

Article

Data-Driven Quantitative Performance Evaluation of Construction Supervisors

Cheng Yang ¹, Jia-Rui Lin ^{2,*}, Ke-Xiao Yan ², Yi-Chuan Deng ³, Zhen-Zhong Hu ⁴ and Cheng Liu ⁵

¹ Zhejiang Supervision on Highway and Water Transportation Construction Engineering Co., Ltd., Hangzhou 310000, China; yangcheng20221003@163.com

² Department of Civil Engineering, Tsinghua University, Beijing 100084, China; quincykx@tsinghua.edu.cn

³ School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, China; ctycdeng@scut.edu.cn

⁴ Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China; hu.zhenzhong@sz.tsinghua.edu.cn

⁵ China-Road Transportation Verification & Inspection Hi-Tech Co., Ltd., Ministry Transport, Beijing 100088, China; liucheng1991@gmail.com

* Correspondence: lin611@tsinghua.edu.cn

Abstract: The performances of construction supervisors are essential for the monitoring, control, and coordination of the construction process of a project in order to adhere to a predefined schedule, cost, quality and other factors. However, it is challenging to evaluate their performance due to limitations such as data deficiency, human error, etc. Thus, this paper proposes an approach to data-driven quantitative performance evaluation of construction supervisors by integrating an analytic hierarchy process (AHP) and activity tracking. The proposed approach contains three parts, namely, index extraction, weighting, data-driven index calculation, and then validation by case study. Firstly, performance indexes were developed based on a literature review as well as surveys and function analysis of the information system for construction supervision (CSI system). Then, the weights of and relationships among of the indexes are determined by AHP. After that, with daily workflow and inspection activities tracked in the CSI system, a method and a software module for automatic calculation of indexes were developed. Lastly, the proposed approach was validated by a real-world case. The result showed that the proposed approach can quantify the performance of a construction supervisor systematically and automatically, which shed lights on how to evaluate the performance of a worker based on the tracking of daily activities. The data-driven process enhanced our strong interpretation of member actions and evaluation indexes, and can boost the performance of every member in an organization.



Citation: Yang, C.; Lin, J.-R.; Yan, K.-X.; Deng, Y.-C.; Hu, Z.-Z.; Liu, C. Data-Driven Quantitative Performance Evaluation of Construction Supervisors. *Buildings* **2023**, *13*, 1264. <https://doi.org/10.3390/buildings13051264>

Academic Editor: Lucio Soibelman

Received: 17 April 2023

Revised: 7 May 2023

Accepted: 9 May 2023

Published: 11 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As the essential phase of a building or infrastructure project, construction accounts for the most cost, time, and resources planning due to the safeguarding, fabrication, and quality control during this period. The construction supervisor, a key component of a project management system, is important for a successful construction process [1]. The construction supervisors have duties related to process management, quality management, safety management, and other various tasks during a construction period, and should accept the commission organized by either the supervisor or enterprise [2]. These duties of construction supervisors contain several typical categories [3], such as activity monitoring, safeguarding, planning the establishment, coordination, reviewing of plans, and assisting. In addition, the construction supervisors must be qualified for education, certification, and experience. Since the staff of the project supervision department are actively involved

in site construction, accurate evaluation of the performance and dynamic assessment of individuals becomes a follow-up task.

In the early 1980s, researchers started to develop a better understanding of how to construct impressions and judgments about the performances of subordinates [4]. For the next fifteen years, the research on performance evaluation shifted in the direction of information processing [1]. After this period, researchers around the world started to rethink the relationships (among owners, construction managers, and architects), links (between construction craft workers and superintendents), and interfaces (between conventional means and digital platform) existing in performance evaluation [2–4]. With the growing trend of studying performance evaluation, the evaluation system has been used to predict behaviors based on the past data and to attach multiple values upon data acquisition, such as position track and rate of progress [5]. The performance result is feedback to clear the current employee performance levels [6]. In addition, evaluation methodology and evolving computer science have been integrated with smart concepts in recent decades to make them more feasible and reliable [7]. Thus, the degree of automation for the performance evaluation system should represent smart characteristics such as relevance, intelligence, synergy, and predictivity [8].

For a long time, the performance evaluation of construction supervisors has been paper-based, which is always tedious, time-consuming, and error-prone [7]. Despite the fact that the efficiency of evaluation has been improved due to the upgrades to computer technology, commonly collected data for performance evaluation are still limited. The lack of details on daily activities and the approaches adopted for performance evaluation still suffer from problems such as subjectivity, unfairness, half-heartedness, and even mistakes [1,6–8]. Thus, a performance evaluation system based on the conventional methods may lead to potential negative effects and impact the success of a construction project [5,7].

In order to tackle the above problems, this study introduces a new approach to data-driven quantitative performance evaluation of construction supervisors by integrating AHP and automatic processing of daily activity tracking data. This has both practical and theoretical implications. The practical implications include the collection of workers' performance data for evaluation, data that are calculated using a data-driven process for automation purposes. The theoretical proposed approach, like the evaluation indexes, applies a more interpretative model to improve the conventional evaluation system. Section 2 discusses previous studies, introduces the empirical study to be applied to construction, and then introduces the modern methods. At the end of this section, a summary of the above discussion is given which introduces the aspects of our proposed method. Section 3 illustrates the data-driven quantitative performance evaluation method, including index extraction, weighting, index calculation, and case analysis and verification. Section 4 shows a case that applies this approach. Section 5 concludes this study and offers suggestions for future work.

2. Literature Review

In the 1960s, only a few performance evaluation methods were available, and the utilized indexes were quite few [9]. From the 1970s to 1980s, the study of performance evaluation methods went deeper and wider to develop core indexes, especially for the performance evaluation system formed by the financial industry [9,10]. In the 1990s, scholars began to try to address the shortcomings of the independence, subjectivity, and enforceability of performance evaluation methods. This stage was the reflective reevaluation phase of performance evaluation [9]. In the following fifteen years, scholars supplemented the previous studies while improving on information processing with computer techniques, developed as shown below in Figure 1. The performance evaluation method began to integrate with computer science developments such as smart sensor systems, digital platform application, data science, and artificial intelligence [7–9]. Figure 1 shows a brief timeline of the progress of performance evaluation methods since the 1980s.

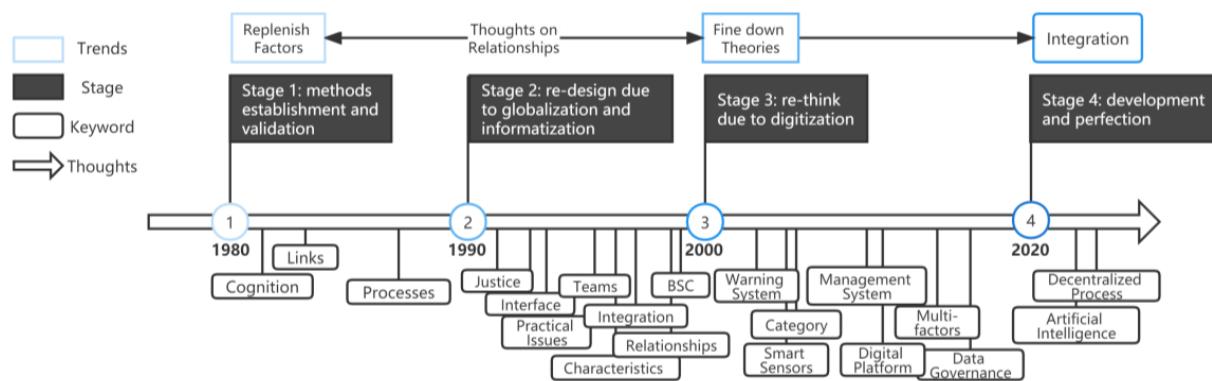


Figure 1. A timeline of the progress of performance evaluation methods [5,10–38].

Historically, performance evaluation has used to describe the periodic assessment of duties and to report on the performance of an individual, as well as the presumed requirements of a job [7]. A typical performance evaluation method includes three parts: index determination, weighting, and evaluation. Firstly, the evaluation method needs to express the job performance based on the related indexes in order to evaluate the strengths and weaknesses of a job. These related indexes are conventionally determined by peers, supervisors, and workers themselves. Secondly, the weighting processes has used both traditional methods, such as in-depth interviews and questionnaires, and modern methods, such as AHP and fuzzy comprehensive evaluation. Lastly, the evaluation portion typically involves subjectivity from peers, supervisors, and human resource management (HRM). A common way of evaluation pattern is to use a checklist table that is offered by a company, to then be scored by HRM. In addition, the newly scoring way is performed by a digital platform as an automated process [39].

2.1. Typical Weighting Determination Methods in Construction

2.1.1. Conventional Weighting Determination Methods

The conventional methods mainly establish empirical links between individuals and contributions. These methods, as shown in Table 1 below, indicate the relationship, evaluation, and weighting based on the perspectives of people. The accuracy of evaluation is strongly referring to the capability of assessors.

Table 1. The comparison of the conventional weighting determination methods [18,21,34].

Methods	Advantages	Disadvantages	Scenarios
Subjective experience method	High efficiency and low cost	Low credibility due to over-subjectivity	Assessors set weights for indexes based on their experience if assessors are familiar with and understand the object.
Questionnaire	Uniformity and generalization, easy to quantify	Non-guarantee of quality	A questionnaire is designed for the assessment and determines the index system and weighting analysis.
In-depth interview	Strong flexibility, deep	Difficulty for the host	The in-depth interview is described as a face-to-face conversation with related individuals to understand their working modes, natures, and other aspects, such as the corresponding evaluation index system and weighting, for reference.
Expert investigation method	Strong representativeness and scientific with reliability and authority	High cost, difficult to organize	Experts are invited to conduct research on the index system. Each expert sets the weight of the index independently, and then takes the average value of the weight of each index as the final weight.

2.1.2. Modern Weighting Determination Methods

1. The Analytic Hierarchy Process

AHP takes a complex multi-objective decision-making problem as a system. The hierarchical structure is drawn by dividing the goal of the decision, the considered factors, and the object of decision into the highest level, the middle level, and the lowest level according to their mutual relations [40]. Hierarchical targets establish single or multi-level analysis structure models. By sorting out the nature of the problem, different factors can be obtained by decomposing the target. According to the relative influence, the correlation, subordination, and importance can be analyzed. The clustering can model the internal relationship of the problem. Finally, the weight relationships among the index factors at the bottom of the model and the top level, namely, the target, are systematically analyzed, and the ranking method is used to sort them according to their importance. The process of sorting calculation requires the 9-point scale method (5-point or 7-point scale method) to evaluate each factor. Then, a judgement matrix should be formed to obtain data that can be compared and quantified. According to the quantified values of these evaluations, the maximum eigenroots and corresponding eigenvectors can be calculated, and then the weights can be given [41,42].

2. Fuzzy Comprehensive Evaluation

Fuzzy mathematics uses mathematical tools to solve the problems of fuzzy things [43]. Therefore, the fuzzy comprehensive evaluation method based on fuzzy mathematics addresses complex problems, which are difficult to describe with precise mathematics, as a clever bridge between formal thinking and complex systems. This evaluation method provides a general assessment of a subject with multiple constraints [44].

3. Data Envelopment Analysis

Data envelopment analysis is a systematic analysis method to evaluate the relative effectiveness or benefit of the same type of unit based on multi-index inputs and multi-index outputs. It is a marvelous way to deal with multi-objective decision-making problems. The relative effectiveness of a decision-making unit (DMU) is called DEA efficiency [45].

2.2. Typical Methods for Performance Evaluation of Construction Workers

1. Critical Incident Method

In 1954, Ferraligen and Bayless proposed an objective evaluation system [46]. This evaluation system analyzed the tasks during a construction period, recorded the best or worst critical events in detail, and collected relevant information to identify the behaviors and performances.

2. Graphic Rating Scale

The schematic diagram is used to display the factors of the performance evaluation and the evaluation information, such as the scoring standard, evaluation grade, and comments [47]. Meanwhile, it is easier to determine the factors in line with the performance status of construction when analyzing the evaluation scores.

