



OPEN ACCESS

EDITED BY

María Teresa Becerra Traver,
Universidad de Extremadura, Spain

REVIEWED BY

Jorge Bonito,
University of Evora, Portugal
Etienne Twizeyimana,
University of Rwanda College of
Education, Rwanda

*CORRESPONDENCE

Gulmira Yeshimbetova
✉ gulmira-eshim@mail.ru

RECEIVED 25 January 2026

REVISED 03 May 2026

ACCEPTED 12 May 2026

PUBLISHED 28 May 2026

CITATION

Yeshimbetova G, Baitasheva G, Koksai EA
and Kozhabekova E (2026) A qualimetric
evaluation of botany field practice:
assessing the alignment between
instructional materials and student
learning outcomes.
Front. Educ. 11:1795880.
doi: 10.3389/feduc.2026.1795880

COPYRIGHT

© 2026 Yeshimbetova, Baitasheva,
Koksai and Kozhabekova. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

A qualimetric evaluation of botany field practice: assessing the alignment between instructional materials and student learning outcomes

Gulmira Yeshimbetova^{1*}, Gaukhar Baitasheva¹, Ela Ayse Koksai²
and Elmira Kozhabekova³

¹Kazakh National Women's Teacher Training University, Almaty, Kazakhstan, ²Nigde Omer Halisdemir
Universitesi, Niğde, Türkiye, ³Uzbekali Zhanibekov South Kazakhstan Pedagogical University, Shymkent,
Kazakhstan

Field practice is a central component of science education; however, its quality is rarely evaluated using integrated, multidimensional frameworks. This study applied a qualimetric approach, defined as a weighted, multi-criteria evaluation method, to assess the quality of botany field-practice learning materials and students' field-practice outcomes. Data were collected from undergraduate students ($n = 387$) using a structured survey and an evaluation rubric covering seven quality dimensions. Dimension-level scores were normalized and aggregated into integral quality indices. The results indicated high overall quality for instructional materials ($KQ = 0.726$) and student outcomes ($KQ = 0.706$), with a strong alignment between instructional inputs and learning outputs. However, student performance showed greater variability across dimensions, with relatively lower scores in inquiry-based fieldwork skills and the application of theoretical knowledge. These findings show that while well-structured materials are associated with effective field-practice implementation, higher-order competencies may require additional pedagogical support. The study contributes to the literature by introducing a systematic qualimetric framework for evaluating the alignment between instructional design and student performance in field-based learning. A key limitation is that the study was conducted within a single institutional context, which may limit generalizability.

KEYWORDS

botany education, educational quality assessment, field practice, inquiry-based learning, qualimetric evaluation, qualimetry, science education

Introduction

Field-based learning experiences, such as field practice and excursions, play a main role in biology and other natural science education by helping students to engage directly with ecological systems, observe phenomena in authentic contexts, and apply classroom knowledge to real-world settings. Research indicates that participation in field courses increase undergraduate scientific knowledge, persistence in STEM, and motivation through immersive, experiential activities that cannot be fully replicated in traditional classroom environments (Fogelman et al., 2025; Hole et al., 2018; Mutseekwa, 2024; Shinbrot et al., 2022). Despite the documented pedagogical value of

field activities, empirical research on higher education field-based learning remains limited, and systematic evaluation methods that bridge instructional design and student performance outcomes are not well studied (Fedesco et al., 2020).

Although field experiences are widely acknowledged as beneficial, the evaluation of these experiences tends to be disjointed, focusing on isolated dimensions or descriptive accounts rather than complete, multidimensional assessment (Fedesco et al., 2020; Seyferth-Zapf et al., 2024). This presents a hole in the literature: educational quality evaluations often overlook the need for integrated frameworks that can jointly assess the quality of educational materials and the quality of student learning products. Addressing this hole is particularly important in field practice contexts, where the alignment between learning supports (e.g., learning materials) and student enactment (e.g., field agendas, herbarium preparation, reports) is main to meaningful outcomes.

Unlike analytic rubrics and authentic assessment approaches, which typically evaluate performance across multiple criteria but report results descriptively or as separate scores, qualimetry provides a structured, quantitative framework that integrates multiple dimensions into a single composite index through normalization, weighting, and aggregation procedures (Azgaldov and Kostin, 2011; Nardo et al., 2005). While analytic rubrics emphasize criterion-referenced judgment and authentic assessment focuses on real-world performance tasks, these approaches often lack systematic methods for combining dimensions into an overall measure of quality (Brookhart, 2013; Wiggins, 1998). In contrast, qualimetry explicitly models the relative importance of each dimension through weighting schemes and produces integral quality indices, enabling transparent, comparable, and multidimensional evaluation of complex educational processes.

The qualimetric approach is theoretically aligned with experiential and inquiry-based learning frameworks, as both emphasize the development of multidimensional competencies through active engagement, reflection, and evidence-based reasoning (Hmelo-Silver et al., 2007; Kolb, 2015). Field practice involves interconnected processes, such as observation, investigation, application, and reflection, that cannot be adequately captured through single indicators or isolated criteria. Qualimetry provides a more suitable evaluative framework than analytic rubrics or authentic assessment because it systematically integrates these interrelated dimensions through normalization, weighting, and aggregation into composite indices, thereby reflecting the holistic and process-oriented nature of experiential and inquiry-based learning (Azgaldov and Kostin, 2011; Nardo et al., 2005). This makes it particularly appropriate for evaluating complex educational contexts where both learning processes and outcomes must be assessed simultaneously.

Recent studies have consistently demonstrated the benefits of field-based and inquiry-oriented learning for improving students' engagement, scientific reasoning, and skill development (Fadillah et al., 2025; Shinbrot et al., 2022). However, most of this research evaluates isolated outcomes (e.g., achievement or attitudes) rather than providing integrated assessments that link instructional design with student performance. While multidimensional assessment approaches have been emphasized in recent literature, they are typically implemented through analytic or descriptive methods that lack systematic aggregation

into composite indices (Guàrdia et al., 2023; Harrison et al., 2022). Therefore, a key gap remains in the development of structured, transparent evaluation frameworks that can simultaneously assess multiple quality dimensions and produce integrated measures of educational effectiveness.

Research questions

RQ1. What is the overall quality level of field-practice learning materials when evaluated using a multidimensional qualimetric framework?

RQ2. What is the overall quality level of students' field-practice outcomes, as reflected in field practice agendas, herbarium products, and written reports, based on qualimetric evaluation?

RQ3. How do quality levels vary across different dimensions of field-practice learning materials and student field-practice outcomes?

RQ4. Does the qualimetric model provide valid and consistent evaluation across dimensions?

Literature review

Field practice in biology and higher education

Field practice is a foundational component of biology and other natural science programs, as it helps students to engage with phenomena in real contexts and to connect abstract theoretical concepts with real-world observations. Research shows that field-based courses support cognitive, affective, and behavioral outcomes, including content knowledge development, scientific skill gaining, and student persistence in STEM disciplines (Shinbrot et al., 2022). Early and inclusive field experiences, in particular, have been shown to strengthen students' sense of belonging and commitment to scientific study, showing the broader educational value of structured field practice beyond content mastery (Race et al., 2021).

A main pedagogical function of field-based learning is the application of theoretical knowledge through inquiry-oriented activities. Field practice provides opportunities for students to observe natural systems, collect data, formulate questions, and test explanations under real environmental conditions. Studies of field excursions show that learning in the field involves active sense-making processes, where observation, interpretation, and discussion are firmly intertwined (Hole et al., 2018). Through such processes, students develop main scientific practices, including observation, comparison, classification, and evidence-based reasoning, which are difficult to do fully in classroom or laboratory settings.

Field practice in biology typically includes a set of structured components designed to support these learning processes. These components often include clearly defined field tasks, systematic data collection activities, preparation of written field reports, and, in botany-focused contexts, the collection and preparation of plant specimens for herbarium submission. Herbarium work requires students to apply taxonomic knowledge, correctly

document morphological features, and follow standardized scientific procedures for specimen preservation and labeling (Florida Museum, 2023). Together, these elements, field tasks, agendas, reports, and specimen preparation, constitute the main products of field practice and serve as key indicators of students' engagement with both theoretical and practical dimensions of biological learning (Fedesco et al., 2020).

