Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Exploring the Relationship Between Project-Based Reflection and Self-Assessed Digital Pedagogical Practices in Teacher Education

Stelios Stefanos Drousiotis

s.drousiotis@alexander.ac.cy

Alexander College https://orcid.org/0009-0009-8073-3485

Giorgos Nicou

Alexander College https://orcid.org/0009-0003-3476-8712

Research Article

Keywords: Pre-service teacher education, Digital teaching practices, Project-based learning, Reflective practice, Self-reported digital competence, Sustainable educational practices

Posted Date: November 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11551303/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Preparing future teachers for flexible digital pedagogical practice requires a clear distinction between practical engagement, perceived capabilities and reflective processes. This exploratory, non-preregistered cross-sectional investigation combined questionnaire responses with administrative records from 150 teacher education students enrolled in four courses, representing 82 independent individual project units and 28 collaborative teams. Statistical models adjusted for course differences and accounted for clustering at the administrative project-unit level. Perceived project design and support were positively associated with project-specific reflection (β = 0.256, 95% CI [0.118, 0.393]; n = 149). Project reflection was also associated with additional variation in a combined index measuring self-reported digital practices and capabilities (β = 0.268, 95% CI [0.075, 0.462]; n = 122; $\Delta R^2 = 0.044$), with the estimated association becoming stronger after adjustment for reported experience. Among 114 participants with complete data across the relevant measures, a post hoc diagnostic identified a positive association with practice-oriented ratings (β = 0.306, 95% CI [0.114, 0.497]). Estimates for capability-oriented ratings were less precise, and the diagnostic neither established differences between the associations nor confirmed the validity of separate measurement scales. Linked written responses distinguished expectations of reusing educational resources, employing digital tools for personal study and rejecting technology use in classroom contexts. Analyses concerning teamwork, anticipated future use and administrative assessment scores offered limited additional support. Overall, the findings identify conditional associations between self-reported measures that may inform discussions about reflective judgement and future pedagogical adaptation. They do not independently establish perceived capability, demonstrated competence development, sustained technology use or environmental sustainability.

1. Introduction

1.1 Reflective Decision-Making in Digital Teacher Education

Preparing future educators to incorporate digital technologies into their teaching requires careful consideration of when, why, and for which learners particular digital resources are pedagogically suitable. Such decisions involve educational objectives, student participation, assessment approaches, and the capacity to adjust instructional strategies when learning needs or teaching circumstances evolve. In this study, educational sustainability refers to the ability to make informed and continuous adaptations to digital pedagogy beyond the immediate requirements of a specific course. This perspective is grounded in sustainable assessment, which emphasises the relationship between current learning experiences and students' ability to evaluate and respond to future learning demands (Boud & Soler, 2016). The broader conceptual context is informed by UNESCO's educational capacity-building agenda (UNESCO, 2020) and research examining responsible approaches to digital teaching (Colás-

Bravo et al., 2021). The central issue is therefore how prospective teachers perceive opportunities to articulate, evaluate, and reconsider their pedagogical decisions.

Project-based activities can provide a tangible means of translating pedagogical decisions into educational artefacts, although maintaining engagement in project tasks and supporting meaningful learning require different forms of instructional provision (Blumenfeld et al., 1991). Reviews of project-based learning distinguish students' perceptions of their experiences from the underlying learning processes and the quality of the products they develop (Guo et al., 2020). Research on teacher preparation offers a complementary perspective, with Tondeur et al. (2012) identifying the importance of role modelling, learning through design, and authentic experiences supported by appropriate scaffolding. Their more recent SQD2 synthesis integrates reflective practice with collaborative design, authentic teaching experiences, and feedback, while also acknowledging the roles of professional identity and affective considerations (Tondeur et al., 2025). Reflection is consequently recognised as an established component of broader approaches to teacher preparation. Nevertheless, the question addressed in this investigation concerns the extent to which students' accounts of project activities within a particular course provide meaningful information within this wider educational framework.

Reflection constitutes the principal conceptual connection underpinning the present investigation. Nguyen et al. (2014) characterise reflection as a process involving attentive, critical, exploratory, and iterative examination of thoughts, actions, and the conceptual frameworks informing them. The educational context is also significant, as the aims, interpersonal relationships, and conditions associated with professional learning influence the opportunities available for reflective engagement (Boud & Walker, 1998). Within teacher education, describing an educational experience is therefore distinct from explaining decisions through the comparison of alternatives or the questioning of underlying assumptions (Hatton & Smith, 1995). Kember et al. (2000) further differentiate reflection and critical reflection from habitual action and understanding. These distinctions justify examining general and critical reflective self-reports as separate constructs rather than combining them into a single comprehensive measure. They also support investigating reflection associated with course-based project activities without assuming that questionnaire responses collected at the same time establish enduring personal characteristics, the depth of reflective engagement, or transformative learning.