3. Behavior Checklist

This approach uses behavioral comparison tables to score or evaluate the objects one by one. The behavior checklist can play an accurate, fast, and easy role in describing the standardization of employee performance [48]. Meanwhile, the objectivity of the supervisors is corrected by the evaluation metrics in the comparison tables. Thus, the fairness is strong. It is easier to trace the source and correct disputes in the performance evaluation.

4. Management by Objectives

Management by objectives (MBO) includes three main processes: object formulation, the execution process, and performance feedback [49,50]. MBO pays attention to the

consistency of higher and lower goals, and regularly evaluates the performance of the construction task completion of the goals to clarify the indexes [49,50].

5. 360-degree Evaluation Method

The 360-degree evaluation method includes evaluation inputs from a number of stakeholders, such as immediate supervisors, team members, customers, peers, and oneself [51]. This method provides information about the influences of actions on construction workers [51].

6. Key Performance Indicator

Key performance indicators (KPI) focus on the input and output of the workflow, identifies the key points according to the process design, and then evaluates them [52]. On the basis of the consistency formed by the target, the KPI method can sort out the execution process effectively and pertinently, establish key points to achieve the target or optimization results, and then fully evaluate whether the construction workers have completed these established points [52].

7. Balanced Score Card

The balanced score card, developed by Kaplan and Norton in the early 1990s [53]. The balanced score card (BSC) is a conceptual framework for translating an organization's strategic objectives into a set of performance indexes according to four perspectives: (1) financial, (2) customer, (3) internal business process, and (4) learning and growing.

In summary, the typical performance evaluation methods indicate the relationship between construction workers and the evaluated indexes [46–53]. The relationship reveals the link between degree of automation and data size, which shows a positive correlation property [6]. As the performance evaluation methods become more comprehensive and systematic, the relationships and dimensions that need to be sorted out become more complex, resulting in an increasing demand for data, as Table 2 shown.

Table 2. The comparison of the typical performance evaluation methods [6,46–53].

Methods	Points of Focus	Degree of Automation	Data Size
Critical Incident Method	Critical events	Manual	Low
Graphic Rating Scale	Factors	Manual	Low
Behavior Checklist	Objects	Manual	Low
Management by Objectives	Goals	Manual or Semi-automatic	Low
Key Performance Indicator	Key points	Manual or Semi-automatic	Low
360-degree Evaluation Method	Multiple dimensions	Manual or Semi-automatic	Middle
Balanced Score Card	Internal process	Manual or Semi-automatic	Middle

2.3. Summary

2.3.1. Advantages of Emerging Information Technologies

The construction supervision industry tends to integrate big data technology and blockchain technology, which continuously and deeply penetrate various fields, bridging previously discrete areas [54–56]. Construction supervision makes progress due to its integration with intelligence, improving the accuracy and strengthening the safety of the work that originally required manual intervention. For example, artificial intelligence technology can be used to fully identify, remind, and record safe clothing or standard actions in construction [55]. This technique greatly improves the safety performance of the project, and this process is dynamic and continuous [55]. The new, intelligent system undertakes and upgrades intelligent equipment to be more efficient, uses data information accurately, and completely changes the method of operation in construction [57]. The current supervision mode tends to form a systematic, digital and intelligent management system, and finally achieves the goal of high standards, requirements, and quality of engineering construction projects [58,59]. For example, combining BIM technology and AI

technology can improve the efficiency of supervision work and, thus, promote the reform of the supervision industry and accurately control projects [55,59,60].

2.3.2. Performance Quantification

Quantification replaces the subjective method of evaluation with mathematical modeling [50]. The quantitative performance is intuitively reflective of the results. Therefore, the evaluation plan must be responsible for the outcome of the evaluation. Meanwhile, the evaluation indexes must be correct and feasible. By quantification of the indexes, construction supervisors are able to respond and track all critical activities and information by logging and recording throughout the life of an engineering project. For the determination of the evaluation indexes, the preliminary conclusions were summarized based on expert interviews and questionnaires. For weighting, a growing number of enterprises, institutions, and other relevant parties have adopted modern evaluation methods to make performance evaluation more scientific, humane, safe, and fair [7,50,58].

2.3.3. Necessity and Significance

Most of the aforementioned performance evaluation methods inherited characteristics from the financial and business field [11,50,52]. These manual evaluation methods still commonly involve subjective-based factors. However, this greatly hinders the progress and manpower of the construction project while potential problems such as the cost of time and human resources, as well as fairness, apply [61]. With IT technologies, it will be possible to track the daily activities of construction workers in an accurate and way, in real time. Thus, a new approach to performance evaluation of construction supervisors, which could quantify a worker's performance while considering both his/her daily activities and feedbacks from his/her partners automatically, is needed and will boost the performance of every member in an organization.

3. Methodology

To address the aforementioned problems, an overall framework and methodology, as shown in Figure 2, was adopted. In this study, we selected an evaluation team composed of researchers, engineers, supervisors, and superintendents in the field of construction engineering. These members were qualified to be consultants in engineering construction and project management, and have certain knowledge and understanding of the evaluation of construction supervisors. First of all, this study combined a literature review, an expert interview, a site survey, and a function analysis of construction supervision information (CSI) systems to identify potential indexes for performance evaluation. Specifically, potential indexes were extracted from the literature and categorized into different groups. The creation of an optimal index system, expert interview, site survey, and function analysis of CSI systems were considered to eliminate the unsuitable indexes and adjust the extracted indexes. Once the index system was optimal, a weight analysis based on AHP was conducted with data collected via questionnaires to determine the weight of each index. Thirdly, an automatic data extraction and index calculation tool was designed and implemented to analyze the daily workflow and inspection activities of each worker, and then to calculate his/her performance index automatically. Lastly, the developed approach was deployed in an updated version of the CSI system and validated through user feedback and beta tests of a case study. The application includes functional development based on evaluating indexes, embedding an algorithm based on the weighting of indexes, and automatically generating a digital document based on the automatic calculation module. A more detailed explanation of each step is provided below.

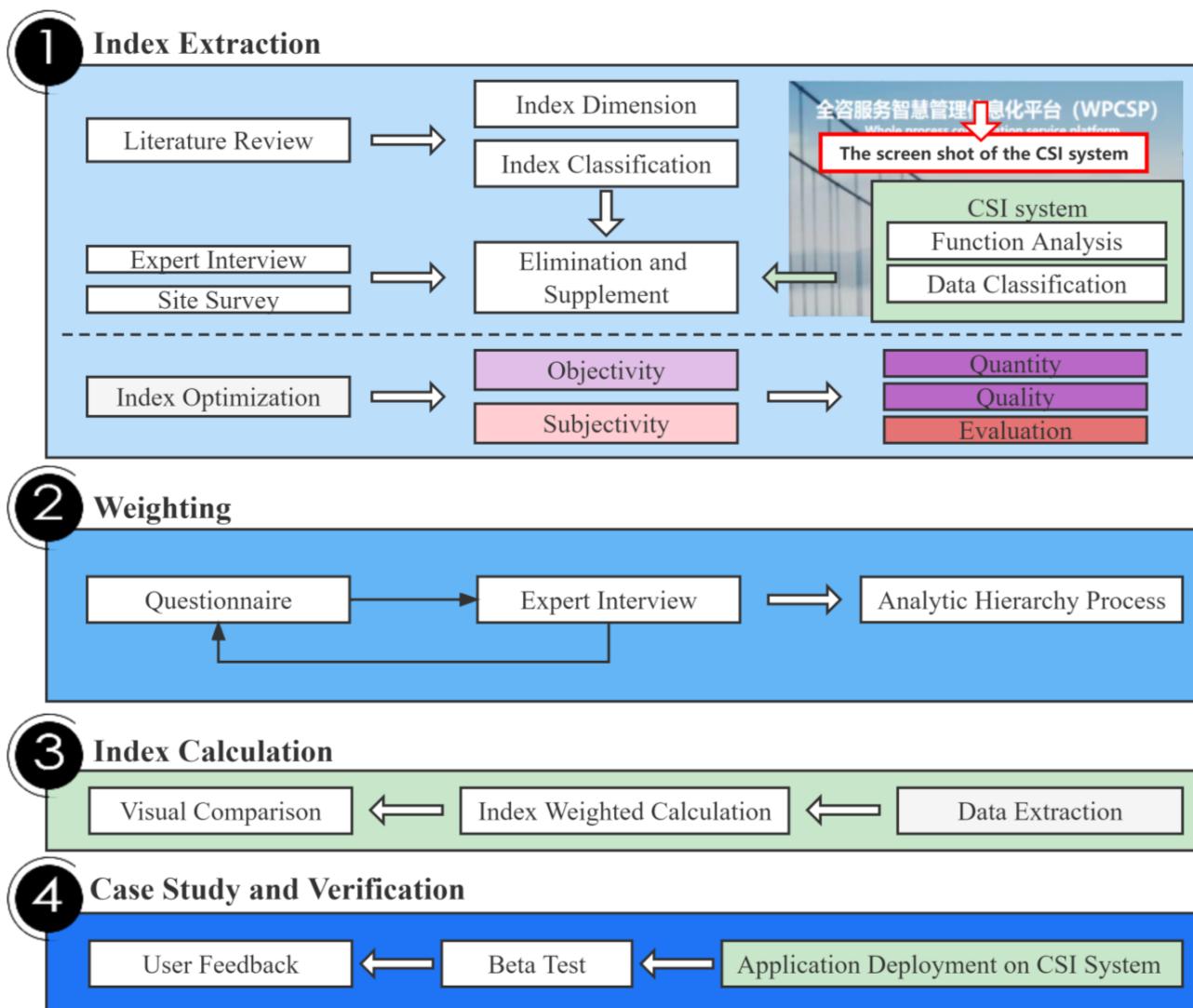


Figure 2. The overall framework of the proposed methodology.

3.1. Index Extraction

On one hand, the review from 1980 to 2020 provided results for evaluation indexes after filtering with keywords such as intelligence, smart, performance evaluation, and smart construction supervision. After the expert interview and questionnaire, the evaluation indexes could be determined as a basic conclusion. As a result, the influence path was sketched, as in Figure 3 below, to clarify the workability. Firstly, the index dimension was divided by two aspects: objectivity and subjectivity. The objectivity aspect contained two evaluation sub-items, namely, work quantity and work quality. Similarly, the subjectivity indicated evaluation by other significant factors. These divided dimensions both had responsibilities regarding work satisfaction and eventual performance evaluation. Furthermore, the work satisfaction was related to performance evaluation as well; the proposed method proved that there are features of relevance and synergy.

On the other hand, the evaluation indexes were extracted from the CSI system. The CSI system is comprised of several modules, such as Coordination Office, Integrated Management, Document Approval, and Data Center. Therefore, it was necessary to analyze the functions of modules before the extraction of indexes. For example, Coordination Office includes sub-items such as Notice, Message, and Personnel. These sub-items contain four parts, such as title, content, date, and status of read. Next, the valuable indexes were extracted based on the function analysis. The evaluation indexes were categorized

and supplemented through clustering. Figure 4, below, shows the procedure of index extraction. As a result, these indexes, chosen based on the influence path and the CSI system, were considered according to properties such as fairness, objectivity, subjectivity, and effectiveness.