Quality of field-practice learning materials

Instructional materials play an important role in guiding students' engagement during field practice by structuring activities, clarifying expectations, and supporting the translation of learning objectives into concrete actions. In field-based contexts, where learning occurs outside controlled classroom environments, the quality of learning materials becomes particularly important for ensuring consistency, coherence, and purposeful engagement with disciplinary content. Research on field-based learning in higher education stresses that well-designed materials help students navigate complex tasks, manage time and resources effectively, and remain focused on intended learning outcomes during field activities (Fedesco et al., 2020).

A key characteristic of high-quality field-practice learning materials is the alignment among learning objectives, field tasks, assessment criteria, and safety requirements. The principle of constructive alignment shows that learning outcomes should be explicitly articulated and systematically connected to instructional activities and assessment practices in order to increase meaningful learning (Biggs, 2014). In field practice, such alignment is essential for cognitive learning and also for procedural and safety-related aspects, as students must understand both what they are expected to learn and how to conduct fieldwork responsibly. Authoritative guidance on science education show that instructional resources shape how learning goals are enacted in practice, influencing both the depth and consistency of implementation (National Academies of Sciences, Engineering, and Medicine [NASEM], 2022).

Empirical research additionally indicates that the quality of instructional materials affects how learning activities are implemented and, consequently, how students engage with learning tasks. Studies examining curriculum materials in science education have shown that materials which clearly articulate objectives, provide coherent task sequences, and contain explicit assessment guidance are more likely to be done as intended and to support student learning (Taylor et al., 2007). In field-practice settings, materials generally include task descriptions, field agendas, data collection guidelines, reporting requirements, and, in biology-focused contexts, instructions for specimen handling and documentation. Clear guidance on these components (including safety and risk considerations) is widely recognized as an essential element of responsible fieldwork design (University of Western Ontario, 2023).

Evaluation of student field-practice outcomes

Student performance in field practice is commonly evaluated through a combination of written reports, reflective accounts,

and documentation of practical activities conducted in authentic learning environments. In experiential and field-based education, such products allow instructors to assess students' engagement with real-world tasks, their ability to connect theoretical knowledge with practical investigation, and their development of scientific reasoning (Moon, 2004; Oxley and Turner, 2022; Sanjito and Hyams-Ssekasi, 2025). In biology field practice specifically, assessment often includes field agendas, observational records, specimen documentation, herbarium preparation, and final written reports, all of which help as evidence of students' conceptual understanding and practical competence.

A key limitation of prior research on field-based learning lies in its reliance on fragmented and single-dimension evaluation approaches. Many studies assess outcomes such as achievement, engagement, or attitudes in isolation, without capturing the interconnected nature of learning processes and instructional design (Fedesco et al., 2020; Shinbrot et al., 2022). Even when multiple dimensions are considered, they are often reported separately through descriptive statistics or rubric scores, limiting the ability to generate integrated conclusions about overall quality (Guàrdia et al., 2023; Jonsson and Svingby, 2007). This lack of methodological integration reduces comparability across studies and constrains the development of comprehensive evaluation frameworks capable of reflecting the complexity of field-based educational practices.

Similarly, single-indicator or checklist-based approaches are unsatisfactory for assessing the complex and multidimensional nature of field-practice performance. Checklists typically focus on task completion rather than quality, reducing complex performances to binary judgments that fail to capture depth, coherence, or scientific rigor (Brookhart, 2013). Theoretical work on performance assessment additionally argues that reducing learning outcomes to a single score or limited indicators undermines construct validity and ignores the interplay of multiple competencies involved in authentic tasks (Guàrdia et al., 2023).

Consequently, there is need for structured, multi-criteria assessment frameworks to evaluate student field-practice outcomes. Authentic assessment literature emphasizes that complex products, such as scientific reports, inquiry records, herbarium specimens, and reflective analyses, should be evaluated across multiple dimensions that represent both process and outcome quality (Wiggins, 1998). Frameworks for authentic and inquiry-based assessment similarly show the importance of assessing observation, analysis, documentation, and reflection as interrelated components of scientific practice (Bonney et al., 2024; National Research Council, 2012).

Multidimensional quality assessment in education

Educational quality is widely recognized as a multidimensional construct that cannot be sufficiently captured by single indicators or isolated outcomes. Early conceptual and empirical work has emphasized that quality in education encompasses multiple interrelated dimensions, such as teaching processes, learning outcomes, contextual relevance, and

institutional conditions, rather than a single universally defined attribute (Harrison et al., 2022). Contemporary perspectives similarly view educational quality as emerging from the interaction of learning goals, instructional processes, learning environments, and learner outcomes, showing the need for evaluation frameworks that can address this complexity in an integrated manner (OECD, 2017).

In response to this complexity, merged indicators have increasingly been used in educational and social evaluation to synthesize information from multiple dimensions into a single, interpretable index. Composite indicators are especially useful when evaluation objects involve diverse criteria that must be considered simultaneously, as they enable systematic aggregation while preserving transparency regarding how individual components contribute to total quality (Nardo et al., 2005). In educational evaluation, such approaches allow researchers and decision makers to move beyond fragmented analyses and to generate complete measures that support comparison, monitoring, and evidence-based improvement (Saisana and Tarantola, 2002).

At the same time, the integration of multiple criteria presents methodological challenges that require careful attention. Scholars have noted that composite indicators may obscure important nuances if weighting schemes, normalization procedures, or aggregation rules are not clearly defined and theoretically justified (Saltelli, 2007). Validity concerns also arise when complex performances are reduced to simplified summary measures without sufficient consideration of the underlying constructs and their relationships. Competence-based education, for example, requires assessment approaches that capture multiple dimensions of performance rather than relying on single indicators (Guàrdia et al., 2023).

Qualimetry as an evaluation framework and its applications in education

Qualimetry initiated as a scientific approach to the quantitative evaluation of quality, grounded in the assumption that complex objects and processes can be systematically assessed through structured criteria and measurable indicators. Early work in qualimetry conceptualized quality as a multidimensional construct that requires explicit definition of indicators, normalization of measurements, and aggregation into integral indices in order to support objective comparison and decision-making. Subsequent developments clarified main qualimetric principles and addressed common methodological errors, emphasizing transparency in indicator selection, weighting procedures, and interpretation of composite results (Azgaldov and Kostin, 2011).

A defining feature of qualimetry is its structured sequence of evaluation steps. These include the identification of relevant quality indicators and criteria, transformation of raw measurements through normalization to dimensionless scales, assignment of weights to reflect the relative importance of criteria, and calculation of integral quality indices that summarize total performance. These procedures closely align with established methodologies for constructing composite indicators, which have been widely applied in social and

educational evaluation when multiple dimensions must be combined into a single interpretable metric (Nardo et al., 2005; Saisana and Tarantola, 2002). By making each step explicit, qualimetry addresses key challenges associated with multidimensional assessment, such as scale incomparability and implicit weighting.

Qualimetry differs from commonly used evaluation frameworks such as authentic assessment and evidence-based evaluation in both purpose and methodological structure. Authentic assessment emphasizes real-world tasks and holistic judgment of student performance but typically relies on qualitative or rubric-based scoring without systematic aggregation across dimensions (Gulikers et al., 2004; Wiggins, 1998). Evidence-based evaluation frameworks focus on the use of empirical data to inform instructional decisions, often emphasizing validity and reliability, yet they frequently assess outcomes through separate indicators rather than integrating them into a unified measure (OECD, 2017; Slavin, 2020). In contrast, qualimetry provides a structured quantitative approach that combines multiple criteria through normalization, weighting, and aggregation into composite indices, enabling transparent and comparable evaluation of complex educational processes (Azgaldov and Kostin, 2011; Nardo et al., 2005).

Within education, qualimetric and composite-index approaches have been increasingly used to evaluate curricula, teaching quality, and learning outcomes. Conceptualizations of educational quality as inherently multidimensional have provided a strong foundation for such applications, showing the need to assess learning processes, instructional design, and outcomes together (Harrison et al., 2022). Large-scale educational frameworks and policy-oriented evaluations similarly rely on composite indicators to synthesize evidence across multiple dimensions of teaching and learning environments (OECD, 2017). At the classroom and program levels, multidimensional assessment frameworks have been applied to authentic learning tasks and learning outcomes, reinforcing the relevance of structured, criteria-based evaluation for complex educational products (Kitts and Perry 2023; National Research Council, 2012).