Students' perceptions of project conditions provide an approach for examining this context-dependent relationship. Constructive alignment connects intended learning outcomes, instructional activities, and assessment procedures, thereby establishing criteria against which learners can evaluate their decisions (Biggs, 1996). Research on educational feedback similarly highlights the importance of learners actively engaging with information about their work, including opportunities to reconsider and revise it (Nicol & Macfarlane-Dick, 2006). Clearly defined objectives, transparent assessment criteria, appropriate levels of challenge, feedback, and instructional guidance may consequently encourage more critical examination of decisions made during digital pedagogical design. In the present study, these elements are represented through a perceived project design/support index. This measure deliberately excludes general perceptions of project relevance and satisfaction, as including these aspects could obscure the

distinction between conditions supporting learning and an overall favourable evaluation of the project experience. Accordingly, the index reflects students' subjective perceptions of the project environment rather than providing an independent evaluation of instructional quality or demonstrating that feedback was subsequently implemented.

Recent investigations further demonstrate the importance of clearly identifying what evidence concerning teacher preparation represents. Combining questionnaire responses with documentary evidence can provide insight into students' experiences of direct instruction, integrated practical activities, and instructional modelling (Momdjian et al., 2025), while successive cycles of design-based research can contribute to refining educational scaffolding and support mechanisms (Bernsteiner et al., 2025). Studies examining reflection within teaching practicum contexts have explored professional learning needs and reflective dialogue (Tiainen et al., 2024), whereas other investigations have differentiated favourable evaluations of project experiences from explanatory and solution-oriented forms of reflection (Kaldi & Zafeiri, 2023). Across these perspectives, documenting educational provision, evaluating it positively, and critically examining the reasoning underlying pedagogical decisions represent distinct forms of evidence. This distinction provides the rationale for analysing perceived project conditions and project-specific reflection separately, while acknowledging that both measures may partly capture participants' evaluations of the same educational experience.

1.2 From Project-Based Reflection to Digital Pedagogical Self-Assessment

The examination of digital pedagogical outcomes requires a clear distinction between professional competence and the forms of evidence used to assess or describe it. The European Framework for the Digital Competence of Educators (DigCompEdu) identifies six interconnected domains of educators' digital competence, encompassing areas such as digital resources, teaching and learning, assessment, and learner engagement (Redecker, 2017). Similarly, a broader professional and ethical perspective positions digital competence as extending beyond the technical ability to operate digital tools (Falloon, 2020). Although these frameworks establish the pedagogical dimensions relevant to teacher preparation, they do not imply that self-reported evaluations are equivalent to demonstrated professional performance. Competence involves dispositions, context-dependent processes, and observable actions rather than perceived ability alone (Blömeke et al., 2015). Research investigating instruments designed to assess digital competence beliefs further highlights the importance of clarifying precisely what respondents are evaluating (Quast et al., 2025). Systematic reviews have also identified considerable variation in the conceptualisation of digital competence and widespread dependence on self-reported measures (Smestad et al., 2023). Within the adapted questionnaire employed in this study, certain items describe reported practices, expressed through statements such as "I use," whereas others address perceived capabilities through statements such as "I can." Consequently, the term digital pedagogical practice/capability self-ratings is used consistently throughout the investigation.

Tondeur et al. (2018) investigated perceptions of teacher preparation through a multilevel survey involving 931 final-year pre-service teachers. Their combined measure of SQD preparation strategies incorporated reflection alongside other approaches intended to support technology integration. The present investigation examines a different aspect of this relationship by considering whether reflection associated with specific course projects provides additional information about digital pedagogical self-ratings beyond general and critical reflection, perceived project conditions, and participants' reported experience. This analytical comparison distinguishes the contribution of these self-reported measures within a common statistical model; however, it does not establish the independent effectiveness of reflection as an instructional strategy. The updated SQD2 framework provides the pedagogical justification for examining this relationship (Tondeur et al., 2025). Accordingly, the contribution of this study lies in empirically examining the additional information provided by project-specific reflective self-appraisal under clearly specified measurement and statistical adjustment procedures, without claiming to identify distinct underlying latent constructs or demonstrate the development of digital competence. In contrast, Yao and Abd Halim (2026) investigated preparation for technology integration and education for sustainable development using a quasi-experimental research design.

Several alternative explanations must nevertheless be considered. Participants evaluated their reflective engagement, perceived project conditions, and digital pedagogical practices or capabilities within the same data collection period. As a result, individual tendencies to respond favourably to positively framed statements may have influenced the reported associations. Furthermore, items measuring project reflection ask respondents to attribute their reconsideration of learners, instructional decisions, and educational experiences to the project itself, creating the possibility that positive perceptions of the project contributed to the covariance between measures. Similarities in the pedagogical content addressed by the questionnaire items represent another potential source of overlap. The statistical model therefore examines conditional associations among self-reported variables, rather than establishing causal relationships. Even when analyses are adjusted using a fixed sample, they cannot determine the extent to which the observed associations arise from participants' reported experiences, shared evaluative tendencies, or overlapping measurement content.

