a clear practical problem source, technical feasibility, or project support can undermine later supervision regardless of the supervisor's effort. The review mechanism therefore examined the origin of the problem, the technical foundation, implementation conditions, and expected applied outcomes. Academic committee members judged scholarly feasibility, while industry experts assessed practical relevance. This arrangement helped prevent topics from becoming either purely academic exercises or loosely defined technical tasks with limited research value.

In the development component, the focal school responded to practice-oriented guidance gaps through two pathways: internal supervisor development and collaborative support from industry mentors. Measures included industrial technology workshops, industry project training, short-term enterprise practice, industry mentor training, and university-industry supervisor roundtables. These activities focused on developing practice-oriented supervision, project guidance, and collaborative education capacity.

The dual-path development arrangement responded to variation in supervisors' backgrounds. Academically oriented supervisors with limited enterprise experience were encouraged to participate in enterprise practice, industrial workshops, and joint projects. Enterprise mentors or practice-oriented supervisors who were less familiar with academic supervision received support on thesis standards, research ethics, student development, and assessment requirements. The intention was not to standardise all supervisors, but to create complementary capability structures. In this way, internal supervisors and industry mentors could jointly support students' academic inquiry and practical problem solving.

In the appraisal component, the focal school incorporated multidimensional indicators and process evidence into monitoring supervisor responsibilities. Supervisor appraisal consisted of three modules: academic contribution, student development, and professional ethics and conduct. Different weights were assigned according to supervisor type. For professional-degree supervisors, appraisal additionally included participation in practical projects, practice-based innovation outputs, contributions to solving industry problems, supervision process records, anonymous student evaluation, and university-industry collaboration. Professional ethics and conduct served as a threshold criterion that could override the overall appraisal result.

The appraisal component shifted attention from output enumeration to process visibility. Supervision records documented meeting frequency, feedback on research progress, project guidance, and support for academic integrity. Student evaluations provided evidence on perceived feedback quality and developmental support. University-enterprise records indicated whether supervisors were involved in joint projects, practice bases, and applied innovation activities. These sources are not free from bias and must be interpreted cautiously, but they make aspects of supervision that are often invisible available for discussion and improvement.

In the feedback-use component, diagnostic feedback and result use were incorporated into supervisor development. Annual appraisal results were used to adjust postgraduate-place allocations, resource support, performance incentives, diagnostic feedback, and periodic review and adjustment of supervisor qualifications. Evaluation results therefore no longer functioned only as annual appraisal conclusions; they entered processes of supervisor capability improvement, resource allocation, and qualification adjustment. Table 4 presents the problem-response mechanism of the model.

**Table 4. Problem-response mechanism of the four-in-one supervisor development model.**

| Component           | Problem addressed                                                                       | Main measures                                                                                                           | Evaluation mechanism                                                                                           |
|---------------------|-----------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|
| <b>Selection</b>    | Supervisor appointment standards were insufficiently aligned with programme type        | Differentiated selection, prior topic review, and expert assessment                                                     | Evaluate the fit among supervisor qualification, topic direction, and programme goals                          |
| <b>Development</b>  | Early-career supervisors had insufficient capacity to supervise practice-based projects | Industry-based practice, industrial workshops, industry-mentor training, and university-industry supervisor roundtables | Evaluate development of practice-based supervision, project guidance, and collaborative education capacity     |
| <b>Appraisal</b>    | Appraisal indicators overemphasized research outputs                                    | Diversified indicators, process records, student feedback, and a non-compensatory professional-conduct criterion        | Evaluate supervision processes, student development outcomes, practical contribution, and professional conduct |
| <b>Feedback use</b> | Evaluation results were insufficiently used                                             | Diagnostic feedback, allocation of postgraduate places, performance incentives, and periodic adjustment                 | Use evaluation results in resource allocation, capability improvement, and qualification management            |

Feedback use is the mechanism through which evaluation becomes developmental rather than merely classificatory. At the individual level, supervisors received evidence on strengths, weaknesses, and recommended development activities. At the organisational level, aggregated evidence informed training themes, industry collaboration priorities, and resource allocation. This dual use matters because supervisor development is not only an individual professional learning issue. It also depends on whether the institution can convert evidence into support structures, incentives, and periodic qualification review.