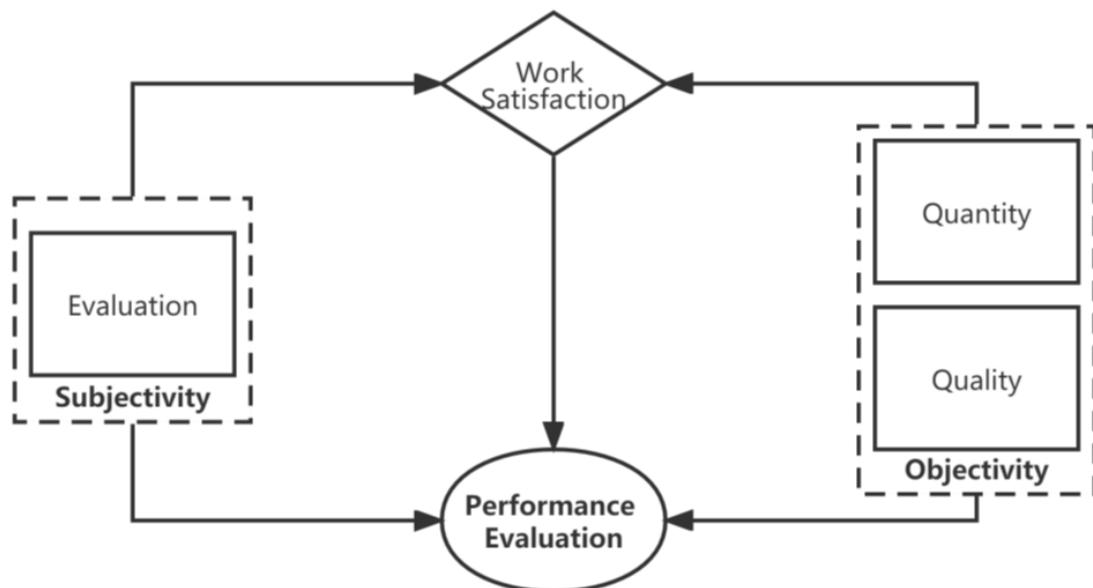


Figure 3. Influence path.

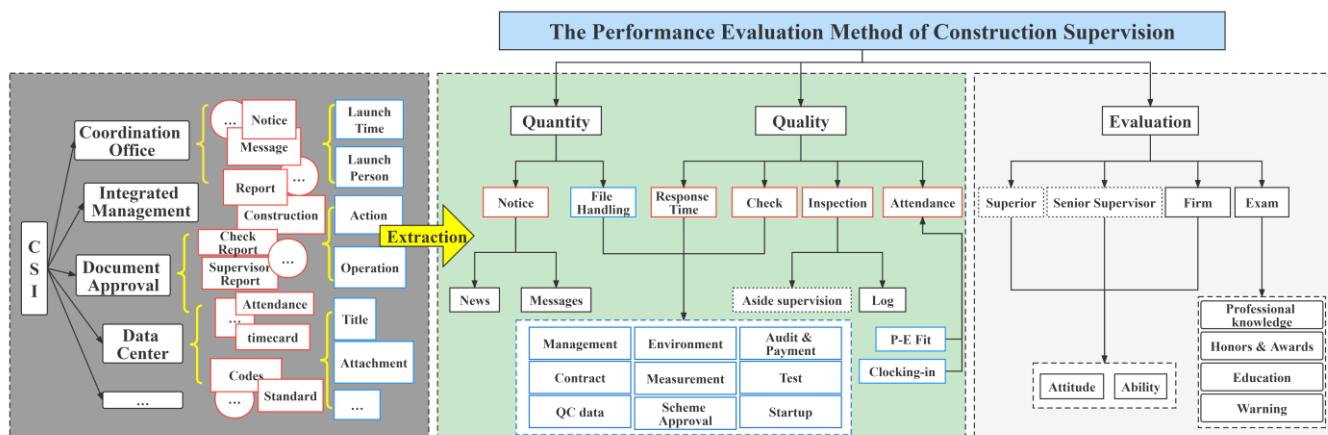


Figure 4. The flow chart of the procedure of index extraction.

3.2. Weighting

3.2.1. Weighting Based on AHP Approach

The weighting procedure started with the collection of information from the questionnaire. The targets of investigation were practitioners, engineers, and scientists with professional knowledge. By a literature review, the architecture of questionnaire was designed with fill-in-the-blank questions to collect non-sensitive data and choice questions to compute and weigh using AHP. The questionnaire required a well-designed scale value for all copies that were given to experts. This questionnaire used 7 scales for each comparison question, and the scale value indicated the weighting procedure. The definitions of the scale values were stated within each copy of the questionnaire, as shown in Table 3 below.

Table 3. Scale value for questionnaire.

Scale Value	Comparative Meaning of Relative Importance of Factor i and Factor j
1	Factor i is strongly unimportant to factor j
3	Factor i is unimportant than factor j
4	Factor i is equally important to factor j
5	Factor i is more important than factor j
7	Factor i is far more important than factor j
2, 6	Factor i is between the two adjacent scales above compared with factor j

Weighting was computed by Formula (1) below:

$$W_i = \bar{w}_i / \sum_{i=1}^n \bar{w}_i, W = \begin{Bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{Bmatrix} \quad (1)$$

CI was computed as Formula (2):

$$CI = \frac{\lambda_{max} - n}{n - 1} \quad (2)$$

where

CI is the consistency index;
 λ_{max} is the principal Eigen value;
 n is the Matrix size.

CR was computed as Formula (3):

$$CR = \frac{CI}{RI} \quad (3)$$

where

CI is the consistency index as a deviation or degree of consistency;
 RI is the random index;
 CR is the consistency ratio.

The random consistency index (RI) is shown in Table 4 below [62]:

Table 4. Random consistency index.

Matrix size	3	4	5	6	7	8	9	10	11	12	13	14	15	16
RI	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49	1.52	1.54	1.56	1.58	1.59	1.5943
Matrix size	17	18	19	20	21	22	23	24	25	26	27	28	29	30
RI	1.6064	1.6133	1.6207	1.6292	1.6358	1.6403	1.6462	1.6497	1.6556	1.6587	1.6631	1.6670	1.6693	1.6724

According to the feedback on the questionnaire and the result of the consistency index (CI), a logical method of constructing a judgment matrix for weighting could be determined. This result required that the CR be less than 0.1. If the value of CR was equal to or less than 0.1, the inconsistency was acceptable. Otherwise, the subjective judgment was revised. Furthermore, the purpose of the reliability analysis, as shown in Table 5, was to study the reliability and accuracy of the answers for use as quantitative data. If alpha was higher than 0.8, the reliability was high. If alpha was between 0.7 and 0.8, the reliability was good. If alpha was between 0.6 and 0.7, the reliability was acceptable. However, if alpha was less than 0.6, the reliability was poor. CITC told the items whether they should be deleted. If the value was less than 0.3, it was considered as inappropriate. Similarly, if the value of

“Alpha if Item Deleted” was significantly higher than alpha, then the item could be deleted and re-analyzed.

Table 5. Reliability analysis.

Item	Correction Item Total Correlation (CITC)	Alpha If Item Deleted	Cronbach's Alpha
Quantity	0.677	0.837	0.845
Quality	0.767	0.747	
Evaluation	0.719	0.777	
Standardized Cronbach's α coefficient: 0.854			

By using the data from the collected questionnaire, in this case, as Table 5 shows, Cronbach's alpha was 0.845, greater than 0.8, indicating a high reliability of the research data. For the “Alpha if Item Deleted”, the values did not increase significantly, indicating that the item should not be deleted. The CITC values of the analysis items were all greater than 0.4, indicating a good correlation between the analysis items and a good level of reliability.

In summary, the questions of the questionnaire had high internal consistency. The reliability coefficient of the research data was higher than 0.8 (Table 6), which indicated that the data's reliability was of high quality and could be used for further analysis.

Table 6. Cronbach reliability analysis.

Item	Correction Item Total Correlation (CITC)	Alpha If Item Deleted	Cronbach Alpha
Certification/Quantity	0.791	0.969	
Certification/Quality	0.827	0.969	
Certification/Evaluation	0.575	0.971	
The level of certification/Quantity	0.863	0.968	
The level of certification/Quality	0.885	0.968	
The level of certification/Evaluation	0.517	0.971	
Professional Engineering Competence/Quantity	0.901	0.968	
Professional Engineering Competence/Quality	0.790	0.969	
Professional Engineering Competence/Evaluation	0.546	0.971	
Attitude/Quantity	0.749	0.969	
Attitude/Quality	0.637	0.970	
Attitude/Evaluation	0.876	0.969	
Creativity/Quantity	0.797	0.969	
Creativity/Quality	0.833	0.969	
Creativity/Evaluation	0.831	0.969	
Practiced time/Quantity	0.737	0.970	
Practiced time/Quality	0.690	0.970	
Practiced time/Evaluation	0.706	0.970	
Skills Proficiency/Quantity	0.810	0.969	
Skills Proficiency/Quality	0.624	0.970	
Skills Proficiency/Evaluation	0.621	0.970	
Examination/Quantity	0.847	0.969	
Examination/Quality	0.878	0.968	
Examination/Evaluation	0.815	0.969	
Standardized Cronbach's α : 0.972			0.971

While Cronbach's alpha refers to the operation of reliability analysis, sensitivity analysis refers to the stability of the result. Figure 5, below, can be seen as the sensitivity analysis of the evaluation indexes. In general, the confidence interval is an estimation interval which is structured by the population parameter based on the sample statistics. The lines located at the values of 5.96 and 6.24 demonstrate the 95% confidence interval of the effect size. The middle line, at the value of 6.10, is the combined effect size. The small circle is the effect size if the item is deleted. If there is an outlier which drops out of the confidence interval, it means that the item has a strong influence on the analysis as an unstable factor. In summary,

the effect size of the evaluation indexes in this study was located in the confidence interval. The evaluation indexes showed a stable performance using the model, and they were able to be used for the following computation.

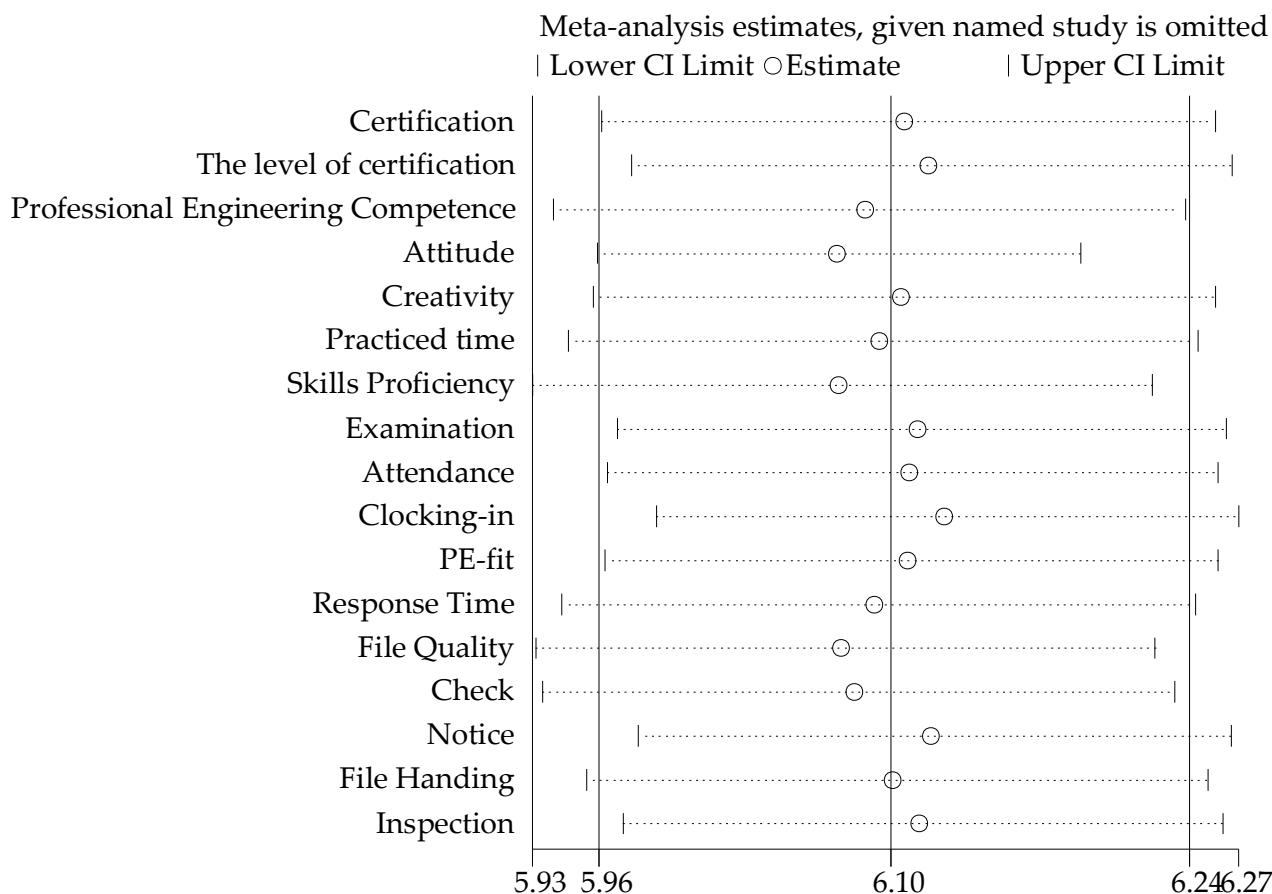


Figure 5. Sensitivity analysis of the evaluation indexes.