Differences in educational opportunities and measurement procedures introduce additional considerations. A recent systematic review of research involving pre-service teachers identified substantial variation in theoretical frameworks, educational contexts, and assessment methods (Zhang et al., 2026). Consequently, differences in digital pedagogical self-ratings may reflect variations in participants' prior learning opportunities and practical experiences, as well as their perceptions of professional preparation. In this investigation, reported experience in developing digital materials, preparing lesson plans, and undertaking project-based activities is examined separately because these activities represent different forms of educational engagement. The scoring procedures also distinguish responses explicitly indicating non-applicability from low self-assessment ratings. Similarly, an unanswered questionnaire item cannot automatically be interpreted as indicating low competence or non-applicability without evidence explaining the missing response. Clearly specified item-coverage requirements and sensitivity analyses are therefore used to make these methodological limitations

transparent, while maintaining the distinction between the available self-reported evidence and an independently established benchmark of digital proficiency.

1.3 Educational Sustainability, Complementary Evidence, and Research Questions

Educational sustainability offers an interpretative framework for understanding this distinction by considering whether digital pedagogical decisions remain appropriate when learners' needs and educational circumstances evolve. The concept of sustainable assessment emphasises the development of informed judgement that continues to support learning beyond the completion of an immediate educational activity (Boud & Soler, 2016). From this perspective, attention is directed towards the considerations that justify retaining, modifying, or discontinuing a digital resource, including its educational relevance, accessibility, and conditions governing its use. UNESCO's Recommendation on Open Educational Resources addresses issues relating to licensing and the adaptation of educational materials (UNESCO, 2019), while the concept of entangled pedagogy highlights the relationship between technological choices, pedagogical purposes, and contextual circumstances (Fawns, 2022). These perspectives inform the interpretation of participants' self-assessments and expressed intentions. However, they do not constitute direct measurements of sustainability outcomes, and the continued use of a digital resource cannot independently demonstrate either long-term pedagogical adaptability or environmental benefits (Selwyn, 2021).

Collaborative experiences, intentions regarding future technology use, and written reflections offer additional perspectives for examining these commitments. The concept of Teamwork Quality encompasses communication, coordination, individual contributions, mutual support, effort, and group cohesion (Hoegl & Gemuenden, 2001). Theoretical perspectives concerning social interdependence and shared regulation further demonstrate why participation in collaborative activities does not necessarily indicate sustained collective reasoning (Johnson & Johnson, 2009; Järvelä & Hadwin, 2013). Similarly, an expressed intention to use digital resources in the future should be distinguished from the transfer of learning, which requires the generalisation and continued application of acquired knowledge or practices beyond the original educational environment (Baldwin & Ford, 1988). Intentions may not translate into actual behaviour when contextual circumstances or limitations in individual control restrict implementation (Ajzen, 1991). Young et al. (2026) illustrate the importance of these contextual constraints in their evaluation of digital competency-based placements, identifying differences in school-level access and institutional support. These considerations provide the theoretical basis for exploratory analyses involving eligible participants engaged in collaborative projects and responses to the item concerning anticipated future use. Brief written accounts and administrative assessment scores are additionally employed to contextualise or question patterns identified through self-reported measures, with each source interpreted according to its evidential limitations.

The investigation addresses two interconnected primary research questions within a cross-sectional sample of students undertaking teacher education. Existing research on project-based learning has

already documented considerable variation in collaborative learning processes and associated educational outcomes (Sjølie et al., 2026). The present study adopts a more focused approach by comparing perceived project conditions, general and critical reflective orientations, and reflection specifically associated with course-based project activities. All statistical analyses are exploratory, and the designation primarily refers solely to the reporting priority assigned to these analyses rather than indicating preregistered hypotheses or confirmatory testing.

The central practical consideration is whether project-specific reflective self-reports provide additional information beyond broader reflective self-assessments and favourable perceptions of project conditions. A sequence of statistical adjustments conducted using a consistent analytical sample is employed to clarify the influence of reported experience, while focused scoring checks assess the sensitivity of the findings to questionnaire item content and respondent coverage. These comparisons address the explanatory information and methodological limitations associated with the specified questionnaire indices within the analysed sample. They cannot establish that the indices represent independent underlying abilities, demonstrate temporal relationships between the measured variables, or determine whether reflection produced changes in pedagogical practice. Consequently, the interpretation of project-specific reflection must consider the wording of the relevant questionnaire items, the specification of the statistical models, and both the estimated magnitude and uncertainty of the observed associations.