### 4.3. First-round formative evidence from implementation

First-round monitoring evidence indicated early changes in supervisor composition, students' perceived supervision quality, and university-industry collaboration after model implementation. Appointment records showed greater attention to substantial industry research and development experience in the appointment of professionally oriented programme supervisors. Internal student monitoring showed higher reported satisfaction with supervision in the first implementation round than at baseline. University-industry collaboration records documented three newly established joint laboratories and five approved key technology research projects. These records provide evidence of appointment decisions, organisational infrastructure, and project approval, but remain proximal indicators. They do not directly measure changes in supervisory practice, student learning, thesis quality, completion, or employment. They are therefore interpreted as formative evidence for judging implementation direction and identifying areas for further improvement, rather than as evidence of long-term impact [18, 20]. Table 5 summarises the evidence and interpretation boundaries.

First-round records documented the initial implementation of the model. Recent appointment records indicate that differentiated selection was being applied. Student satisfaction monitoring provided an additional source of formative feedback on students' supervision experiences. The establishment of joint laboratories and key technology projects indicates movement from informal collaboration toward more stable platforms. Taken together, these findings show that elements of the model were being operationalised, while leaving open the question of sustained educational impact.

Operational evidence from appraisal reform appeared in three areas: process records, feedback use, and institutional linkage. In annual appraisal, the focal school added supervision process records, anonymous student feedback, university-industry collaboration, and professional ethics and conduct. Supervisor evaluation therefore no longer depended only on outcomes such as publications, projects, and funding. After the first round of appraisal, the focal school used supervision process records and student feedback to generate diagnostic feedback for supervisors and to guide subsequent training arrangements.

Some supervisors participated in industry project practice, industrial workshops, or supervision-capability training in response to appraisal feedback. For professional-degree supervisors, opportunities to participate in projects, postgraduate-place allocations, and resource support also began to refer to practical education outcomes and collaborative guidance. Periodic review of supervisor qualifications, linkage to

postgraduate-place allocations, and performance incentives were incorporated into institutional texts. Evaluation results therefore began to enter the processes of supervisor development support and resource allocation.

**Table 5. First-round formative evidence and interpretation boundaries.**

| Outcome area                             | Observed indicator                                                                   | First-round evidence                                                                                                          | Interpretation boundary                                                                                                                        |
|------------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Supervisor composition</b>            | Practice background of newly appointed professionally oriented programme supervisors | Appointment records showed greater attention to substantial industry R&D experience                                           | Appointment standards were being applied, but the evidence does not demonstrate long-term student-development effects                          |
| <b>Guidance capacity</b>                 | Student monitoring of supervision                                                    | Recorded satisfaction was 72.3% at baseline and 86.5% in the first implementation round (descriptive comparison)              | Comparability of the samples and survey administration was not established; the figures are presented descriptively without causal attribution |
| <b>University-industry collaboration</b> | Joint platforms and collaborative projects                                           | Three joint laboratories were established, and five key technology research projects were approved                            | Platform construction progressed, but the quality of project-based cultivation requires continued tracking                                     |
| <b>Development motivation</b>            | Use of appraisal results                                                             | Institutional texts linked appraisal results to postgraduate-place allocation, performance rewards, and qualification review. | Institutional linkage was established, but behavioural effects require long-term observation                                                   |

Differentiated selection, dual-path development, diversified appraisal, and diagnostic feedback respectively addressed the major problems in supervisor appointment, practice-oriented guidance, appraisal indicators, and use of results. However, these changes remain first-round formative evidence. They do not yet support conclusions about stable long-term effects on supervisors' guidance behaviour or graduate education quality.