Despite its advantages, the qualimetric approach has several limitations that should be acknowledged. One key concern is the potential subjectivity in assigning weights to different dimensions, which may influence the resulting composite index and introduce evaluator bias (Nardo et al., 2005; Saltelli, 2007). Additionally, the aggregation of multiple indicators into a single index can lead to a loss of nuance, as important variations across individual dimensions may be obscured in the overall score. Scholars have also noted that composite indicators may oversimplify complex educational constructs if the underlying assumptions, normalization procedures, and weighting schemes are not sufficiently transparent and theoretically justified (Saisana and Tarantola, 2002; Saltelli, 2007).

Thus, the reviewed literature suggests that field-practice quality should be understood as the alignment between instructional design, learning processes, and student outputs. Experiential and inquiry-based learning explain why students should engage in observation, investigation, application, and reflection, while qualimetry offers a method for evaluating these dimensions systematically through normalized, weighted, and aggregated

indicators (Azgaldov and Kostin, 2011; Hmelo-Silver et al., 2007; Kolb, 2015). Therefore, the present study conceptualizes botany field-practice quality as a multidimensional construct consisting of instructional-material quality and student-performance quality. This framework directly informs the research questions by examining the overall quality of each component, the variation across dimensions, and the validity of using a qualimetric model to evaluate field-based learning.

To synthesize the literature and establish a coherent link between theoretical foundations, learning processes, and evaluation, a multidimensional conceptual framework is proposed. The model integrates experiential learning and inquiry-based learning with a qualimetric evaluation approach to systematically assess the quality of field-practice instruction and student outcomes.

Figure 1 presents a five-stage model illustrating how field-practice quality is constructed and evaluated. The first stage (Input) represents instructional design features, including content coverage, alignment, and inquiry-oriented tasks. The second stage (Mediator) captures learning processes grounded in experiential and inquiry-based learning, such as observation, investigation, and reflection. The third stage (Analytical Layer) introduces the qualimetric framework, where multiple dimensions of quality are evaluated through normalization, weighting, and aggregation to produce composite indices. This stage also acknowledges methodological limitations, including subjectivity in weighting and potential loss of nuance. The fourth stage (Proximal Outcomes) presents immediate evaluation results, including quality indices and dimension-level

variations. The fifth stage (Distal Outcomes) reflects broader educational interpretations, such as alignment between instructional inputs and student outputs and the development of higher-order skills. Moderating factors (e.g., learner, task, and contextual variables) and a feedback loop are incorporated to highlight boundary conditions and the iterative improvement of instructional design. The framework directly informs the study's research questions by linking theoretical, methodological, and evaluative components.

Methods

This study employed a quantitative evaluative research design based on a qualimetric framework to systematically assess the quality of botany field-practice learning materials and students' field-practice outcomes. The study involved undergraduate students enrolled in a botany field practice course at the Kazakh National Women's Teacher Training University. A total of 387 students participated, primarily first- and second-year students, and the field practice was conducted in natural botanical environments including forested areas and botanical gardens.

Participants

The participants of the study were undergraduate students enrolled in the botany field practice course at the Kazakh National Women's Teacher Training University. Using a

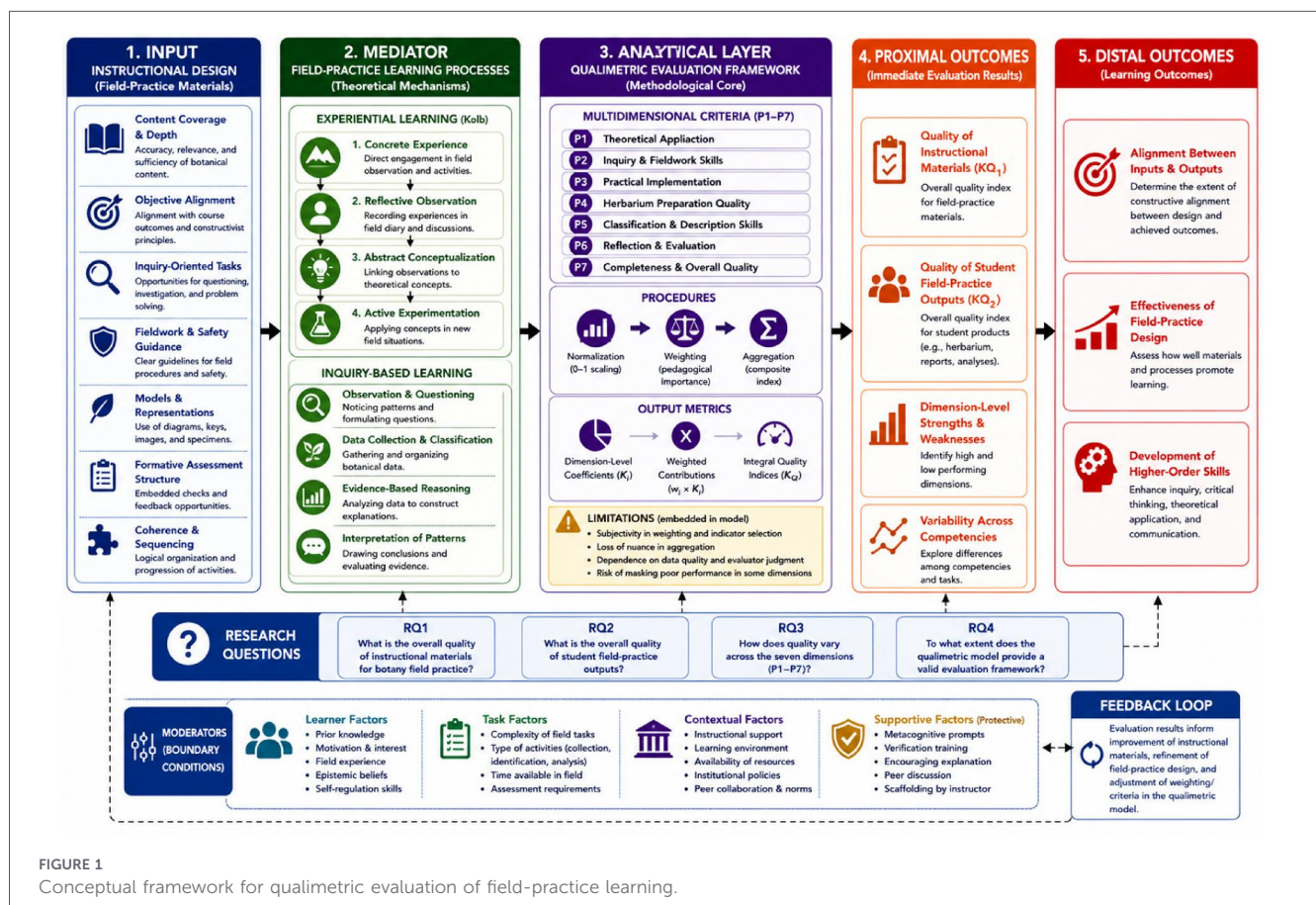


FIGURE 1
Conceptual framework for qualimetric evaluation of field-practice learning.

convenient sampling method a total of 387 students participated in the evaluation of the field-practice learning materials. Participation was aligned with students' mandatory involvement in the field practice and followed institutional and ethical guidelines. The age of the participants ranged from 17 to 28 years, with a mean age of approximately 18.5 years, showing that the sample primarily consisted of early undergraduate students. In terms of gender distribution, the sample was mostly female ($n = 329$; 85.0%), reflecting the institutional profile of the university, while 58 participants (15.0%) identified as male. Regarding academic standing, most participants were in the first and second years of study, with 133 students (34.4%) enrolled in the first year and 162 students (41.9%) in the second year. A smaller proportion of students represented upper academic levels, including third- and fourth-year students, who together accounted for less than one quarter of the sample. Among these students, the field practice agenda of 25 of them were randomly selected and qualitatively evaluated.