- RQ1. To what extent are perceived project design and support, together with general and critical reflection, associated with project-specific reflection after accounting for reported experience and course context?
- RQ2. Does project-specific reflection explain additional variation in self-reported digital pedagogical practices and capabilities beyond these factors, and how is this association influenced by statistical adjustment and scoring procedures?

Supplementary analyses investigate teamwork among eligible participants involved in collaborative projects, intentions concerning future digital resource use, the correspondence between self-reported measures and administrative assessment scores, and brief written responses. These analyses situate the primary statistical associations within specific educational judgements and differing expectations regarding future technology use, thereby identifying the types of evidence required to assess informed pedagogical adaptation beyond the immediate course context. Participants' open-ended responses were independently coded by two individuals using separately developed coding frameworks. The resulting coding outputs were subsequently compared and reconciled through a retrospective review of the original written responses. The qualitative synthesis was informed by this reconciliation process, with the procedures and limitations described in Section 2.4.

2. Materials and Methods

2.1. Research Design, Educational Context, and Data Sources

This exploratory cross-sectional investigation examined technology-intensive project-based activities undertaken within four teacher education courses described as being delivered at Alexander College, Cyprus, in association with Canterbury Christ Church University (CCCU), UK. Data were collected using an online questionnaire administered through the 1KA platform. Questionnaire responses were subsequently linked to administrative project identifiers and course assessment scores. The research design did not incorporate baseline measurements collected before the projects, a comparison group, or longitudinal follow-up assessments. The reporting approach was informed by the cross-sectional items of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (von Elm et al., 2007).

All 173 students enrolled across the four courses received invitations to participate. Following data cleaning, the final workbook contained 151 respondent records and one summary row representing non-respondents. Excluding the summary row and one respondent whose record lacked an administrative project identifier resulted in an analytical dataset of 150 linked participants. Among these respondents, 27 did not satisfy the minimum requirement of 17 substantive responses across the 22 items used to calculate the digital pedagogical score. Consequently, 123 participants had sufficient data to generate this score, although one had missing information concerning practical experience (Q5), leaving 122 complete cases for the analysis addressing RQ2. The cleaned workbook did not permit reconstruction of the questionnaire access or consent-screening processes for the 22 invited students who were not represented among its respondent records. The underlying research records and associated analytical documentation are retained confidentially in a restricted-access Google Drive repository to support research integrity and protect participant information. Administrative project identifiers recorded in source column CL distinguished 82 independently completed individual project units and 28 collaborative teams involving 68 respondents, producing a total of 110 administrative project units. The original investigator coordinated and taught the four courses, evaluated students' project submissions, invited participation in the research, and retained exclusive access to the research response data. The original investigator personally delivered the Bachelor and Master ICT in Education courses, whereas practical sessions in Multimedia and Hypertext and ICT Didactics were conducted by two teaching assistants, each assigned to one course. The original investigator confirmed that all projects administratively classified as individual were completed independently and that responses concerning project conditions (Q20) and project-specific reflection (Q24) referred to the corresponding administratively linked projects. As these confirmations were provided by the instructor responsible for the original investigation, they should not be regarded as independently verified evidence. Disagreement between self-reported working arrangements and administrative project classification was identified for 73 of the 150 respondents. Administrative records were therefore used to determine the applicability of teamwork measures: Q21 and Q22 responses associated with individual projects were treated as

inapplicable, while missing collaborative-project ratings were not estimated or replaced. The resulting teamwork analysis included 50 respondents distributed across 25 teams.

Administrative project units were constructed by retaining both course and project prefixes, ensuring that independently completed work undertaken by each student was assigned a separate unit. This procedure also applied to the three-artefact portfolio completed within Bachelor ICT in Education. Course indicators were incorporated into the analytical models to account for differences in average outcomes across the four observed educational contexts. Nevertheless, this adjustment did not establish experimental comparability between programme levels, participants' reported experience, or project formats, nor did it permit their individual effects to be identified independently. Administrative project format was not treated as an intervention, and the inconsistent self-reported format variable was excluded from the primary adjustment procedures. Course membership was established using the administrative subject and project prefixes, including the single case in which this classification differed from the participant's response to Q3. The original investigator additionally confirmed that the administratively identified groups completed separate projects, and no further grouping associated with shared project outputs was identified.

The investigation took place during the 2025/26 academic year. Multimedia and Hypertext and Master ICT in Education were delivered during the winter semester, while Bachelor ICT in Education and ICT Didactics were offered in the summer semester. Invitations were extended to all 173 enrolled students, comprising 92 students in Bachelor ICT in Education, 10 in Multimedia and Hypertext, 26 in ICT Didactics, and 45 in Master ICT in Education. Questionnaire responses were collected between 4 May and 14 September 2026, following the granting of ethical approval on 8 April 2026. Following completion of the project activities, the questionnaire link was made available through the relevant course's online classroom. Although completing an individual reflection constituted a compulsory course requirement, participation in the research was independent of this academic obligation. Consent to the research use of responses was voluntary, and declining participation did not influence course requirements or assessment grades. Question Q1 requested permission to use anonymised or pseudonymised responses for research purposes, whereas Q2 separately sought consent to publish anonymised quotations from open-ended answers. Only responses accompanied by this additional permission were included in the qualitative analysis. The introductory information accompanying the questionnaire informed participants that research analyses would be undertaken after grading had been completed.