3.2.2. Comparison of Other Approaches

As shown in Table 7 below, the DEA method is different from AHP and fuzzy comprehensive evaluation in a number of respects, such as the trend of application and the amount of required data. Despite the fact that fuzzy comprehensive evaluation and AHP share key features, such as the trend of application, the quantity of required data, and the method of calculation, fuzzy comprehensive evaluation was unsuitable for this study due to its high complexity. Compared with the fuzzy comprehensive evaluation method, AHP uses a simple operation to achieve goals, such as hierarchicalization and weighting.

Table 7. A comparison between AHP, fuzzy comprehensive evaluation, and the DEA method.

	AHP	Fuzzy Comprehensive Evaluation	DEA
Advantages	(1) Systematic method of analysis; (2) Simple and useful; (3) Less required quantitative data.	(1) Membership involved; (2) Precise digital processing; (3) Vectorized evaluation results.	Applicable to complex decision-making problems;
Disadvantages	(1) Unable to offer a new option; (2) Less quantitative data but more qualitative data; (3) Complex calculation.	(1) Complex calculation; (2) Strong subjectivity on index weight vector; (3) Superfuzzy phenomenon.	(1) A mass of data required; (2) Strong subjectivity.

3.3. Index Calculation

After the AHP procedure, the function development was scheduled, including data extraction, index calculation, and comparative analysis. For example, the work quantity of the senior construction supervisor was calculated by the statistic of the quantity of tasks that launched in the CSI system. The work quantity comprised Notice and File Handling, which are the monthly number of notifications read and the monthly number of tasks processed, respectively. The work's quantity was computed by Formula (4):

$$M_{Quantity} = (M_{N_i} - \sum D_{N,i}) + (M_{FH_j} - \sum D_{FH,j}) \quad (4)$$

where

$M_{Quantity}$ is the mark of the work quantity of senior construction supervisor;

M_{N_i} is the default mark of “Notice”;

M_{FH_j} is the default mark of “File Handling”;

$D_{N,i}$ is the mark of deduction item i of “Notice”;

$D_{FH,j}$ is the mark of deduction item j of “File Handling”.

Similarly, the work quality included five indexes, including Response Time, Check, Inspection, and Attendance. The indexes of Response Time and Check were based on the contract terms and the company regulations. The Inspection contained the location, track, and log. The Attendance depended on Clocking-in and Person–Environment Fit (P-E Fit). The work's quality was computed by Formula (5):

$$M_{Quality} = \sum M - \sum D, \quad M = \begin{Bmatrix} M_{RT} \\ M_{Che} \\ M_{Ins} \\ M_{Att} \end{Bmatrix}, \quad D = \begin{Bmatrix} D_{RT} \\ D_{Che} \\ D_{Ins} \\ D_{Att} \end{Bmatrix} \quad (5)$$

where

$M_{Quality}$ is the mark of the work quality of the senior construction supervisor;

M_{RT} is the default mark of Response Time;

M_{Che} is the default mark of Check;

M_{Ins} is the default mark of Inspection;

M_{Att} is the default mark of Attendance;

D_{RT} is the mark of the deduction item of Response Time;

D_{Che} is the mark of the deduction item of Check;

D_{Ins} is the mark of the deduction item of Inspection;

D_{Att} is the mark of the deduction item of Attendance.

Overall, data acquisition was realized by Internet of Things front-end devices, the input of an intelligent mobile terminal, and association indexes, and then the generation of electronic quality assurance data was synchronized. Data extraction based on the generation and record then pushed the data stream towards the procedure of weighted calculation of the indexes. The method and software module for automatic calculation of the indexes were developed to achieve complete automated evaluation and visualized analysis. For example, if a senior construction supervisor were to re-work due to a failure in the Check process or an out-of-track assignment, the work quality mark would be deducted automatically in the CSI system. Construction, design, supervision, testing, and participating parties conducted collaborative management of the inspection and evaluation online, and achieved traceability throughout the construction process (Figure 6). The visual comparison ended with this step, and produced a visual result on the system for the users.

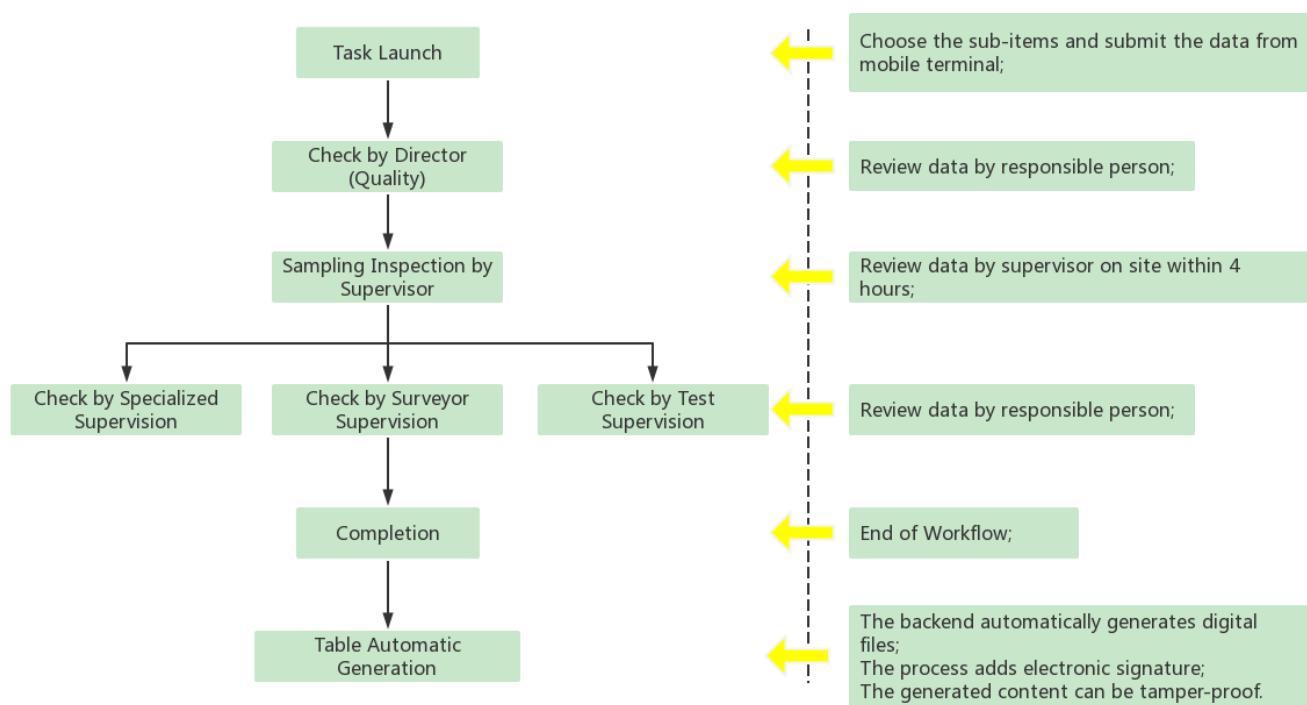


Figure 6. Traceability and visualization of the construction process.

3.4. Case Analysis and Verification

A survey with the co-operative construction supervision company was conducted to determine the requirements and insufficiency of the current CSI system. According to the survey, the method and software module for automatic calculation were developed to perfect the current management of performance evaluations. After a beta test and improvement based on users' feedback, the application was deployed on the CSI system for the performance evaluation of construction supervision.

Once the function was initially deployed, the beta version of the application was ready to be released for the beta test. The beta test collected data such as statistics on user operations, actions, reports, errors, and suggestions. Firstly, the beta test was required to check whether the functions were running properly or not. This check protocol referred to the requirements of the survey. The content of the check protocol included logic, interface, function, and typos. For example, a group of volunteers operated the CSI system on both PC and mobile terminals. Volunteers were chosen from among construction craft workers, construction supervisors, and superintendents. They clicked the buttons, submitted commands to the system, and checked whether the feedback was correct or whether bugs were present. Finally, the application was updated by analyzing the feedback and the debugging process.

4. Case Study

4.1. Background

Zhejiang Supervision on Highway and Water Transportation Construction Engineering Co., Ltd., Hangzhou, China has conducted the supervision business since 1993. The current business scope of the company includes construction supervision, test and inspection services for the highway and water transportation, and engineering procurement construction.

4.2. Index Extraction on Case Study

A new approach to data-driven quantitative performance evaluation of construction supervisors by integrating AHP and automatic processing of daily activity tracking data was applied to Zhejiang Supervision on Highway and Water Transportation Construction

Engineering Co., Ltd., Hangzhou, China. In this case, as Figure 7 shown, 110 questionnaires were collected in total. The job title distribution of the questionnaire included construction craft workers, at 10.87%; senior engineers, at 4.35%; intermediate engineers, at 21.74%; junior engineers, at 56.52%; and others, at 6.52%.

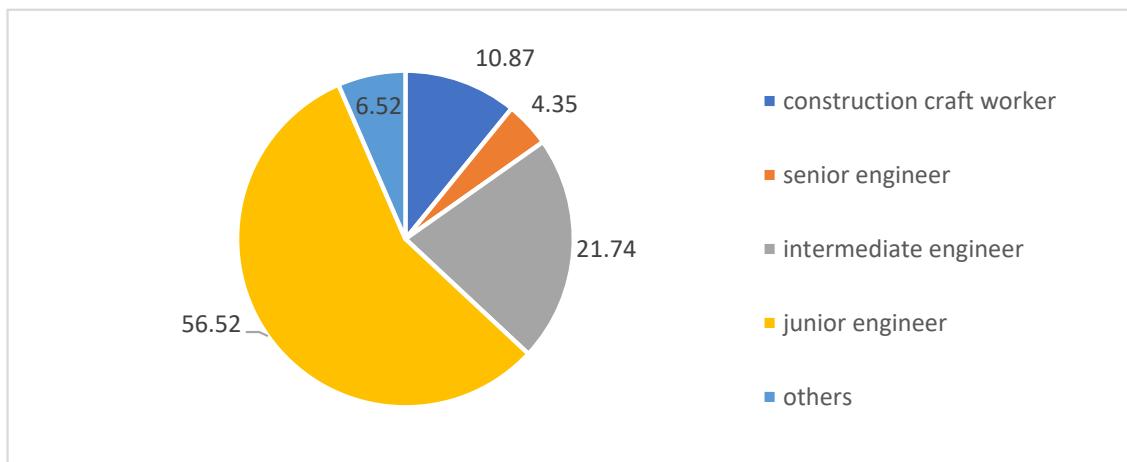


Figure 7. The job title distribution of the questionnaire's respondents.

First of all, this study invited 15 experts as a consultant group. These experts comprised professors, professor-level senior engineers, and senior engineers. The evaluation plan was responsible for the outcome of the evaluation. Meanwhile, the evaluation indexes were required to be correct and feasible. Therefore, for the determination of the evaluation indexes, most of the preliminary conclusions were reached through expert interviews and questionnaires.

Next, the series indexes are summarized in Table 8 below based on the literature review, expert interviews, and questionnaires.

Table 8. The preliminary evaluation indexes.