Instruments

Two structured instruments were used to collect data for the qualimetric evaluation: (a) a survey assessing the quality of field-practice learning materials ([Supplementary Appendix A](#)) and (b) an evaluation rubric assessing students' field-practice outcomes ([Supplementary Appendix B](#)). Both instruments were explicitly structured to support qualimetric evaluation. Items were grouped into theoretically grounded dimensions, responses were collected on uniform ordinal scales, and dimension-level scores were calculated to function as elementary quality indicators. This design ensured compatibility with subsequent normalization, weighting, and aggregation procedures used to calculate integral quality indices.

Evaluation of field-practice learning materials survey

This survey was developed to assess the perceived quality and effectiveness of instructional materials used in biology field practice. The instrument (see [Supplementary Appendix A](#)) consisted of 22 items organized into seven dimensions: (P1) Content Coverage and Depth, (P2) Objective Alignment, (P3) Development of Scientific and Inquiry Skills, (P4) Implementation of Field/Laboratory Work and Safety, (P5) Use of Models and Representations, (P6) Formative Assessment, and (P7) Coherence and Sequencing. Each dimension counted between three and four items pointing out key aspects of material quality relevant to field-based learning.

Items were rated on a four-point Likert-type scale ranging from 1 (Strongly disagree) to 4 (Strongly agree). The use of a four-point scale was intended to reduce neutral responding and to encourage participants to make evaluative judgments. Dimension scores were calculated by averaging the item responses within each dimension, providing dimension-level indicators for subsequent qualimetric normalization and aggregation.

This survey was developed by the authors concurrently with the second survey. Feedback from two experts led to minor

revisions. The survey was not pilot tested with students, as it was used solely by the authors of the study. The survey was administered to biology students who involved in field practice, and the collected responses showed perceptions of how well the learning materials supported field-practice objectives, task implementation, safety, skill development, and assessment processes.

Field practice agenda analysis survey

This survey ([Supplementary Appendix B](#)) was designed as an evaluator-based instrument to assess the quality of students' field-practice outcomes, including field agendas, herbarium specimens, and written reports. The instrument comprised 28 items, grouped into same seven dimensions of the Evaluation of Field-Practice Learning Materials Survey.

Items were rated using a four-point performance scale: 1 (Not evident), 2 (Partially evident), 3 (Mostly evident), and 4 (Fully evident). This scale was selected to capture qualitative differences in the extent to which expected competencies and outputs were demonstrated in students' field-practice work. Each student product was independently evaluated by two evaluators across seven predefined dimensions, each assigned a specific weight, with some dimensions sharing identical weights.

This survey was initially developed by the authors and refined through several meetings. The first draft was reviewed by experts in the field, and minor revisions were made based on their feedback. Prior to administration, the survey was pilot tested with two students while the first author observed the process. No issues were identified, and the survey was then administered to the students.

As with the learning materials survey, dimension scores were calculated as the mean of the corresponding item ratings. These dimension-level scores served as the basis for the qualimetric analysis, enabling systematic comparison across dimensions and integration into a total quality index.

Validity and reliability of the research instruments

The reliability of the first instrument was examined using Cronbach's alpha coefficient. [Table 1](#) shows coefficient for individual dimensions of Field-Practice Learning Materials Instrument.

The overall internal consistency of the 22-item scale was excellent ($\alpha = .985$), indicating a very high level of reliability. Reliability coefficients for the seven dimensions were also high, ranging from $\alpha = .941$ to $\alpha = .967$. The reliability of the field-practice agenda evaluation was examined through inter-rater reliability analysis. Two independent evaluators assessed student agendas using the same rubric across the seven evaluation dimensions. Pearson's correlation coefficient was calculated to determine the level of agreement between the evaluators. The analysis indicated a strong positive correlation ($r = .91$, $p < .001$), demonstrating a high level of scoring consistency.

A confirmatory factor analysis (CFA) was conducted to test the proposed seven-factor structure of the 22-item instrument. The seven-factor model showed mixed fit, chi-square

TABLE 1 Reliability analysis of the evaluation of field-practice learning materials instrument.

Dimension	Description	Number of items	Cronbach's α
P1	Content Coverage and Depth	4	.951
P2	Objective Alignment	3	.967
P3	Development of Scientific and Inquiry Skills	3	.962
P4	Implementation of Field/Laboratory Work and Safety	3	.957
P5	Use of Models and Representations	3	.949
P6	Formative Assessment	3	.941
P7	Coherence and Sequencing	3	.955
Overall scale	Field-Practice Learning Materials Survey	22	.985

(188) = 832.34, $p < .001$, CFI = 0.952, TLI = 0.941, RMSEA = 0.094, and SRMR = 0.018. All standardized factor loadings were high (0.874–0.958), indicating strong item-factor associations. However, latent factor correlations were also high (0.751–0.935), indicating limited discriminant validity among the seven proposed dimensions. Compared with a one-factor model, the seven-factor model showed better fit; nevertheless, the high factor correlations suggest that the instrument is best interpreted as measuring an overarching evaluation construct with strongly related subdomains. For the details of the CFA see [Supplementary Appendix C](#).

Although only 25 student field-practice outputs were evaluated, this sample size is considered adequate for the purposes of detailed qualitative and rubric-based assessment. The selection was conducted randomly to ensure representativeness of the broader cohort, thereby reducing selection bias. Furthermore, the consistency of ratings, supported by high inter-rater reliability ($r = .91$), suggests that the evaluated sample provides stable and reliable estimates of student performance.

Data collection

Data collection was conducted during the summer semester of 2025, specifically between 9 June and 20 June, at Kazakh National Women's Teacher Training University. The study was organized within the framework of field-based educational practice, and data were collected under field conditions.

Data were collected face-to-face using paper-based materials. In addition, the Google Forms platform was used to record

certain data. The field practice and the data collection process were carried out by the appointed supervising instructor. The practice and data collection activities lasted two weeks.

Participants did not take part in the field practice through individual invitations; instead, participation followed the official academic calendar. For first- and second-year students assigned to the field practice, the dean's office issued an official order specifying the list of student groups, the location and duration of the practice, and the responsible supervising instructor. Students participated in the practice based on this approved order.

Prior to the commencement of data collection, all participants were informed about the purpose of the study, the research procedures, and the voluntary nature of their participation. Students were assured that all collected data would be treated confidentially and used exclusively for research purposes. Written or electronic informed consent was obtained from all participants, and no personal identifying information was collected. Anonymity and confidentiality were fully ensured throughout the study.

Data were collected from participants involved in the biology field-practice course during the study period. The sampling framework followed a census-based approach, where all students participating in the field-practice module were invited to complete the survey evaluating the learning materials. In total 387 students participated in the survey, however, 362 students submitted valid questionnaire responses, resulting in a 93% response rate for the distributed survey forms. This approach was used to minimize sampling bias and ensure that the dataset reflected the full cohort of participants. In addition, 25 field-practice agendas were randomly collected and evaluated using the rubric-based assessment instrument.

All ethical procedures were carefully followed throughout the study. Prior to data collection, participants were informed about the purpose of the research, the procedures involved, and their rights as participants. Participation was entirely voluntary, and students were assured that they could withdraw at any time without any academic consequences. Informed consent was obtained from all participants in electronic form. To ensure confidentiality and anonymity, no personally identifying information was collected, and all data were analyzed in aggregate form.

Field-based educational practice

The field practice was conducted during the summer semester of 2025, from June 9 to June 20, with first- and second-year students enrolled at the Kazakh National Women's Teacher Training University. The practice was implemented in accordance with the official academic calendar and an institutional order approved by the dean's office, and all activities were carried out under the supervision of assigned instructors. Participation in the field practice was a mandatory component of the botany curriculum and aimed at integrating theoretical knowledge with practical learning experiences in natural environments.

Data collection was carried out under authentic field-based conditions over a two-week period. Students completed practical

tasks in forested areas, botanical gardens, and regions characterized by natural plant communities. The field practice focused on developing core professional competencies in botany, including the preparation of herbarium specimens, accurate use of scientific nomenclature, systematic classification of plants, biomorphological description of local flora, and analysis of plant distribution and ecological conditions. Students also conducted comparative analyses and presented their findings using diagrams and tables.