Each of the four courses extended over 15 weeks, with project-based activities occupying approximately 10–12 weeks. In Multimedia and Hypertext, students developed an interactive multimedia website focused on educational robotics, incorporating guidance for teachers, lesson plans, and learning resources for pupils. Students enrolled in Master ICT in Education designed educational websites intended for the learner groups associated with their respective programmes. Within Bachelor ICT in Education, students independently completed three projects involving web-based and photographic materials, an infographic, and an educational video, which were presented through a portfolio managed by the instructor. Participants in ICT Didactics organised a Moodle learning environment (Moodle Pty

Ltd., West Perth, WA, Australia) and developed interactive educational resources addressing selected digital technologies. Accordingly, the expression project-specific refers to project activities undertaken within the relevant course, including the three-component portfolio in Bachelor ICT in Education, rather than necessarily referring to one individual artefact.

The organisation of project activities followed the ADDIE framework, comprising analysis, design, development, implementation, and evaluation. Comprehensive project instructions were distributed during the first teaching week and made accessible through the courses' online learning environments. Following a brief introductory presentation of relevant content, students progressed through project-related problems with guidance from mentors and feedback provided throughout the different stages of development. Further assistance was available through scheduled consultations and email communication. Students were also permitted to revise their completed work in response to feedback. These descriptions represent the instructional arrangements reported for the courses; students' perceptions of these arrangements were examined separately through the questionnaire index measuring perceived project design and support.

2.2. Measurement Instruments and Scoring Procedures

The questionnaire incorporated Greek and English language items adapted to the educational context or developed by the investigator with reference to established conceptual frameworks relating to reflective thinking, digital competence, and teamwork. These items were not administered as unchanged versions of previously validated measurement instruments. Substantive questionnaire responses were evaluated using a five-point scale ranging from 1 to 5. Composite scores comprising multiple items were calculated as unweighted arithmetic means, provided that respondents had completed at least 75% of the relevant items, with the minimum required number of responses rounded upward. Internal consistency estimates were calculated using cases with complete responses across the relevant items and were interpreted as descriptive indicators of response consistency rather than evidence of factorial validity or measurement invariance.

General reflection and critical reflection were assessed independently using questionnaire items Q11a-d and Q11e-h, respectively. Each measure required valid responses to at least three of its four constituent items. This separation was informed by the conceptual distinctions established in the original reflective thinking framework (Kember et al., 2000), thereby avoiding the assumption that different forms of reflective engagement constitute a single undifferentiated construct. Both measures were collected at the same time as the project-related evaluations. Project-specific reflection was assessed using six items, Q24c-h, with a minimum of five substantive responses required for score calculation. These items addressed consideration of learners' experiences, pedagogical effectiveness, professional identity, reported experience, the reasoning behind decisions and alternative approaches, and retrospective learning. Items Q24a-b, which focused primarily on routine implementation or adherence to instructions, were excluded from the principal project-reflection measure. Their inclusion as additional items, together with an alternative scoring procedure involving reverse coding, was examined separately through sensitivity analyses.

Perceived project design and support were measured using five questionnaire items, Q20a, Q20b, and Q20d-f, with at least four substantive responses required to calculate the composite score. These items examined the clarity of learning objectives, transparency of assessment criteria, appropriateness of project challenges, availability of feedback, and provision of instructional assistance. Item Q20c, which concerned the perceived relevance of project activities to future teaching, was deliberately omitted to minimise conceptual overlap with the measure of anticipated future use. The resulting index was treated as an exploratory representation of students' perceived project conditions rather than as a validated measure of instructional implementation fidelity. Four items concerning previous experience, Q8a-d, were incorporated into the statistical models as separate predictors. These addressed experience in developing digital materials, planning technology-supported instruction, producing digital content, and participating in project-based activities. Practical teaching-related experience during the preceding 12 months, assessed through Q5, included activities such as teaching practicum, classroom observation, tutoring, substitute teaching, and workshops. Responses were coded on a four-category scale from 0 to 3, representing no experience, fewer than 10 hours, 10–50 hours, and more than 50 hours, respectively. An alternative categorical specification was examined through sensitivity analysis. Items Q8a-d were entered individually into the models as linear predictors using their original five-point response scales. Their shared question stem explicitly referred to experiences undertaken before the beginning of the course. Nevertheless, these measures were collected retrospectively through the questionnaire administered after project completion and should not be interpreted as prospectively recorded baseline observations.