Indexes			
①	Certification	⑥	The level of the holding certificate
②	Professional engineering competence	⑦	Attitude
③	Creativity	⑧	Motivation
④	Practiced time	⑨	Qualification
⑤	Skills proficiency	⑩	Examination

Lastly, with several discussions and meetings with experts, as well as a site survey, a second round questionnaire was released to adjust and supplement the preliminary evaluation indexes, as shown in Table 9 below. During this procedure, experts firstly gave suggestions for the classification and leveling of indexes based on the feedback analysis. The updated indexes have been modified according to name and clustered according to the national engineering guidelines and standards. Secondly, the inclusion relation of the indexes was investigated during the site survey with superintendents, supervisors, and construction craft workers. The result of a periodical discussion with experts is shown in Table 10 below. According to the suggestions, there were intersecting indexes that should have been merged or deleted, and valuable indexes that were lacking in the system that should have been added. For example, Attitude, Motivation, Practice Time, Skills Proficiency, etc., are all summaries of personal comprehensive quality, which can be recombined into Personal Ability and summarized in Superior Evaluation or Peer Evaluation. This would not only increase the feasibility of the assessment method, but also eliminate repetition of the content. By mapping the indexes to the CSI system using

our literature review, questionnaire, site survey, and expert interview, the hierarchy of the performance evaluation index system was formed, resulting in Figure 8. A performance evaluation method and rating table with three roles were proposed to evaluate the work effectiveness of technicians, junior supervisors, and senior supervisors.

Table 9. The adjustment operation.

The Preliminary Evaluation Index		Adjustment	Supplement Item
①	Certification	①	
②	Professional engineering competence	⑥	Deleted
③	Creativity	③	
④	Practiced time	④	Merge into index Evaluation by significant others, and then divided into Attitude and Ability
⑤	Skills proficiency	⑤	
⑥	The level of the holding certificate	⑦	
⑦	Attitude	⑧	
⑧	Motivation	⑨	Merge into index Quality, and then divide into Inspection, Attendance, and Check.
⑨	Qualification	⑩	
⑩	Examination	⑩	Categorize into index Evaluation

Table 10. The evaluation indexes with fine-up adjustment.

Lv 1 Index		Lv2 Index	Lv3 Index		
Quantity	Evaluation	Notice	Messages	News	
		File Handling	The content is based on the different responsibilities of construction supervisors.		
Quality		Inspection	Log	Aside supervision	
		Attendance	Clocking-in	PE-fit	
Evaluation	Company	Response Time	The content is based on the different responsibilities of construction supervisors.		
		Check	Attitude	Ability	
Company		Superior Senior Construction Supervisor			
Examination			Based on the firm regulations		

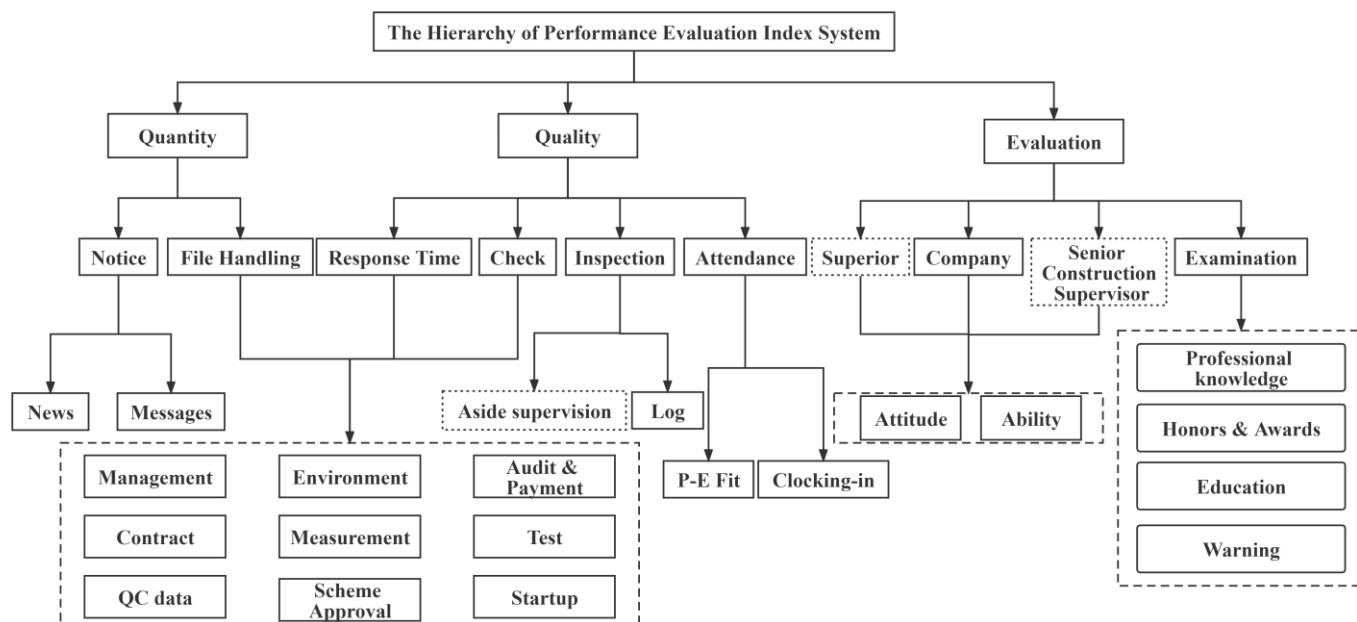


Figure 8. The hierarchy of the performance evaluation index system.

4.3. Weighting on Case Study

Using the statistics based on the two-time questionnaires (110 copies in total) and the summary based on several discussions with experts, in addition to the site survey, this approach used Quantity, Quality, and Evaluation as the first-level metrics. In addition, the quantity and quality of the work were shown to correlate with objectivity, and the evaluation was associated with subjectivity. Then, by constructing a judgment matrix, the scoring of each expert on the weight of the index in the system was determined first, and then the final weight of the index was determined according to the weighted geometric average method of the ranking vector of each expert. All the judgement matrices passed the consistency test. In this case, the index dimension was separated into subjectivity and objectivity, and the weight of each index is shown in Figure 8 above and Table 11 below.

Table 11. Weight analysis of the construction supervisor of each criterion.

Lv 1 Index	Weight	Lv 2 Index	Weight
Quantity	0.298	Notice	0.134
		File Handing	0.164
Quality	0.375	Attendance	0.086
		Response Time	0.1015
		Inspection	0.086
		Check	0.1015
Evaluation	0.327	Superior	0.109
		Company/Senior	0.109
		Construction Supervisor	
		Examination	0.109

For example, a judgement matrix was structured by Attendance, Response Time, Inspection, and Check, as in Table 12 below. Using the AHP method allowed us to obtain the results of key values such as eigenvector, weights, max-eigen, and CI, shown in Table 13 below. The result passed the consistency test, as shown in Table 14 below. This result proves that the items work in coordination, without inner conflict.

Table 12. An example of the AHP method.

AHP Data	Attendance	Response Time	Inspection	Check
Attendance	1.000	0.833	1.000	0.833
Response Time	1.200	1.000	1.200	1.000
Inspection	1.000	0.833	1.000	0.833
Check	1.200	1.000	1.200	1.000

Table 13. The result of the AHP method.

Items	Eigenvector	Weights	Max-Eigen	CI
Attendance	0.909	22.727%		
Response Time	1.091	27.273%		
Inspection	0.909	22.727%	4.000	0.000
Check	1.091	27.273%		

Table 14. The result of the consistency test.

Max-Eigen	CI	RI	CR	Test
4.000	0.000	0.890	0.000	Pass

Using the fuzzy comprehensive evaluation method validated the result obtained using the AHP method in Table 15. The result shows the same trend of distribution for the indexes.

Table 15. The result of the fuzzy comprehensive evaluation method.

	Response Time	Inspection	Check	Attendance
Membership	0.258	0.236	0.263	0.242
Weights	0.258	0.236	0.263	0.242

After a practitioner lost marks, the index affected the total score statistics within a certain period until the score loss ceased to occur within the influence range (period). Then, it was regarded as credit repair, that is, the statistical total score was no longer affected. According to the proposal, the influence of scope (period) to the relevant index was confirmed after discussions with senior practitioners in expert meetings, as shown in Table 16.

Table 16. Definition table of the evaluation index.

Lv 1 Index		Lv 2 Index		Scope (Month)		Lv 3 Index	
A	Quantity	A1	Notice	1	Receive notices	The quantity of processing files	
		A2	File Handling	6	The quantity of processing files		
B	Quality	B1	Response Time	6	The processing speed of the pending portion of the workflow	Identify problems Inspection log and aside supervision Clocking-in and P-E fit	
		B2	Check	6	Identify problems		
		B3	Inspection	6	Inspection log and aside supervision		
		B4	Attendance	6	Clocking-in and P-E fit		
C	Evaluation	C1	Superior	6	Evaluate the attitude and ability based on superior	Evaluate the attitude and ability based on company Evaluate the attitude and ability based on senior supervisor Examine professional knowledge and award outstanding behavior (bonus), or warnings (deduction)	
		C2	Company	6	Evaluate the attitude and ability based on company		
		C3	Senior Construction Supervisor	12	Evaluate the attitude and ability based on senior supervisor		
		C4	Examination	12	Examine professional knowledge and award outstanding behavior (bonus), or warnings (deduction)		

Weight analysis of the first-level indexes ranked them from high to low in terms of Quality (0.375), Evaluation (0.327), and Quantity (0.298). It can be seen that the weights of the first-level indicators were relatively average, which reflects that the above indexes did impact the work effectiveness evaluations of supervisors. Among them, the quality of work, which had quality as its first consideration, was more important than the other two first-level indexes.

There were 11 s-level indexes, and the top 3 weights were Attendance (0.086), File Handling (0.164) and Response Time (0.1015), indicating that the quality priority and the quantity priority were the most important in the evaluation of the work effectiveness of construction supervisors, reminding the superior management to pay more attention to it. In terms of Certification, Evaluation (including Superior Evaluation, Company Evaluation, and Senior Evaluation) was more important, suggesting that supervisors need to pay attention to improving their work attitudes, professional quality, and professional abilities.

In terms of the quantity priority, the weight of File Handling (0.164) was the highest under the Quantity index, indicating that the amount of online file processing based on the platform was significant and effective in expressing the amount of work carried out by supervisors. In terms of the quality priority, the weight of Attendance (0.086) under the index of quality was the highest, and compared with the weight of the overall secondary index, it was in the highest position, demonstrating the importance of the attendance system to the supervision work. At the same time, the same results were shown after expert selection and the questionnaire survey. Finally, under the Evaluation index, the sub-items related to evaluation work attitude, business ability, and collaboration office were classified and aggregated into Superior Evaluation and Company/Senior Supervisor Evaluation, according to the previous research.

As described in the previous overall process, this study initially established an evaluation index system for the work effectiveness of supervisors and calculated the weights. After the index system and scoring standards were sent to industry experts and senior practitioners for review and consultation, multiple rounds of adjustment and detail supplementation were carried out according to the suggestions of the participants, and the evaluation index system and weight interpretation table (Table 16 and Figure 9 below) were formed.