The field practice was organized into four main stages: (1) introduction to the practice tasks, (2) plant collection, (3) scientific processing and herbarium preparation, and (4) preparation of the final report. Each day was divided into two parts. During the first half of the day, students participated in field excursions and observational activities. The second half was devoted to processing, describing, and systematizing the collected plant materials.

During the introductory stage, students were familiarized with the objectives of the field practice, safety regulations, and their individual responsibilities. In the plant collection stage, students gathered plant specimens, photographed them, and prepared collections for further analysis. Particular attention was paid to the examination of morphological structures to determine taxonomic affiliation at the family level. The third stage constituted the core scientific component of the practice, during which students conducted detailed analyses of both above-ground and below-ground plant parts. Morphological, anatomical, and ecological characteristics were examined, and each specimen was documented through individual descriptions.

The primary objective of the field practice was to deepen students' theoretical understanding of botany while developing practical skills related to plant identification, classification, and ecological analysis. Students studied different types of root systems (taproot, fibrous, and adventitious), leaf forms, flower structures, and reproductive organs using real plant specimens. Special emphasis was placed on flower morphology, including sepals, petals, stamens, pistils, and gynoecium structure. Plant collection was conducted primarily under dry weather conditions, and protected species listed in the Red Book were not damaged; instead, they were documented exclusively through photographic records.

Based on the collected materials, students prepared herbarium specimens in accordance with established methodological requirements. Plant samples were properly dried, mounted on herbarium sheets, and labeled with the scientific name, taxonomic classification, root and leaf types, and detailed morphological and ecological descriptions. Completed herbarium collections and written reports were submitted to supervising instructors and evaluated using predefined assessment criteria.

Throughout the field practice, students maintained a field practice diary, which served as the primary document recording daily activities and learning experiences. The diary included dates, descriptions of completed tasks, information on studied plant species, reflective comments, encountered difficulties, self-assessments, and personal suggestions. Diaries were reviewed and approved regularly by supervising instructors and played an important role in fostering students' research skills, reflective thinking, and academic writing competencies.

Data analysis

Data analysis was conducted in several consecutive stages, consistent with the principles of applied qualimetry and the objectives of the study. Initially, responses from both instruments were screened for completeness and correctness. Item-level responses were accumulated at the dimension level by calculating arithmetic mean scores for each predefined dimension (P1–P7). These dimension-level means helped as the primary empirical inputs for the qualimetric analysis.

For each instrument, dimension scores were normalized to a dimensionless 0–1 scale using linear normalization, based on the minimum and maximum possible values of the four-point response scale of both surveys. This step helped comparability across dimensions and eliminated scale-related bias. Normalized coefficients (K_i) were calculated for all dimensions separately for the evaluation of field-practice learning materials and the evaluation of student field-practice outcomes.

Following normalization, a non-equal weighting scheme was applied to all dimensions. The weighting coefficients assigned to each dimension were determined through an expert-informed judgment process. Specifically, an expert panel consisting of specialists in biology education and field-based instruction evaluated the relative importance of each dimension in relation to the objectives of botany field practice. Based on this evaluation, higher weights were assigned to core competencies (e.g., application of theoretical knowledge, inquiry and fieldwork skills, and herbarium preparation), while supporting dimensions received relatively lower weights. To strengthen the robustness of the weighting scheme, a sensitivity analysis was conducted by testing alternative weight distributions, which showed that the overall quality indices remained stable across reasonable variations in weights.

The total quality of each evaluation object was then expressed through an integral quality index (KQ), calculated as the sum of weighted dimension coefficients. Integral indices were calculated independently for field-practice learning materials and student field-practice outcomes. These indices provided merged measures of quality while retaining a clear link to the underlying dimensions.

In addition to composite indices, dimension-level normalized coefficients and weighted contributions were examined descriptively to identify patterns of relative strengths and weaknesses across dimensions. Results were interpreted using predefined qualimetric quality thresholds to classify total quality levels. All analyses were done using spreadsheet-based statistical measures, and results were stated in tabular form to ensure transparency and reproducibility.

Qualimetric analysis procedure

The evaluation was done using a qualimetric approach, which is designed for the systematic assessment of complex, multidimensional objects through the integration of multiple indicators into one quality indices.

Each survey item was treated as an elementary quality indicator. Indicators were grouped into predefined dimensions (P1–P7) based on theoretical and practical considerations of

field practice quality. For each dimension i , a dimension score was calculated as the arithmetic mean of its corresponding item scores:

$$x_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

Where x_{ij} is the score for item j within dimension i , and n_i is the number of items in dimension i .

To help comparability across dimensions and to remove scale dependency, raw dimension scores were transformed into dimensionless normalized quality coefficients. Given the four-point response scale used in both instruments, linear normalization was applied following standard qualimetric formula:

$$K_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

Where K_i is the normalized quality coefficient for dimension i , and $x_{\min} = 1$ and $x_{\max} = 4$ represent the minimum and maximum possible scale values.

In line with qualimetric recommendations emphasizing transparency and avoidance of random prioritization, non-equal weights were assigned to all quality dimensions. With seven dimensions included in the evaluation, the weight for each dimension was calculated as $w_i = \frac{1}{m}$ where $m = 7$ is the total number of dimensions.

The total quality of each evaluation object was expressed through an integral quality index, computed as the weighted sum of normalized dimension coefficients:

$$K_Q = \sum_{i=1}^m w_i \cdot K_i$$

The resulting index provides a single total measure of quality while retaining a clear link to the underlying dimensions and indicators.

Integral quality indices were interpreted using predefined qualimetric thresholds commonly applied in applied quality evaluation: $K_Q \geq 0.80$: High quality, $0.60 \leq K_Q < 0.80$: Acceptable–high quality, $0.40 \leq K_Q < 0.60$: Moderate quality, and $K_Q < 0.40$: Low quality

Results

This section presents the results of the qualimetric evaluation of field-practice learning materials and students' field-practice agendas. Descriptive statistics, normalized dimension coefficients, weighted contributions, and integral quality indices are reported to summarize the quality profiles of both objects of evaluation. The results are organized to allow comparison between instructional inputs and student learning outputs.

Descriptive results

Table 2 presents the dimension-level descriptive statistics for the Evaluation of Field-Practice Learning Materials survey. The results show that all seven quality dimensions received mean scores above 3.14 on a four-point scale, indicating generally positive evaluations of the learning materials. Among the

dimensions, Objective Alignment (P2) recorded the highest mean score ($M = 3.195$), followed by Scientific and Inquiry Skills (P3) ($M = 3.190$) and Content Coverage and Depth (P1) ($M = 3.186$). The lowest mean score was calculated for Coherence and Sequencing (P7) ($M = 3.143$); however, this value remained within the upper range of the scale.

Table 3 summarizes the descriptive results for the Field Practice Agenda Analysis survey. Mean scores across dimensions ranged from 2.925 to 3.625. The highest mean score was obtained for Purposefulness and Structure (P2) ($M = 3.625$). Comparatively lower mean scores were observed for Inquiry and Fieldwork Skills (P3) ($M = 2.925$) and Application of Theoretical Knowledge (P1) ($M = 2.950$). The remaining dimensions; Herbarium Preparation Quality (P4), Description and Classification Skills (P5), Reflection and Evaluation (P6), and Completeness and Overall Quality (P7), yielded mean scores above 3.10, showing generally favorable evaluations of students' field-practice performance.

Normalized qualimetric coefficients

To make comparability possible across quality dimensions and to enable subsequent aggregation, the dimension-level mean scores were transformed into normalized qualimetric coefficients (K_i) using a linear normalization procedure. Given the four-point response scale, normalization was performed by converting raw mean values to a dimensionless 0–1 scale. This procedure allows each dimension to be interpreted as a relative contribution to overall quality, independent of the initial measurement scale.

TABLE 2 Dimension-level descriptive statistics for evaluation of field-practice learning materials.

Dimension	Description	Mean (1–4)
P1	Content Coverage & Depth	3.186
P2	Objective Alignment	3.195
P3	Scientific & Inquiry Skills	3.190
P4	Field/Laboratory Implementation & Safety	3.171
P5	Models & Representations	3.168
P6	Formative Assessment	3.158
P7	Coherence & Sequencing	3.143

TABLE 3 Dimension-level descriptive statistics for field practice agenda analysis.