The remaining issue is organisational coordination. A four-in-one model requires data on selection, development participation, student feedback, appraisal, and industry collaboration to be interpreted across administrative boundaries. If these evidence streams remain fragmented, the improvement cycle will be difficult to sustain. The first implementation round therefore points to the need for a more integrated evidence platform and clearer responsibility for using evaluation results. Subsequent monitoring should examine whether feedback changes supervisory behaviour and whether enterprise engagement becomes regular rather than episodic. Future evaluation should incorporate more objective outcome indicators, including graduate achievement, thesis quality, employment development, and independent external evaluation results.

## 5. Discussion

This study examined graduate supervisor development in a Chinese private university through a formative educational evaluation framework. The findings

suggest that supervisor development should be understood not only as an individual training issue but also as an institutional quality-improvement issue. The baseline diagnosis indicated concerns regarding differentiated selection, practice-oriented guidance capability, university-industry support, and appraisal use. These concerns are consistent with research showing that supervision quality cannot be assessed only through degrees, academic titles, publications, grants, and research projects [1-4]. In professionally oriented postgraduate programmes, supervisors are also expected to support project-based learning, practical problem solving, industry engagement, and students' professional development.

The findings further suggest that the four diagnostic areas may be related. Limited recognition of practice-oriented capability in appointment criteria may reduce its visibility in appraisal, while unstable enterprise collaboration may constrain supervisors' access to authentic projects and opportunities for practice-based professional learning. However, the present study does not establish a causal sequence among these factors. These relationships should therefore be understood as plausible institutional mechanisms that require further examination rather than as confirmed causal findings.

The contribution of the four-in-one model lies in linking differentiated selection, dual-path capability development, multidimensional process appraisal, and diagnostic feedback use within a continuous quality-improvement cycle. Selection, professional development, appraisal, and feedback are familiar elements of higher education management. The model does not claim that these elements are individually new. Instead, it makes their relationships explicit: appointment criteria should reflect programme goals; development activities should address identified capability gaps; appraisal should include evidence of supervision processes; and feedback should inform subsequent decisions about training, resources, incentives, allocation of postgraduate places, and qualification review.

This integrated logic is consistent with utilization-focused evaluation, the Context, Input, Process, and Product model, school self-evaluation research, and educational development evaluation [14-20, 23, 24]. These perspectives emphasize that evaluation should connect needs analysis, resources, implementation, evidence collection, feedback, and adjustment. The four-in-one model therefore contributes an evaluable institutional logic for supervisor development rather than a fixed administrative template for all universities.

The early implementation evidence supports cautious conclusions. Appointment records, revised appraisal arrangements, joint laboratories, collaborative projects, and student satisfaction monitoring indicate that elements of the model were being operationalised. However, these indicators represent institutional arrangements, organisational outputs, and stakeholder perceptions. They do not directly demonstrate changes in supervisory behaviour, student learning, thesis quality, completion, employment, or sustained enterprise participation. The findings should therefore be interpreted as formative implementation evidence rather than causal evidence of long-term effectiveness.

The potential value of the study for other institutions lies primarily in its evaluation logic. Institutions offering professionally oriented postgraduate programmes may examine whether supervisor appointment criteria reflect programme needs, whether development activities address identified gaps, whether appraisal includes process evidence, and whether evaluation findings are used in

subsequent institutional decisions. However, the framework requires local adaptation because institutional authority, disciplinary traditions, supervisor composition, industry resources, and partnership conditions vary across settings. The study focuses on a single academic school within one institution during an early implementation stage. The evidence primarily concerns stakeholder perceptions, implementation processes, and organisational arrangements. The internally developed questionnaire supported item-level descriptive analysis but had not undergone full psychometric validation. Reliance on self-reported and internal monitoring data, the absence of a comparison group, limited objective measures of educational outcomes, and the lack of long-term follow-up limit conclusions about causal effectiveness, sustainability, and institutionalisation. The applicability of the framework to other institutional contexts remains to be examined.