Digital pedagogical practice/capability self-ratings were assessed through 22 questionnaire items covering the six competence areas identified in the DigCompEdu framework (Redecker, 2017). Six items were phrased in terms of perceived capability, whereas the remaining 16 referred to reported pedagogical practices. All items employed common response anchors progressing from not yet undertaking a particular practice to its systematic application in more demanding educational circumstances. The overall mean across the 22 items was retained as the principal descriptive outcome because it represented the complete content administered to participants, including pedagogical activities addressed through capability-oriented statements. The subset consisting exclusively of practice-worded items also covered all six DigCompEdu areas; consequently, coverage of competence areas alone did not provide a justification for prioritising the combined measure. The decision to retain the full index did not establish that the items represented a single underlying latent construct, nor did it resolve the conceptual inconsistency between capability-oriented wording and response anchors describing frequency of practice. Neither of the subsets defined by item wording had undergone pilot testing or independent validation as a distinct measurement scale. Calculation of the overall score required at least 17 substantive responses from the 22 items. This criterion was satisfied by 123 participants, of whom 89 provided responses to every item. The scoring requirement did not ensure that each respondent contributed information from all six competence areas. No proficiency classifications corresponding to DigCompEdu levels A1-C2 were assigned. Across the questionnaire, 22 item responses were coded as 6, indicating "not relevant/no opportunity", while 349 cells were already blank for reasons

that could not be reconstructed. Both categories were excluded from numerical scoring, although only code 6 explicitly represented non-applicability. Missing or inapplicable values were neither imputed nor substituted with low ratings.

Supplementary assessment of teamwork quality was based on 12 items aligned with the Teamwork Quality framework, Q21a-l, requiring at least nine substantive responses and applying exclusively to participants undertaking collaborative projects (Hoegl & Gemuenden, 2001). Individual contribution was assessed separately through items Q22a–e and included as an additional self-reported covariate in a sensitivity analysis. Anticipated future use was measured through the single ordinal item Q20j, which asked respondents about their intention to apply knowledge or skills acquired through the project in subsequent studies or professional practice. This item was not restricted to future classroom teaching. Items Q20g–i were not incorporated into the future-use measure because perceptions of applicability, understanding, and satisfaction do not represent the same construct as behavioural intention.

Administrative assessment scores were available for all respondents whose questionnaire records could be linked to project identifiers. These scores were retained on their original proportional scale, ranging from 0.400 to 0.958. They were subsequently standardised within each course and analysed at the respondent level, as different scores had been awarded to members of the same team in 19 of the 28 collaborative groups. Five supplementary regression analyses examined administrative assessment scores as the dependent variable. Each model considered one focal predictor, perceived project design/support, project-specific reflection, digital pedagogical self-ratings, teamwork quality, or future-use intention, while adjusting for course membership and accounting for clustering within administrative project units. These analyses investigated associations with course-based assessment outcomes and were not equivalent to RQ2 models incorporating grades as adjustment variables. The assessment rubrics differed across courses and project artefacts, incorporating criteria relating to individual learning processes or reflective documentation alongside the quality of completed products. The analyses used the administrative proportional scores as recorded rather than recalculating grades from the maximum values specified in the rubrics. Although standardisation within courses addressed differences in score distributions, it did not establish equivalence between assessment criteria. Furthermore, project assessment scores did not represent an established measure collected before project participation. Consequently, these administrative results were interpreted as contextual evidence rather than as a standardised indicator of digital pedagogical competence.

The original Greek language questionnaire items and their corresponding English translations were reviewed and approved by the investigator. Initial English translations were produced using DeepL Pro (DeepL SE, Cologne, Germany), after which specialists in ICT in education, computer science, and computer science didactics reviewed the translated material for conceptual consistency and terminological accuracy. No preliminary pilot study was undertaken.