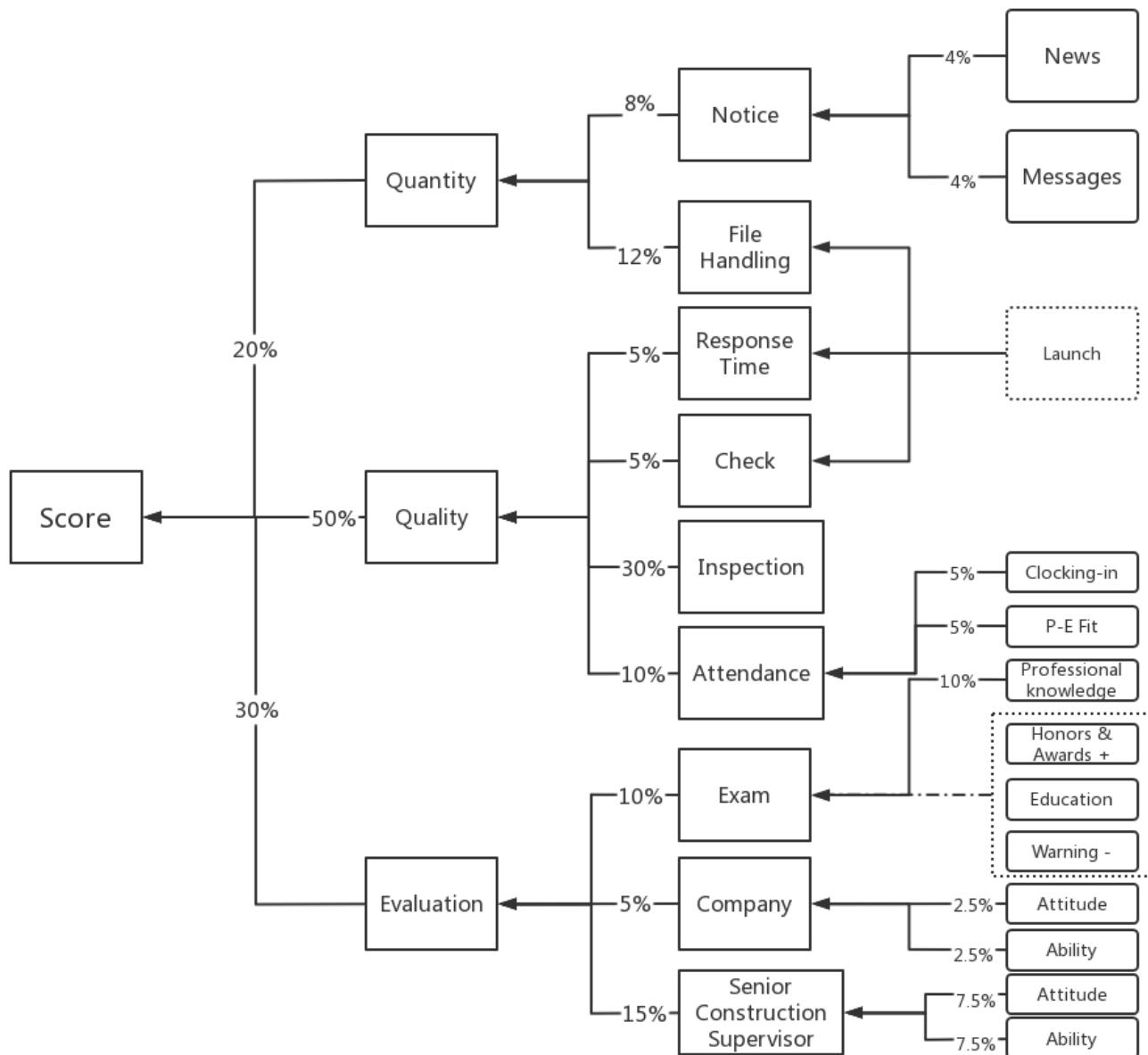


Figure 9. An example of an evaluation index system and weight interpretation table for a technician.

The indexes inside the dotted lines in Figure 9 were adjusted according to different evaluation objects (roles). For example, in the evaluation of the technicians, the relevant indexes in the Evaluation module were Junior Supervisor Evaluation and Company Evaluation and Examination, while Superior Evaluation was applicable to the Superior Evaluation of the assessment object below the senior supervisor engineers. In addition, the corresponding index weights were also adjusted according to the results of expert scores and inquiries.

4.4. Index Calculation on Case Study

The automatic calculation module applied the evaluation index system and weights to generate the performance evaluation report. This system was deployed in the server and also released as a mobile app. The data acquisition mainly used the mobile app to process information such as time, location, user authentication, site inspection with defect classification, reporting upload, etc. The CSI system received the information and data along with their weights, and then the calculation module automatically ran the logical algorithm based on the extracted data. Furthermore, the subjective evaluation indexes were supplemented by a professional exam, senior supervisor, superintendent, and the company, depending on the job title of the evaluation object.

The CSI system has constantly generated performance evaluation reports since the beta test was released in 2020. The beta test of the CSI system was initially released to the construction craft workers, construction supervisors, and superintendents. From 2021 to 2022, the functions of performance evaluation improved gradually. Through the aforementioned research and the proposed approach to data-driven quantitative performance evaluation indexes for construction supervisors, improvement to the performance evaluation system have been promoted and the process of job evaluation has been pushed in the direction of intelligentization.

According to the supervision performance evaluation index system and the corresponding proposed checklist, the function of performance evaluation was developed to be deployed in the CSI system, as shown in Figures 10–13, so as to give the quantitative indexes a stronger interpretation. These included:

1. Considering the fields and key values in the database, combined with the characteristics of the project, through literature research, site survey, and expert discussion, the indexes suitable for the performance evaluation of supervisors were supplemented and sorted out, and, finally, the evaluation system was formed. The performance-evaluating baseline was based on the national engineering guidelines and standards as well as local government policy. For the evaluating indexes, we considered suggestions from construction craft workers, construction supervisors, superintendents, and experts.
2. The objective evaluation index (work quantity and work quality) and the subjective evaluation index (work evaluation) modules were used to further divide the logical relationships in the data. After the evaluating procedure was performed during the discussion and meetings with experts, the automatic calculation module began to develop based on the indexes, weighting, daily work flow, and the data extracted from digital files.
3. Finally, the module was applied to the CSI system for scoring and grading. The verification of the project showed that the module was properly operational. The statistic of workload could be calculated automatically, and the statistical time range was selectable.

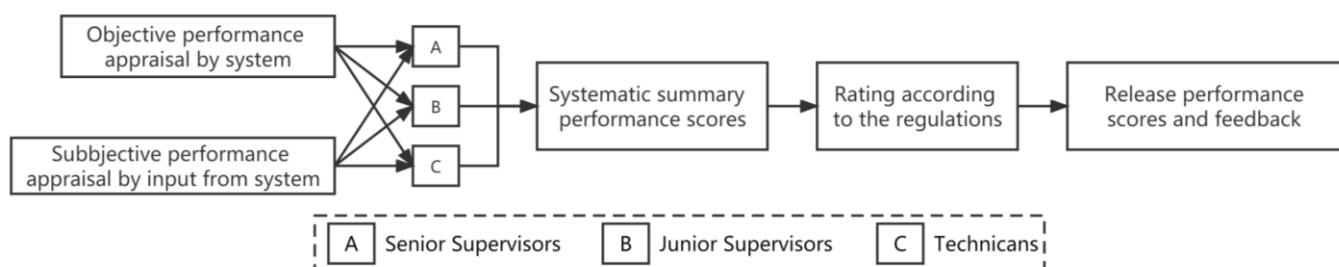


Figure 10. Simplified process diagram of the evaluation process.

Figure 11 shows the interactions of the CSI intelligent performance evaluation system of construction supervisors. The red box in the upper left selected the module of login. This module had two functions: (1) ID authentication and (2) data preparation. The CSI

system returned the content according to the ID authentication. For example, the junior supervisors had less of a right to operate the system compared to the superintendents. The red box in the upper right shows the other modules. The arrow pointing to the bottom right corner shows the interface of project management. The main functions are the menu bar, navigation bar, and notice. The last interface, in the bottom right corner, shows data visualization according to the selected button in the navigation bar.

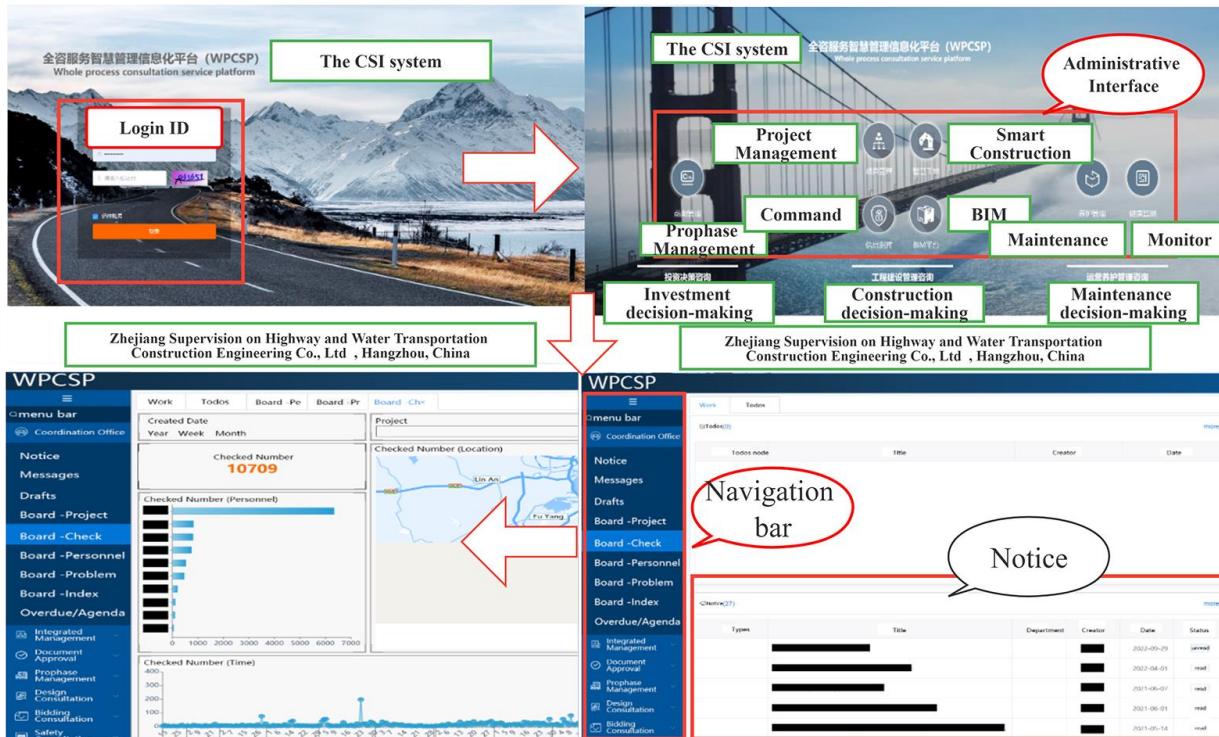


Figure 11. The interactions of an intelligent CSI performance evaluation system of construction supervisors.

The Personnel interface visualized the statistics of Attendance, Workload, and Work Briefing. As Figure 12 shows below, the menu bar was applied to select details such as a particular person, a specific time, or a work area. The CSI system automatically calculated the data, and then generated results as data-driven graphs.

The Issue interface visualized the statistics of construction management, work stats, workload, and trend risk. As Figure 13 shows below, the menu bar was applied to select details such as data stats, creator, table name, and date. The construction management on the left shows the statistics of launched tasks, such as Inspection and Check. The work stats in the middle provide information on the location and work area. The colored points on this map represent workers with different ID authentications. The closed ring on the right indicates the progress of a task. The statistics of monthly and daily workload are shown at the bottom. At the bottom right corner, a trend risk graph changes with the data statistic of the construction management module.

4.5. Application and Feedback

According to interviews with users, the positive experience of using this CSI system can be summarized. Firstly, the generation speed of the digital performance evaluation report from the CSI system was fast due to the data processing being close to a real-time state. The CSI system generated a large amount of data in the form of a continuous stream of events and stored it in a monitored state. Secondly, the operations were simple and straightforward, according to the feedback from construction craft workers and engineers. Current real-time streaming of data is conducted according to the “at least once” processing paradigm. In this paradigm, the last received event location is saved after the event is

processed. Thirdly, the CSI system was made stable by meetings with experts. It can be seen that the objectivity and authenticity of the data sources have an impact on the evaluation results, but this is strongly related to the data generated by the continuous event flow and filtered by the evaluation index system and evaluation table, shown in Table A1 in Appendix A. Moreover, superintendents highly praised the system's convenience and traceability. It was convenient to pull the original data that needed to be traced through the labels, marks, and form of index, and, finally, to locate the responsible party according to the data. In the CSI intelligent monitoring system, the evaluation system algorithm was formed for recording and settlement purposes. The validity, timeliness, security, and traceability of the data were well guaranteed.