Future research should examine whether the early implementation signals reported in this study persist across cohorts and whether feedback use produces sustained changes in supervisor development practices. Longitudinal studies could track supervision processes, student experiences, project participation, and university-industry collaboration over time. Comparative case studies could investigate how similar frameworks operate across disciplines, institutional types, and regional contexts. Stronger outcome evidence should also be incorporated, including thesis quality, project completion, practical innovation capacity, employment relevance, and independently assessed supervision quality. Such research would make it possible to test the proposed quality-improvement logic rather than infer long-term impact from early implementation evidence alone.

## 6. Conclusions

This study developed a formative evaluation framework for examining graduate supervisor development in professionally oriented postgraduate programmes. It shows how differentiated selection, capability development, process appraisal, and feedback use can be connected as institutional functions within a quality-improvement cycle.

The study contributes an evaluative perspective that shifts attention from isolated supervisor-training activities to the institutional relationships among appointment, support, evidence collection, and subsequent decision making. This perspective provides a way to examine whether supervisor-development arrangements are aligned with programme goals and whether evaluation findings are used for improvement.

The framework is intended as an adaptable analytical tool rather than a universal administrative model. Its application should be shaped by local programme missions, supervisory structures, disciplinary conditions, and university-industry resources.

## Data Availability Statement

The data are not publicly available because they include institutional monitoring records and participant responses collected for internal formative evaluation. De-identified materials may be made available from the corresponding author upon reasonable request, subject to institutional permission.

## References

1. Bearman, M.; Tai, J.; Henderson, M.; Esterhazy, R.; Mahoney, P.; and Molloy, E. (2024). Enhancing feedback practices within PhD supervision: A qualitative framework synthesis of the literature. *Assessment & Evaluation in Higher Education*, 49(5), 634-650.
2. Lee, A. (2008). How are doctoral students supervised? Concepts of doctoral research supervision. *Studies in Higher Education*, 33(3), 267-281.
3. Mainhard, T.; Van Der Rijst, R.; Van Tartwijk, J.; and Wubbels, T. (2009). A model for the supervisor-doctoral student relationship. *Higher Education*, 58(3), 359-373.
4. Pearson, M.; and Brew, A. (2002). Research training and supervision development. *Studies in Higher Education*, 27(2), 135-150.
5. Ministry of Education of the People's Republic of China; National Development and Reform Commission; and Ministry of Finance of the People's Republic of China. (2020). Opinions on accelerating the reform and development of graduate education in the new era [in Chinese; title translated by authors]. Retrieved September 30, 2026, from [https://www.moe.gov.cn/srcsite/A22/s7065/202009/t20200921\\_489271.html](https://www.moe.gov.cn/srcsite/A22/s7065/202009/t20200921_489271.html)
6. Shi, Y. (2022). On the selection and application of policy tools in the construction of supervisors teams since the reform and opening-up. *Academic Degrees & Graduate Education*, 2022(4), 20-27.
7. Yin, J.; Ye, F.; and Tian, P. (2024). On modeling the ability factors of professional degree postgraduate supervisors from the perspective of Juche education - An analysis on knowledge graphs with supervisors for Master's Degree students in Education as an example. *Journal of Graduate Education*, 2024(5), 71-77.
8. Zhang, X. (2023). The connotation and realization path of humanistic care in the responsibilities of postgraduate supervisor. *Academic Degrees & Graduate Education*, 2023(6), 15-19.
9. Fossland, T. (2023). Becoming a professional supervisor: Doctoral supervisors' development in a mandatory, large-scale development programme. *Journal of Praxis in Higher Education*, 5(2), 27-52.
10. Halse, C. (2011). 'Becoming a supervisor': the impact of doctoral supervision on supervisors' learning. *Studies in Higher Education*, 36(5), 557-570.
11. Kiley, M. (2011). Developments in research supervisor training: Causes and responses. *Studies in Higher Education*, 36(5), 585-599.
12. Lee, A. (2018). How can we develop supervisors for the modern doctorate? *Studies in Higher Education*, 43(5), 878-890.
13. Vähämäki, M.; Saru, E.; and Palmunen, L.M. (2021). Doctoral supervision as an academic practice and leader-member relationship: A critical approach to relationship dynamics. *The International Journal of Management Education*, 19(3), 100510.
14. Patton, M.Q. (2008). *Utilization-focused evaluation*. (4th ed.). Thousand Oaks, CA: SAGE Publications.
15. Stufflebeam, D.L. (2003). The CIPP model for evaluation. In Kellaghan, T.; and Stufflebeam, D.L. (Eds.). *International handbook of educational evaluation*. Kluwer Academic Publishers, Dordrecht. Vol. 9, 31-62.