2.3. Quantitative Analysis and Statistical Procedures

For RQ1, project-specific reflection was specified as the dependent variable, with perceived project design/support serving as the principal predictor. The model additionally incorporated general reflection, critical reflection, four distinct self-reported experience measures (Q8a-d), practical teaching experience (Q5), and course indicators as adjustment variables. For RQ2, digital pedagogical practice/capability self-ratings constituted the dependent variable, while project-specific reflection was the main predictor, alongside perceived design/support and the same adjustment variables. The incremental analysis for RQ2 compared two models that were identical except for the inclusion of project-specific reflection. To ensure comparability across the diagnostic adjustment sequences, the complete-case samples remained unchanged, comprising 149 respondents for RQ1 and 122 for RQ2. Dependent variables and continuous predictors were standardized using the sample standard deviations calculated within their respective fixed analytical samples. Course indicators remained unstandardised, with bachelor-level ICT serving as the reference category where applicable. These regression models were intended to estimate conditional associations rather than causal effects. Although Q8 captured participants' retrospective accounts of experience preceding the course, it did not represent a prospectively collected baseline measure. Ordinary least squares regression was estimated using a CR1 cluster-robust covariance estimator based on administrative project units. The small-sample correction multiplier was calculated as (G/(G-1)\times(n-1)/(n-k)), where (G) denotes the number of administrative project units, (n) represents the number of observations, and (k) is the number of estimated model parameters, including the intercept. Two-sided confidence intervals and p-values were calculated using a (t(G-1)) reference distribution, while residual degrees of freedom were defined as (n-k). This approach accommodates dependence among observations belonging to the same recorded project unit while assuming independence between separate units. Additional null-imposed studentised wild cluster tests were performed for the two principal focal coefficients and the supplementary teamwork coefficient, using the same administrative project units, Rademacher weights, and 4,999 resampling iterations. The two primary focal hypothesis tests were treated as one family and adjusted using the Holm procedure (Holm, 1979). Reported confidence intervals were marginal rather than simultaneous, while p-values associated with adjustment sequences and sensitivity analyses were interpreted as unadjusted diagnostic results. Changes in (R^2) were used to describe variation explained within the observed sample and were not interpreted as evidence of competence development or predictive performance in new samples.

The supplementary teamwork analysis specified project-specific reflection as the dependent variable and teamwork quality as the focal predictor. Perceived design/support, general reflection, critical reflection, and course indicators were included as covariates, with the analysis restricted to eligible respondents who had undertaken collaborative projects. A separate model examining future-use intention treated Q20j as the dependent variable and project-specific reflection as the principal predictor, adjusting for design/support, both broader reflection measures, the four reported experience variables (Q8a-d), and course indicators. Practical teaching experience (Q5) was omitted from the principal intention model to preserve the complete sample of 150 respondents with intention data, although an additional sensitivity analysis incorporated Q5. These models addressed different conditional relationships and were therefore interpreted separately. The two supplementary focal hypothesis tests

formed an independent Holm-adjusted family. Similarly, the six analyses concerning digital competence areas and the five analyses involving administrative assessment scores were organised into separate families, with Holm corrections applied within each group. Although these procedures addressed multiple testing within the predefined analytical families, they did not eliminate uncertainty arising from the wider range of analytical decisions informed by the observed data.

Sensitivity analyses investigated the influence of digital-score completeness requirements, weighting across DigCompEdu areas, overlap between questionnaire content, alternative approaches to scoring reflection, adjustment for routine or implementation-related responses, exclusion of individual courses, influential observations, and alternative coding of practical experience. A separate item-wording diagnostic calculated mean scores for the 16 practice-oriented items, requiring at least 12 substantive responses, and the six capability-oriented items, requiring at least five. The complete 22-item digital index and both wording-based indices were analysed using identical RQ2 covariates and the intersection of respondents meeting the eligibility requirements for all three outcomes. These item groupings were treated as descriptive subsets rather than independently validated measurement scales, and their marginal confidence intervals were not interpreted as formal tests of differences between the outcomes. An exact omitted-variable decomposition was undertaken to investigate the change associated with adjustment for Q8 within the fixed RQ2 sample. Availability weighting was additionally applied to examine selection associated with observed covariates. Further checks of the future-use analysis included ranked and proportional-odds specifications for Q20j, adjustment for practical experience, and an expanded index of perceived project design/support. The supplementary statistical results are provided in Appendix A (Supplementary File S1: Statistical Tables). Additional analytical records and supporting research documentation are retained securely in a restricted-access research repository and are not included in the supplementary materials accompanying this manuscript. No automated selection of predictors, mediation modelling, or causal-effect estimation was undertaken. An additional content-based diagnostic compared the full digital self-rating index with alternative indices excluding Q13c and excluding both Q13c and Q15d. These comparisons were performed using the same 122 respondents included in the principal RQ2 analysis, with identical covariates and clustering procedures.

The investigation was not preregistered, and the final analytical specification was established after the investigator had access to the observed dataset and regression results. Accordingly, the designation primarily identifies the reporting hierarchy rather than indicating a prospectively established confirmatory analysis. Decisions documented retrospectively included the requirement for at least 75% item completion, the exclusion of inconsistent self-reported project format from the principal adjustment models, and the definition of the hypothesis-testing families. The approved ethics application outlined general intentions to construct framework-informed composite measures, examine internal consistency, apply regression or multilevel analytical methods, and undertake qualitative analysis. However, it did not prospectively specify the precise questionnaire item combinations, minimum response-coverage thresholds, covariate groupings, cluster-robust estimation procedures, or multiplicity-adjustment families used in the final analyses. The comparison between practice-worded and capability-worded digital items, together with the Q8 covariance decomposition, constituted post hoc diagnostic investigations.