Dimension	Description	Mean (1–4)
P1	Application of Theoretical Knowledge	2.950
P2	Purposefulness & Structure	3.625
P3	Inquiry & Fieldwork Skills	2.925
P4	Herbarium Preparation Quality	3.100
P5	Description & Classification Skills	3.125
P6	Reflection & Evaluation	3.300
P7	Completeness & Overall Quality	3.250

TABLE 4 Normalized qualimetric coefficients for evaluation of field-practice learning materials.

Dimension	Description	Mean (1–4)	Ki (0–1)
P1	Content Coverage & Depth	3.186	0.729
P2	Objective Alignment	3.195	0.732
P3	Scientific & Inquiry Skills	3.190	0.730
P4	Field/Laboratory Implementation & Safety	3.171	0.724
P5	Models & Representations	3.168	0.723
P6	Formative Assessment	3.158	0.719
P7	Coherence & Sequencing	3.143	0.714

Table 4 include the normalized coefficients for the Evaluation of Field-Practice Learning Materials survey. The normalized values ranged from 0.714 to 0.732, indicating a consistently high level of quality across all dimensions. The highest normalized coefficient was observed for Objective Alignment (P2) ($K_i = 0.732$), while the lowest coefficient corresponded to Coherence and Sequencing (P7) ($K_i = 0.714$). The narrow range of coefficients shows a relatively balanced quality structure of the learning materials.

Table 5 displays the normalized coefficients for the Field Practice Agenda Analysis survey. In contrast to the learning materials survey, these coefficients exhibited greater variability, ranging from 0.642 to 0.875. The highest normalized value was recorded for Purposefulness and Structure (P2) ($K_i = 0.875$), while the lowest coefficients were associated with Inquiry and Fieldwork Skills (P3) ($K_i = 0.642$) and Application of Theoretical Knowledge (P1) ($K_i = 0.650$). The remaining dimensions gave normalized coefficients above 0.70, indicating generally favorable evaluations of students' field-practice performance across multiple quality features.

Weighted dimension contributions

In accordance with qualimetric theory, the overall quality of a complex educational process should not be calculated by assuming equal importance among its constituent dimensions (Azgaldov and Kostin, 2011). Field practice in botany is a complex educational activity in which different dimensions contribute to learning outcomes to changing degrees (Mautenbaev et al., 2018). Consequently, assigning equal weights to all dimensions would oversimplify the evaluation and reduce the validity of the resulting quality index.

To address this issue, differentiated weight coefficients (G_i) were assigned to each evaluation dimension (P1–P7) based on their relative input to the effectiveness of field practice. The weighting scheme reflects the pedagogical significance of each dimension, emphasizing those that directly support the integration of theoretical knowledge with practical fieldwork skills.

Table 6 presents the weight coefficients assigned to each dimension. Table 7 shows the weighted dimension contributions for the evaluation of field-practice learning materials. The

TABLE 5 Normalized qualimetric coefficients for field practice agenda analysis.

Dimension	Description	Mean (1–4)	Ki (0–1)
P1	Application of Theoretical Knowledge	2.950	0.650
P2	Purposefulness & Structure	3.625	0.875
P3	Inquiry & Fieldwork Skills	2.925	0.642
P4	Herbarium Preparation Quality	3.100	0.700
P5	Description & Classification Skills	3.125	0.708
P6	Reflection & Evaluation	3.300	0.767
P7	Completeness & Overall Quality	3.250	0.750

TABLE 6 Weight distribution for field practice evaluation dimensions.

Dimension	Description	Weight (W_i)
P1	Application of theoretical knowledge	0.20
P2	Purposefulness and structural alignment	0.10
P3	Inquiry and fieldwork skills	0.20
P4	Herbarium preparation quality	0.20
P5	Description and classification skills	0.15
P6	Reflection and evaluation	0.10
P7	Completeness and overall quality	0.05
Total		1.00

TABLE 7 Weighted dimension contributions for evaluation of field-practice learning materials.

Dimension	K_i	Weight (w_i)	$w_i \times K_i$
P1	0.729	0.20	0.146
P2	0.732	0.10	0.073
P3	0.730	0.20	0.146
P4	0.724	0.20	0.145
P5	0.723	0.15	0.108
P6	0.719	0.10	0.072
P7	0.714	0.05	0.036
Total		1.00	0.726

weights sum to 1.00, satisfying the fundamental requirement of additive qualimetric models.

Higher weights were assigned to P1, P3, and P4, as these dimensions represent the core learning objectives of botany field practice: applying theoretical knowledge in natural settings, conducting systematic field investigations, and producing scientifically valid herbarium specimens. Moderate weight was assigned to P5, which reflects students' abilities to describe and classify plant specimens correctly. Lower weights were assigned

to P2, P6, and P7, as these dimensions support but do not directly define the central learning outcomes of field practice.

Each dimension score was first converted into a relative qualimetric index using the following transformation:

$$K_i = \frac{\bar{S}_i - 1}{3}$$

where:

- $\bar{S}_i$ is the mean score for dimension i based on a four-point rating scale, and
- K_i represents the normalized quality index for that dimension, ranging from 0 to 1.

The overall quality of field practice was calculated using a weighted additive model, which integrates the normalized dimension indices and their respective weights:

$$K_{FPQI} = \sum_{i=1}^7 K_i \times G_i$$

Expanded, the formula is expressed as:

$$K_{FPQI} = 0.20K_{P1} + 0.10K_{P2} + 0.20K_{P3} + 0.20K_{P4} + 0.15K_{P5} + 0.10K_{P6} + 0.05K_{P7}$$

The resulting Field Practice Quality Index ranges from 0 to 1, with higher values indicating a higher level of effectiveness in students' field practice performance.

The total Field Practice Quality Index was 0.726, showing a high level of effectiveness in the implementation of field-practice learning materials. The largest contributions to the total quality index came from Application of Theoretical Knowledge (P1), Inquiry and Fieldwork Skills (P3), and Herbarium Preparation Quality (P4), showing the central role of hands-on scientific practice in botany field education. Lower-weighted dimensions contributed proportionally less to the final index.

Table 8 displays the weighted contributions for the Field Practice Agenda Analysis survey. The total Field Practice Agenda Quality Index was 0.706, showing a high level of quality in students' field practice agendas. The strongest contributions to the total index were observed in Herbarium Preparation

Quality (P4) and Application of Theoretical Knowledge (P1), showing students' ability to translate theoretical understanding into structured fieldwork documentation. Although Purposefulness and Structural Alignment (P2) showed the highest individual dimension score, its contribution to the total index was moderated by its lower weight, consistent with its supportive role within the qualimetric framework.

Integral quality indices

The total quality of field-practice learning materials and student field-practice performance was assessed through the calculation of integral quality indices (KQ). These indices were obtained by aggregating the weighted dimension coefficients, thereby producing a single composite measure that represents the combined contribution of all evaluated quality dimensions within the qualimetric framework.

For the Evaluation of Field-Practice Learning Materials survey, the integral quality index was calculated as $KQ_1 = 0.726$. According to the predefined qualimetric criteria, this value corresponds to a high level of quality, indicating that the learning materials provide effective support for field-practice implementation in terms of content accuracy, inquiry orientation, practical skill development, and instructional coherence.

Similarly, the integral quality index for the Field Practice Agenda Analysis survey was $KQ_2 = 0.706$, which also falls within the high quality range. This finding shows a generally strong level of student performance during field practice, especially with respect to the application of theoretical knowledge, execution of fieldwork activities, preparation of herbarium materials, reflective components, and the total completeness of field-practice outputs.

At the dimension level, the evaluation of field-practice learning materials revealed a relatively balanced quality profile, as evidenced by the limited dispersion of normalized dimension coefficients and their weighted contributions. This pattern reflects a stable and coherent instructional design that consistently supports multiple aspects of field-practice implementation. In contrast, the analysis of students' field-practice agendas showed greater variability across dimensions, particularly in areas related to the application of theoretical knowledge and inquiry-based fieldwork skills. Nevertheless, all dimensions met or exceeded the acceptable quality threshold, indicating that variations in performance did not weaken the total effectiveness of field-practice outcomes.