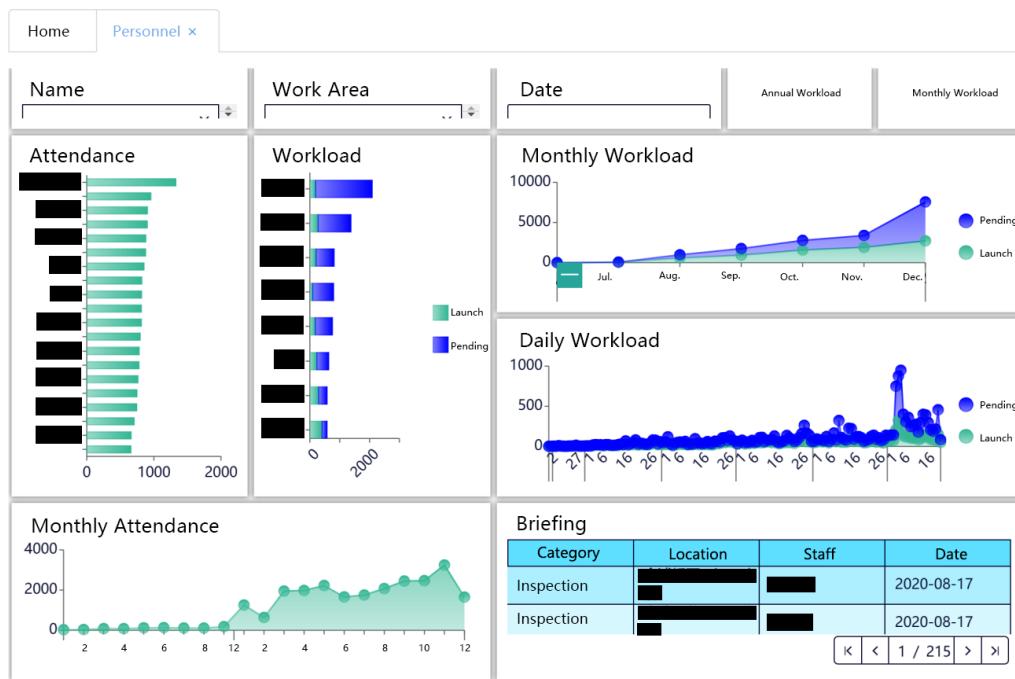


Figure 12. The application of an intelligent CSI performance evaluation system of construction supervisors: Personnel (Data Masking).

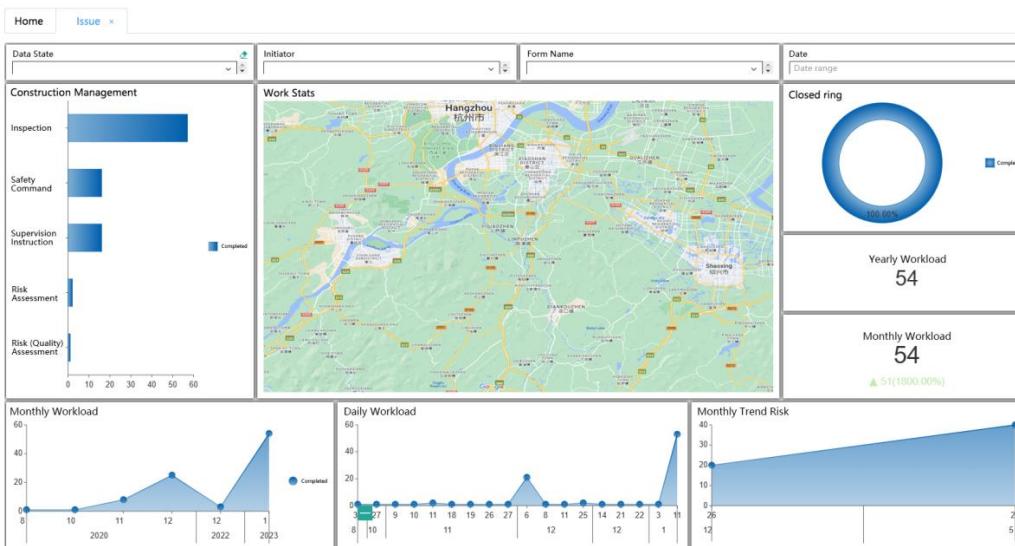


Figure 13. The application of an intelligent CSI performance evaluation system of construction supervisors: Issue.

5. Conclusions

This new approach to data-driven quantitative performance evaluation of construction supervisors by integrating AHP and automatic processing of daily activity tracking data was applied to the CSI system. The analytic hierarchy process (AHP) was used to calculate the weight of the evaluation indexes, and the reliability of the index, at all levels, was analyzed. The objective evaluation modules and the subjective evaluation module provided an overall measure of a supervisor's power of action. In terms of ensuring the authenticity and objectivity of data sources, data confirmation and hierarchical authorization were necessarily based on the job title. With the support of the CSI system, the work achieved traceability of the data and strengthened its abilities of safe construction and problem traceability. The data tracking used key tags in the database as indexes to quickly locate and pull data that needed to be validated. Therefore, the data processing procedure in this study was able to be recorded throughout the process, tracked throughout the period, and monitored throughout the path. Since the data were generated as a continuous stream of events, they were less susceptible to independent negative data.

This approach can systematically, automatically, and scientifically reflect the performance of supervisors. It covers the lack of data feedback and the subjective and emotional evaluation seen in previous performance evaluation efforts. The evaluation portion combines professional knowledge assessment, examination, and manual scoring, which not only retains the advantages of subjective evaluation, but also achieves scientific coordination in terms of weighting. However, data acquisition suffers from the complex surroundings of the construction site. Therefore, data acquisition and the data communication environment are at risk of interference. As a result, the result from the CSI system is affected by this interference.

In the future, the capture and recognition of various actions will be improved in terms of recognition rate and accuracy. These upgrades are helpful for intelligent intervention in the review of drawings, logs, and other links involving manual operations. It is worth expecting that new intelligent systems based on intelligent supervision technology will undertake and upgrade intelligent equipment in order to use data information more efficiently and accurately, and will completely change the operation mode of site construction. Since the intelligence of engineering supervision technology is based on the overall data of the engineering field process, it will also bring great challenges to the data itself. This includes data storage security, authority management, and other aspects.

Author Contributions: Conceptualization, J.-R.L.; methodology, K.-X.Y. and J.-R.L.; validation, Z.-Z.H. and Y.-C.D.; formal analysis, K.-X.Y., C.Y. and J.-R.L.; investigation, K.-X.Y. and J.-R.L.; resources, C.Y.; data curation, K.-X.Y. and C.L.; writing—original draft preparation, C.Y., K.-X.Y. and J.-R.L.; writing—review and editing, K.-X.Y., J.-R.L. and Y.-C.D.; visualization, K.-X.Y. and C.L.; supervision, J.-R.L., Z.-Z.H. and Y.-C.D.; project administration, J.-R.L., Z.-Z.H. and C.Y.; funding acquisition, J.-R.L. and C.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Plan Project of the Zhejiang Provincial Department of Transportation (grant number 2020061).

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Performance rating sheet for supervision: technician.

No.	Weighting	Lv.1	Weighting	Lv.2	Content	Period (Month)	Weighting	Lv.3	Evaluation Item	Label	Technician (Object)
1	20%	Quantity	8%	Notice	The number of notifications read: Monthly (Read/Total)	1	4.00%	News	Read pushed news from CSI system by APP or PC	non-daily	>90%
							4.00%	Messages	Read pushed messages from the CSI system by APP or PC	non-daily	
								Management	Acceptance records, etc.	day-to-day	
								Environment	Inspection records, etc.	day-to-day	
								Audit and Payment		day-to-day	
								Contract		non-daily	
			12%	File Handling	Number of tasks processed and launched	6	12.00%	Measurement	Approval form, etc.	day-to-day	>1
								Test		day-to-day	
								QC data	Review signature, etc.	day-to-day	
								Scheme Approval		non-daily	
								Startup		non-daily	
			5%	Response Time	The time of processing and handling tasks shall be in accordance with the contract	6	5.00%	Management	Acceptance records, etc.	day-to-day	
								Environment	Inspection records, etc.	day-to-day	
								Audit and Payment		day-to-day	
								Contract		non-daily	
								Measurement	Approval form, etc.	day-to-day	
								Test		day-to-day	
								QC data	Review signature, etc.	day-to-day	
								Scheme Approval		non-daily	
								Startup		non-daily	
2	50%	Quality	5%	Check	Number of unfinished tasks: (0, full mark; <10 cases, one point is deducted for every 10 cases; >30 cases, 0)	6	5.00%	Management	Acceptance records, etc.	day-to-day	
								Environment	Inspection records, etc.	day-to-day	
								Audit and Payment		day-to-day	
								Contract		non-daily	
								Measurement	Approval form, etc.	day-to-day	
								Test		day-to-day	
								QC data	Review signature, etc.	day-to-day	
								Scheme Approval		non-daily	
								Startup		non-daily	
			30%	Inspection	Patrol times and logging * Side station	6	10.00%	Log	Number of task flows that participate in patrol initiation	day-to-day	Inspect the construction site regularly or irregularly
							20.00%	Aside supervision	The number of side stations is subject to actual occurrence		Not less than once a day for major projects under construction

1. Tasks with the label of day-to-day:
processing time < 24 working hours
(one point is deducted if over than 12 h)

2. Tasks with the label of non-daily:
in accordance with the contract

Up until now, the special supervision inspection was qualified and had no accidents, full marks

The task was completed and qualified without accident, full marks

Number of unfinished cases:
(-0, full mark; <10 cases,
one point is deducted for every
10 cases; >30 cases, 0)

Table A1. *Cont.*

* Side station according to the actual work needs.