16. Stufflebeam, D.L.; and Coryn, C.L.S. (2014). *Evaluation theory, models, & applications*. Jossey-Bass/Wiley, San Francisco, CA.
17. Chalmers, D.; and Gardiner, D. (2015). An evaluation framework for identifying the effectiveness and impact of academic teacher development programmes. *Studies in Educational Evaluation*, 46, 81-91.
18. Hum, G.; Amundsen, C.; and Emmioğlu, E. (2015). Evaluating a teaching development grants program: Our framework, process, initial findings, and reflections. *Studies in Educational Evaluation*, 46, 29-38.
19. Richter, E.; and Richter, D. (2024). Measuring the quality of teacher professional development: A large-scale validation study of an 18-item instrument for daily use. *Studies in Educational Evaluation*, 81, 101357.
20. Taylor, K.L.; and Znajda, S.K. (2015). Demonstrating the impact of educational development: The case of a course design collaborative. *Studies in Educational Evaluation*, 46, 39-46.
21. Straub, R.; Kulin, S.; and Ehmke, T. (2021). A transdisciplinary evaluation framework for the assessment of integration in boundary-crossing collaborations in teacher education. *Studies in Educational Evaluation*, 68, 100952.
22. Veugen, M.J.; Gulikers, J.T.M.; and den Brok, P. (2021). We agree on what we see: Teacher and student perceptions of formative assessment practice. *Studies in Educational Evaluation*, 70, 101027.
23. Kyriakides, L.; and Campbell, R.J. (2004). School self-evaluation and school improvement: A critique of values and procedures. *Studies in Educational Evaluation*, 30(1), 23-36.
24. Charalambous, E.; Kyriakides, L.; and Creemers, B.P.M. (2018). Promoting quality and equity in socially disadvantaged schools: A group-randomisation study. *Studies in Educational Evaluation*, 57, 42–52.
25. Paufler, N.A.; King, K.M.; and Zhu, P. (2020). Promoting professional growth in new teacher evaluation systems: Practitioners' lived experiences in changing policy contexts. *Studies in Educational Evaluation*, 65, 100873.
26. Martinez, F.; Taut, S.; and Schaaf, K. (2016). Classroom observation for evaluating and improving teaching: An international perspective. *Studies in Educational Evaluation*, 49, 15-29.
27. White, M.; and Maher, B. (2024). What value do standardized observation systems add to summative teacher evaluation systems? *Studies in Educational Evaluation*, 83, 101414.
28. Lindorff, A.; Jentsch, A.; Kaiser, G.; Walkington, C.; and Sammons, P. (2020). Hybrid content-specific and generic approaches to lesson observation: Possibilities and practicalities. *Studies in Educational Evaluation*, 67, 100919.
29. Creswell, J.W.; and Plano Clark, V.L.P. (2017). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: SAGE Publications.
30. Yin, R.K. (2018). *Case study research and applications: Design and methods* (6th ed.). Thousand Oaks, CA: SAGE Publications.
31. Braun, V.; and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
32. Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). Thousand Oaks, CA: SAGE Publications.