Discussion

The results of this study indicate that both the field-practice learning materials and students' field-practice performance achieved a high level of overall quality, as reflected by their integral qualimetric indices ($KQ_1 = 0.726$ for instructional materials and $KQ_2 = 0.706$ for student outcomes). The relatively close values of these indices suggest a strong alignment between instructional inputs and student learning outputs. This finding indicates that higher perceived quality of field-practice learning

TABLE 8 Weighted dimension contributions for field practice agenda analysis.

Dimension	K_i	Weight (w_i)	$w_i \times K_i$
P1	0.650	0.20	0.130
P2	0.875	0.10	0.088
P3	0.642	0.20	0.128
P4	0.700	0.20	0.140
P5	0.708	0.15	0.106
P6	0.767	0.10	0.077
P7	0.750	0.05	0.038
Total		1.00	0.706

materials is associated with more favorable student performance outcomes, although this relationship should be interpreted as associative rather than causal due to the non-experimental design of the study.

Despite the overall high quality levels, dimension-level analysis revealed notable differences between instructional design and student performance. The evaluation of learning materials showed a relatively balanced quality structure across all seven dimensions, as reflected in the narrow range of normalized coefficients. In contrast, student performance exhibited greater variability, particularly in dimensions related to the application of theoretical knowledge and inquiry-based fieldwork skills. This pattern suggests that while instructional materials may provide consistent structural support, the development of higher-order competencies is more variable and may require additional instructional scaffolding and repeated practice.

The relatively lower scores observed in inquiry and fieldwork skills are particularly important, as these competencies represent core objectives of experiential and inquiry-based science education. Experiential learning theory emphasizes that knowledge develops through cycles of experience, reflection, conceptualization, and experimentation (Kolb, 2015). While field practice provides opportunities for such processes, the present findings suggest that exposure to authentic tasks alone may not ensure uniform development of inquiry-related competencies. Instead, structured guidance and iterative engagement may be necessary to support deeper learning.

Similarly, inquiry-based learning frameworks show the importance of investigation, evidence-based reasoning, and the construction of scientific explanations (Hmelo-Silver et al., 2007; National Research Council, 2012). The findings of this study indicate that although students were generally successful in completing structured tasks, dimensions associated with independent inquiry and the application of theoretical knowledge were comparatively more challenging.

Another important observation concerns the balanced quality structure of the instructional materials. The relatively consistent scores across dimensions shows that the materials addressed multiple aspects of effective field-based learning, including content coverage, alignment, safety, and assessment. Such balance is important for supporting coherent learning experiences; however, it may also contribute to students perceiving instructional quality in a holistic rather than differentiated manner. This interpretation is supported by the results of the CFA, which indicated high correlations among the seven dimensions. These findings show that respondents may not clearly distinguish between different aspects of instructional quality, instead evaluating them as components of a broader, unified construct. As a result, dimension-level distinctions should be interpreted with caution, as they may reflect overlapping perceptions rather than fully independent constructs.

The application of weighted dimension coefficients further strengthened the interpretive value of the evaluation results by emphasizing competencies central to field-practice learning. However, the qualimetric approach itself introduces certain limitations. The aggregation of multiple indicators into a single composite index may obscure important variation across dimensions, particularly when those dimensions are highly interrelated. In addition, the assignment of weights, although

informed by expert judgment, involves a degree of subjectivity that may influence the resulting quality indices.

Several additional limitations should be considered when interpreting the findings. First, the study relied on self-reported data, which may be subject to response bias and may reflect overall impressions rather than precise evaluations of specific dimensions. Second, although the exclusion rate for survey responses was relatively low (6.5%), a substantially larger proportion of student field-practice products (approximately 40%) was excluded based on quality criteria. While this ensured that only valid and assessable outputs were analyzed, it may have resulted in a slight overestimation of overall performance levels. Finally, the generalizability of the findings is limited by the specific institutional and educational context in which the study was conducted. The sample was predominantly composed of early undergraduate students and was not demographically balanced, which may limit the applicability of the results to other populations or disciplines.

The findings of our study are consistent with existing research emphasizing the value of field-based and inquiry-oriented learning in science education. Previous studies have shown that field experiences support cognitive, affective, and skill-based outcomes (Esparza et al., 2024; Shinbrot et al., 2022), while inquiry-based approaches promote critical thinking and scientific reasoning (Arifin et al., 2025; Fadillah et al., 2025). The present study contributes to this body of literature by demonstrating that the perceived quality of instructional materials and student performance outcomes are closely aligned within a qualimetric evaluation framework, while also highlighting the importance of scaffolding for the development of higher-order competencies.

Conclusions

This study demonstrated that both field-practice learning materials and students' field-practice performance attained a quality when evaluated through a multidimensional qualimetric framework. The close alignment between the integral quality indices of instructional inputs and student outputs indicates that well-designed field-practice materials can support the translation of theoretical knowledge into practical scientific activities. These findings confirm the value of structured, inquiry-oriented field practice as a component of science education and show the importance of coherence between instructional design and learning implementation.

The application of weighted qualimetric modeling provided understanding of the relative contributions of different quality dimensions, revealing that competencies related to scientific practice were central to total field-practice effectiveness. At the same time, dimension-level variability in student performance shows the need for targeted pedagogical support to strengthen higher-order inquiry skills.

Implications of the study

The findings of this study have several implications for the design and implementation of field-based learning in biology education. First, the results show the importance of well-

structured instructional materials that clearly align learning objectives, practical tasks, and assessment criteria. Effective instructional design for field practice should integrate theoretical knowledge with authentic field activities, allowing students to apply biological concepts through observation, data collection, and interpretation in real environmental contexts. Second, the results show that inquiry-oriented tasks and clear procedural guidance play a key role in supporting students' engagement during field practice. Instructional materials should therefore include step-by-step guidelines, safety instructions, and opportunities for students to formulate questions, collect evidence, and interpret findings. Third, the study demonstrates that systematic evaluation frameworks, such as the qualimetric approach used in this research, can provide useful feedback for improving both instructional materials and student performance in field-practice courses. Finally, the results indicate that incorporating formative assessment and reflective activities within field-practice tasks can enhance students' ability to connect theoretical knowledge with practical experiences.

Limitations of the study

Several limitations of this study should be acknowledged. First, the research was conducted within a single institutional context, which may limit the generalizability of the findings to other educational settings or disciplines. In addition, the sample was predominantly composed of early undergraduate students and was not demographically balanced. Second, certain limitations are associated with the measurement instruments. The survey instrument was not pilot tested with students, and although internal consistency was very high, this may indicate redundancy among items. Furthermore, CFA revealed high correlations among dimensions, suggesting limited discriminant validity and potential overlap between constructs. Finally, the study is limited by its analytical approach. The use of descriptive statistics and qualimetric aggregation, while appropriate for the study's objectives, may obscure variability across dimensions and does not allow for strong inferential conclusions.

Future research can apply the qualimetric framework across diverse disciplines, institutions, and educational levels to test its robustness and transferability. Longitudinal studies examining changes in quality indices over time would provide deeper vision into how field-practice competencies change with repeated exposure and instructional refinement. In addition, integrating qualitative data, such as student reflections, field observations, or instructor interviews, could enrich the interpretation of qualimetric results and illuminate the mechanisms underlying performance differences across dimensions. Such extensions would strengthen the empirical basis of qualimetric evaluation in education and also support the development of more effective, evidence-informed field-practice models.

Data availability statement

The datasets presented in this article are not publicly available due to privacy and research ethics considerations. Requests to access the datasets should be directed to gulmira-eshim@mail.ru.

Ethics statement

The studies involving humans were approved by Kazakh National Women's Teacher Training University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin because A written survey was not conducted.