References

1. Sedykh, E. Project Management: Process Approach. *Vestn. Samara State Tech. Univ. Psychol. Pedagog. Sci.* **2019**, *16*, 181–192. [[CrossRef](#)]
2. Lee, M.-W.; Shin, E.-Y.; Lee, K.-S.; Park, H.-P. The Improvement Plan of Project Management for Highway Construction Supervisor Using Construction Management Benchmarking of Developed Countries. *Korean J. Constr. Eng. Manag.* **2006**, *7*, 138–148.
3. Zhang, J.-P.; Lin, J.-R.; Hu, Z.-Z.; Wang, H.-W. Intelligent Construction Driven By Digitalization. *Archit. Technol.* **2022**, *53*, 1566–1571.
4. Landy, F.J.; Barnes-Farrell, J.L.; Cleveland, J.N. Perceived Fairness and Accuracy of Performance Evaluation: A Follow-Up. *J. Appl. Psychol.* **1980**, *65*, 355–356. [[CrossRef](#)]
5. Cardy, R.L.; Dobbins, G.H. *Performance Appraisal: Alternative Perspectives*; South-Western Publishing Company: La Jolla, CA, USA, 1994; ISBN 978-0-538-81383-9.
6. Shahat, A. A Novel Big Data Analytics Framework for Smart Cities. *Future Gener. Comput. Syst.* **2018**, *91*, 620–633. [[CrossRef](#)]
7. Dransfield, R. *Human Resource Management*; Heinemann: London, UK, 2000; ISBN 978-0-435-33044-6.
8. Costa, N.; Rodrigues, N.; Seco, M.; Pereira, A. SL: A Reference Smartness Level Scale for Smart Artifacts. *Information* **2022**, *13*, 371. [[CrossRef](#)]
9. Arvey, R.D.; Murphy, K.R. Performance evaluation in work settings. *Annu. Rev. Psychol.* **1998**, *49*, 141–168. [[CrossRef](#)]
10. Banks, C.G.; Murphy, K.R. Toward narrowing the research-practice gap in performance appraisal. *Pers. Psychol.* **1985**, *38*, 335–345. [[CrossRef](#)]
11. Freinn-von Elverfeldt, A.C. Performance Appraisal: How to Improve Its Effectiveness. Available online: <http://essay.utwente.nl/58960/> (accessed on 26 December 2022).
12. Greenberg, J. *Advances in Organizational Justice*; Stanford University Press: Redwood City, CA, USA, 2002; ISBN 978-0-8047-6458-2.
13. Murphy, K.R.; Cleveland, J.N.; Skattebo, A.L.; Kinney, T.B. Raters Who Pursue Different Goals Give Different Ratings. *J. Appl. Psychol.* **2004**, *89*, 158–164. [[CrossRef](#)]
14. Perlow, R.; Latham, L.L. Relationship of Client Abuse With Locus of Control and Gender: A Longitudinal Study in Mental Retardation Facilities. *J. Appl. Psychol.* **1993**, *78*, 831. [[CrossRef](#)]
15. Waldman, D.A. The Contributions of Total Quality Management to a Theory of Work Performance. *Acad. Manage. Rev.* **1994**, *19*, 510–536. [[CrossRef](#)]
16. Bobko, P.; Coella, A. Employee Reactions to Performance Standards: A Review and Research Propositions. *Pers. Psychol.* **1994**, *47*, 1–29. [[CrossRef](#)]
17. Cardy, R.L.; Keefe, T.J. Observational Purpose and Evaluative Articulation in Frame-of-Reference Training: The Effects of Alternative Processing Modes on Rating Accuracy. *Organ. Behav. Hum. Decis. Process.* **1994**, *57*, 338–357. [[CrossRef](#)]
18. Borman, W.C.; Motowidlo, S.J. Task Performance and Contextual Performance: The Meaning for Personnel Selection Research. *Hum. Perform.* **1997**, *10*, 99–109. [[CrossRef](#)]
19. Murphy, K.R.; Cleveland, J.N. *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives*; SAGE: Newbury Park, CA, USA, 1995; ISBN 978-0-8039-5475-5.
20. Cascio, W.F.; Outtz, J.; Zedeck, S.; Goldstein, I.L. Statistical Implications of Six Methods of Test Score Use in Personnel Selection. *Hum. Perform.* **1995**, *8*, 133–164. [[CrossRef](#)]
21. Fletcher, C. Performance Appraisal and Management: The Developing Research Agenda. *J. Occup. Organ. Psychol.* **2001**, *74*, 473–487. [[CrossRef](#)]
22. Jackson, S.E.; Ruderman, M.N. Introduction: Perspectives for Understanding Diverse Work Teams. In *Diversity in Work Teams: Research Paradigms for a Changing Workplace*; Jackson, S.E., Ruderman, M.N., Eds.; American Psychological Association: Washington, DC, USA, 1995; pp. 1–13. ISBN 978-1-55798-333-6.
23. Ayman, R.; Chemers, M.M.; Fiedler, F. The Contingency Model of Leadership Effectiveness: Its Levels of Analysis. *Leadersh. Q.* **1995**, *6*, 147–167. [[CrossRef](#)]
24. Lewin, D.; Mitchell, O.S.; Sherer, P.D. *Research Frontiers in Industrial Relations and Human Resources*; Cornell University Press: Ithaca, NY, USA, 1992; ISBN 978-0-913447-53-6.
25. Lawler, E.E., III. *Strategic Pay: Aligning Organizational Strategies and Pay Systems*; Jossey-Bass/Wiley: Hoboken, NJ, USA, 1990; pp. xvii, 308. ISBN 978-1-55542-262-2.
26. Guzzo, R.A.; Dickson, M.W. Teams in organizations: Recent Research on Performance and Effectiveness. *Annu. Rev. Psychol.* **1996**, *47*, 307–338. [[CrossRef](#)]
27. Prendergast, C.; Topel, R.H. Favoritism in Organizations. *J. Polit. Econ.* **1996**, *104*, 958–978. [[CrossRef](#)]
28. DeNisi, A.S.; Robbins, T.L.; Summers, T.P. Organization, Processing, and Use of Performance Information: A Cognitive Role for Appraisal Instruments1. *J. Appl. Soc. Psychol.* **1997**, *27*, 1884–1905. [[CrossRef](#)]
29. Hoque, Z. Jodie Moll Public Sector Reform—Implications for Accounting, Accountability and Performance of State-Owned Entities—An Australian Perspective. *Int. J. Public Sect. Manag.* **2001**, *14*, 304–326. [[CrossRef](#)]

30. Kaplan, R.S.; Norton, D.P. Transforming the Balanced Scorecard from Performance Measurement to Strategic Management: Part I. *Account. Horiz.* **2001**, *15*, 87–104. [CrossRef]
31. Hagood, W.O.; Friedman, L. Using the Balanced Scorecard to Measure the Performance of Your HR Information System. *Public Pers. Manag.* **2002**, *31*, 543–557. [CrossRef]
32. Deem, J.W. *The Relationship of Organizational Culture to Balanced Scorecard Effectiveness*; Nova Southeastern University: Fort Lauderdale, FL, USA, 2009.
33. Salleh, M.; Amin, A.; Muda, S.; Halim, M.A.S.A. Fairness of Performance Appraisal and Organizational Commitment. *Asian Soc. Sci.* **2013**, *9*, 121. [CrossRef]
34. Batra, M.S.; Bhatia, D.A. Significance of Balanced Scorecard in Banking Sector: A Performance Measurement Tool. *Indian J. Commer. Manag. Stud.* **2014**, *5*, 47–54.
35. Neha, S.; Himanshu, R. Impact of Performance Appraisal on Organizational Commitment and Job Satisfaction. *Int. J. Eng. Manag. Sci.* **2015**, *6*, 95–104.
36. Azaria, A.; Ekblaw, A.; Vieira, T.; Lippman, A. MedRec: Using Blockchain for Medical Data Access and Permission Management. In Proceedings of the 2016 2nd International Conference on Open and Big Data (OBD), Vienna, Austria, 22–24 August 2016; pp. 25–30.
37. Uchenna, O.; Agu, A.G.; Uche, E.U. Performance Appraisal and Employee Commitment in Abia State Civil Service: A Focus on Ministries of Education and Works. *ABR* **2018**, *6*, 335–346. [CrossRef]
38. Huang, E.Y.; Paccagnan, D.; Mei, W.; Bullo, F. Assign and Appraise: Achieving Optimal Performance in Collaborative Teams. *IEEE Trans. Autom. Control* **2022**, *68*, 1614–1627. [CrossRef]
39. Guo, H.; Lin, J.-R.; Yu, Y. Intelligent and Computer Technologies’ Application in Construction. *Buildings* **2023**, *13*, 641. [CrossRef]
40. Modak, M.; Pathak, K.; Ghosh, K.K. Performance Evaluation of Outsourcing Decision Using a BSC and Fuzzy AHP Approach: A Case of the Indian Coal Mining Organization. *Resour. Policy* **2017**, *52*, 181–191. [CrossRef]
41. Lin, J.-R.; Hu, Z.-Z.; Li, J.-L.; Chen, L.-M. Understanding On-Site Inspection of Construction Projects Based on Keyword Extraction and Topic Modeling. *IEEE Access* **2020**, *8*, 198503–198517. [CrossRef]
42. Erkan, T.; Erdebili, B. Selection of Academic Staff Using the Fuzzy Analytic Hierarchy Process (FAHP): A Pilot Study. *Teh. Vjesn.* **2012**, *19*, 923–929.
43. Tzeng, G.-H.; Huang, J.-J. *Multiple Attribute Decision Making: Methods and Applications*; CRC Press: Boca Raton, FL, USA, 2011; ISBN 978-1-4398-6157-8.
44. Hongxing, L.I. Multifactorial Functions in Fuzzy Sets Theory. *Fuzzy Sets Syst.* **1990**, *35*, 69–84. [CrossRef]
45. Wei, Q. Data Envelopment Analysis. *Chin. Sci. Bull.* **2001**, *46*, 1321–1332. [CrossRef]
46. Ogbu Edeh, F.; Nonyelum Ugwu, J.; Gabriela Udeze, C.; Chibuike, O.N.; Onya Ogwu, V. Understanding Performance Management, Performance Appraisal and Performance Measurement. *Am. J. Econ. Bus. Manag.* **2019**, *2*, 129–146.
47. Bracken, D.W.; Timmreck, C.W.; Church, A.H. *The Handbook of Multisource Feedback*; John Wiley & Sons: Hoboken, NJ, USA, 2001; ISBN 978-0-7879-5856-5.
48. Aman, M.G.; Singh, N.N.; Stewart, A.W.; Field, C.J. The Aberrant Behavior Checklist: A Behavior Rating Scale for the Assessment of Treatment Effects. *Am. J. Ment. Defic.* **1985**, *89*, 485–491.
49. Adekunle, D.S. *The Philosophy and Practice of Management by Objectives*; Troy State University: Troy, AL, USA, 2005.
50. Aggarwal, A.; Mitra Thakur, G. Techniques of Performance Appraisal-A Review. *Int. J. Eng. Adv. Technol.* **2013**, *2*, 617–621.
51. Jafari, M.; Bourouni, A.; Hesamamiri, R. A New Framework for Selection of the Best Performance Appraisal Method. *Eur. J. Soc. Sci.* **2009**, *7*, 92–100.
52. Pan, W.; Wei, H. Research on Key Performance Indicator (KPI) of Business Process. In Proceedings of the 2012 Second International Conference on Business Computing and Global Informatization, Shanghai, China, 12–14 October 2012; pp. 151–154.
53. Johanson, U.; Eklöv, G.; Holmgren, M.; Mårtensson, M. *Human Resource Costing and Accounting versus the Balanced Scorecard: A Literature Survey of Experience with the Concepts*; School of Business Stockholm University: Stockholm, Sweden, 1998.
54. Phanden, R.K.; Aditya, S.V.; Sheokand, A.; Goyal, K.K.; Gahlot, P.; Jacso, A. A State-of-the-Art Review on Implementation of Digital Twin in Additive Manufacturing to Monitor and Control Parts Quality. *Mater. Today Proc.* **2022**, *56*, 88–93. [CrossRef]
55. Zheng, Z.; Lu, X.-Z.; Chen, K.-Y.; Zhou, Y.-C.; Lin, J.-R. Pretrained Domain-Specific Language Model for Natural Language Processing Tasks in the AEC Domain. *Comput. Ind.* **2022**, *142*, 103733. [CrossRef]
56. Meng, Q.; Zhang, Y.; Li, Z.; Shi, W.; Wang, J.; Sun, Y.; Xu, L.; Wang, X. A Review of Integrated Applications of BIM and Related Technologies in Whole Building Life Cycle. *Eng. Constr. Archit. Manag.* **2020**, *27*, 1647–1677. [CrossRef]
57. Lin, C.; Hu, Z.Z.; Yang, C.; Deng, Y.C.; Zheng, W.; Lin, J.R. Maturity Assessment of Intelligent Construction Management. *Buildings* **2022**, *10*, 1742. [CrossRef]
58. Neely, A. *Business Performance Measurement: Theory and Practice*; Cambridge University Press: Cambridge, UK, 2002; ISBN 978-0-521-80342-7.
59. Deng, H.; Xu, Y.; Deng, Y.; Lin, J. Transforming Knowledge Management in the Construction Industry through Information and Communications Technology: A 15-Year Review. *Autom. Constr.* **2022**, *142*, 104530. [CrossRef]

60. Zheng, Z.; Zhou, Y.-C.; Lu, X.-Z.; Lin, J.-R. Knowledge-Informed Semantic Alignment and Rule Interpretation for Automated Compliance Checking. *Autom. Constr.* **2022**, *142*, 104524. [[CrossRef](#)]
61. Jiang, C.; Li, X.; Lin, J.-R.; Liu, M.; Ma, Z. Adaptive Control of Resource Flow to Optimize Construction Work and Cash Flow via Online Deep Reinforcement Learning. *Autom. Constr.* **2023**, *150*, 104817. [[CrossRef](#)]
62. Saaty, T.L. How to Make a Decision: The Analytic Hierarchy Process. *Eur. J. Oper. Res.* **1990**, *48*, 9–26. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.