Author contributions

GY: Writing – original draft, Writing – review & editing. GB: Writing – review & editing. EAK: Writing – review & editing. EK: Supervision, Writing – review & editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2026.1795880/full#supplementary-material>

References

- Arifin, Z., Sukarmin, Saputro, S., and Kamari, A. (2025). The effect of inquiry-based learning on students' critical thinking skills in science education: a systematic review and meta-analysis. *Eurasia J. Math. Sci. Technol. Educ.* 21 (3), em2592. doi: 10.29333/ejmste/15988
- Azgalov, G. G., and Kostin, A. V. (2011). Applied qualimetry: its origins, errors and misconceptions. *Benchmarking Int. J.* 18 (3), 428–444. doi: 10.1108/14635771111137796
- Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Rev. High. Educ.* 1, 5–22. https://www.tru.ca/_shared/assets/Constructive_Alignment36087.pdf
- Bonney, E. A., Dah, E. M., Doodoo, C., and Asamoah, D. Y. (2024). Improving pre-service science teachers' reflective practice through problem-based learning approach at colleges of education in Ghana. *J. Res. Didact. Sci.* 3 (1), 15–35. doi: 10.51853/jorids/15181
- Brookhart, S. M. (2013). How to create and use rubrics for formative assessment and grading. ASCD. Available online at: [https://www.geocities.ws/bdktraining/pdfkur/How%20to%20Create%20and%20Use%20Rubrics%20for%20Formative%20Assessment%20and%20Grading%20\(%20PDFDrive%20\).pdf](https://www.geocities.ws/bdktraining/pdfkur/How%20to%20Create%20and%20Use%20Rubrics%20for%20Formative%20Assessment%20and%20Grading%20(%20PDFDrive%20).pdf) (Accessed November, 2025).
- Esparza, D., Wagler, A., Olimpo, J. T., and Ashbrook, L. (2024). How students' science, social, and personal identities interact with immersive field biology experiences. *CBE—Life Sci. Educ.* 23 (2), ar23. doi: 10.1187/cbe.24-02-0046
- Fadillah, M. A., Usmeldi, U., Lufri, L., Mawardi, M., and Festiyed, F. (2025). Systematic literature review of inquiry-based learning models for optimizing learning outcomes at various educational levels. *J. Pedagog. Sociol. Psychol.* 7 (4), 278–293. doi: 10.33902/JPSp.202532540
- Fedesco, H. N., Cavin, D., and Henares, R. (2020). Field-based learning in higher education: exploring the benefits and possibilities. *J. Scholarsh. Teach. Learn.* 20 (1), 65–84. doi: 10.14434/josotl.v20i1.24877
- Florida Museum. (2023). Preparation of plant specimens for deposit as herbarium vouchers. University of Florida. Available online at: <https://www.floridamuseum.ufl.edu/herbarium/methods/vouchers/> (Accessed November 2025).
- Fogelman, S. E., Moreno Vera, M. A., and McNeill, K. L. (2025). A curriculum evaluation of a genetics unit from a science identity perspective. *J. Res. Sci. Math. Technol. Educ.* 8 (SI), 219–240. doi: 10.31756/jrsmt.419SI
- Guàrdia, L., Maina, M., Mancini, F., and Martinez Melo, M. (2023). Key quality factors in digital competence assessment: a validation study from teachers' perspective. *Appl. Sci.* 13 (4), 2450. doi: 10.3390/app13042450
- Gulikers, J. T. M., Bastiaens, T. J., and Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educ. Tech. Res. Dev.* 52 (3), 67–86. doi: 10.1007/BF02504676
- Harrison, R., Meyer, L., Rawstorne, P., Razee, H., Chitkara, U., Mears, S., et al. (2022). Evaluating and enhancing quality in higher education teaching practice: a meta-review. *Stud. High. Educ.* 47 (1), 80–96. doi: 10.1080/03075079.2020.1730315
- Hmelo-Silver, C. E., Duncan, R. G., and Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and Clark. *Educ. Psychol.* 42 (2), 99–107. doi: 10.1080/00461520701263368
- Hole, T. N., Jönsson, L. M., and Sandell, K. (2018). Working and learning in a field excursion. *CBE—Life Sci. Educ.* 17 (2), ar18. doi: 10.1187/cbe.17-08-0185
- Jonsson, A., and Svingby, G. (2007). The use of scoring rubrics: reliability, validity, and educational consequences. *Educ. Res. Rev.* 2 (2), 130–144. doi: 10.1016/j.edurev.2007.05.002
- Kitts, K., and Perry, S. (2023). I was a good classroom teacher, why do I feel so lost now? Examining a doctoral student's transition with becoming a teacher educator during and post pandemic. *Eur. Educ. Res.* 6 (2), 55–70. doi: 10.31757/euer.625
- Kolb, D. A. (2015). *Experiential Learning: Experience as the Source of Learning and Development*. 2nd ed. Upper Saddle River, NJ: Pearson Education, Inc.
- Mautenbaev, A., Atanbaeva, G., Kyrbassova, E., Parmanbekova, M. K., Baitasheva, G. U., Ashirova, Z. B., et al. (2018). Using the mindmap method (associogram) in the study of biology. *J. Pharm. Sci. Res.* 10 (12), 3214.
- Moon, J. A. (2004). *A Handbook of Reflective and Experiential Learning*. London: Routledge.
- Mutseekwa, C. (2024). Towards an integrated model for orienting preservice science teachers for 21st-century teaching. *Int. Educ. Rev.* 3 (1), 1–19. doi: 10.58693/ier.311
- Nardo, M., Saisana, M., Saltelli, A., and Tarantola, S. (2005). *Tools for Composite Indicators Building*. European Commission, Joint Research Centre.
- National Academies of Sciences, Engineering, and Medicine. (2022). *Science and Engineering in Preschool Through Elementary Grades: The Brilliance of Children and the Strengths of Educators (Chapter 7)*. National Academies Press. doi: 10.17226/26215
- National Research Council. (2012). *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. National Academies Press. doi: 10.17226/13165
- OECD. (2017). *The OECD Handbook for Innovative Learning Environments*. OECD Publishing. doi: 10.1787/9789264277274-en
- Oxley, S. P., and Turner, D. (2022). Field-Based environmental chemistry course for interdisciplinary, project-based learning. *J. Chem. Educ.* 99 (8), 2877–2884. doi: 10.1021/acs.jchemed.1c01003
- Race, A. I., Hillerislambers, J., and Levine, A. S. (2021). How an early, inclusive field course can build persistence in ecology and evolutionary biology undergraduates. *CBE—Life Sci. Educ.* 20 (3), ar40. doi: 10.1187/cbe.21-01-0017
- Saisana, M., and Tarantola, S. (2002). *State-of-the-art Report on Current Methodologies and Practices for Composite Indicator Development*. Ispra: European Commission, Joint Research Centre. doi: 10.13140/RG.2.1.1505.1762
- Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Soc. Indic. Res.* 81 (1), 65–77. doi: 10.1007/s11205-006-0024-9
- Sanjito, B. P., and Hyams-Ssekasi, D. (2025). The impact of science practical work on secondary school Students' learning outcomes: a meta-analysis. *J. Res. Sci. Math. Technol. Educ.* 8 (3), 7–24. doi: 10.31756/jrsmt.832
- Seyferth-Zapf, C., Bärnreuther, C., Stephan, M., Ehmann, M., and Seyferth-Zapf, M. (2024). SELFIEfor(STUDENT)TEACHERS? A mixed-methods study on the suitability and validity of the self-reflection tool for student teachers. *Eur. Educ. Res.* 7 (3), 7–33. doi: 10.31757/euer.732
- Shinbrot, X. A., Treibergs, K., Hernández, L. M. A., Esparza, D., Ghezzi-Kopel, K., Goebel, M., et al. (2022). The impact of field courses on undergraduate knowledge, affect, behavior, and skills: a scoping review. *BioScience* 72 (10), 1007–1017. doi: 10.1093/biosci/biac070
- Taylor, J. A., Van Scotter, P., and Coulson, D. (2007). Bridging research on learning and student achievement: the role of instructional materials. *Sci. Educ.* 16 (1), 36–43. Available online at: <https://files.eric.ed.gov/fulltext/EJ783420.pdf> (Accessed November, 2025).
- University of Western Ontario, Department of Biology. (2023). Fieldwork risk assessment guide. Available online at: https://www.uwo.ca/biology/fieldwork/risk_assessment.html (Accessed November, 2025).
- Wiggins, G. (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco, CA: Jossey-Bass.