Conceptual considerations informed the selection of the six-item project-specific reflection measure, the separate general and critical reflection measures, and the five-item perceived design/support index. The principal quantitative analyses were conducted using R version 4.3.3 (R Foundation for Statistical Computing, Vienna, Austria). All findings were therefore interpreted as exploratory and hypothesis-generating. The procedures used to compare the two sets of qualitative coding results and establish reconciliation rules were developed retrospectively during manuscript revision, after access to participants' responses, and should not be interpreted as a prospectively designed assessment of coding reliability.

2.4. Written Responses and Integration of Evidence

The questionnaire included three open-ended prompts inviting participants to describe an experience involving teamwork (Q23), identify their most significant learning insight (Q25), and outline a future development objective together with a specific action they intended to undertake (Q26). The complete coding inventory comprised 397 non-empty responses submitted by 146 participants who had authorised the use of anonymised quotations. These consisted of 105 responses to Q23 and 146 responses each to Q25 and Q26. However, the manuscript applied more restrictive eligibility requirements: participants needed an identifiable administrative project link, while responses to Q23 were additionally limited to those associated with eligible collaborative projects. Following the exclusion of the respondent without an administrative link, Q23 responses that did not satisfy the collaboration criterion, and the remaining entry containing only a slash, the eligible dataset consisted of 338 written responses from 145 contributors. This included 48 accounts concerning teamwork, 145 statements describing learning insights, and 145 responses identifying future development goals. One eligible response to Q25 did not clearly relate to digital pedagogy and was retained in the inventory without receiving substantive codes. Consequently, the final denominator for substantive qualitative coding comprised 337 responses: 48 for Q23, 144 for Q25, and 145 for Q26. The broader inventory was retained for auditing and comparing the two coding outputs rather than extending the target sample reported in the manuscript. Given their brief and prompted nature, these written responses were not considered sufficient evidence to establish the depth of participants' reflective engagement or the subsequent implementation of their intended actions.

The original investigator initially applied a directed descriptive coding approach [39], organising responses around pedagogical purpose, reflective reconsideration, coordination and responsibility, accessibility and responsible technology use, reuse and adaptation, and perspectives expressing disagreement or intentions unrelated to classroom teaching. Twelve illustrative quotations, including examples that challenged the dominant interpretations, were verified against exact passages in the eligible original Greek responses before being translated into English. Each quotation was subsequently associated with the corresponding participant's response to the future-use intention item (Q20j). The selection of illustrative quotations was not revised on the basis of these intention ratings. Only the initial 12 quotations were included in the linkage display. The resulting joint display [40] was used to support interpretation of the written responses and their relationship to questionnaire findings. It was not

intended to quantify the prevalence of qualitative themes or independently validate participants' self-reported assessments.

The original investigator and a second individual independently produced separate sets of coding outputs using domains defined by the questionnaire prompts, while developing their own semantic subcategories. The original investigator subsequently examined both coding outputs and reviewed the reconciliation process against the original written responses. Records were matched according to respondent identity, questionnaire item, and source text. Before individual coding assignments were compared, the definitions and inclusion or exclusion criteria of the two coding systems were examined to identify semantically equivalent categories, broader or narrower conceptual relationships, partially overlapping meanings, and concepts represented in only one coding scheme. Each code was counted no more than once for an individual response, irrespective of how many passages within that response supported its application. The comparison distinguished exclusions, differences in textual segmentation, and substantive differences in interpretation. The resulting harmonised coding framework contained 85 labels distributed across Q23, Q25, and Q26, including categories serving technical or administrative purposes. A single overall Cohen's kappa coefficient was not calculated because the original coding frameworks were not identical. Accordingly, the comparison of the two coding outputs and the original investigator's subsequent review were not presented as evidence of independent human intercoder reliability.

Coding disagreements were resolved through examination of supporting passages in relation to the complete response and the corresponding questionnaire prompt. More specific codes were retained only when explicitly supported by the respondent's statement, rather than inferred from broader conceptual similarities. Multiple codes were assigned where distinct meanings were identifiable, while negative, qualified, and conditional perspectives were preserved. Reconciliation considered source traceability, eligibility, semantic correspondence, and consistent application of coding rules. Initial coding decisions, supporting passages, final interpretations, and explanations for their resolution were retained in a restricted audit record. Separately, the Q26 rubric classified the specificity of participants' professional development goals and proposed actions. These classifications described the content of written intentions rather than demonstrated professional competence.

2.5. Ethical Considerations and Data Protection

Although completing an individual reflective activity formed part of the compulsory course requirements, consent for the use of participants' responses in the research was obtained independently through questionnaire item Q1. Separate permission to reproduce anonymised quotations from written responses was requested through Q2. Students who declined to authorise the research use of their responses were not subject to any consequences concerning their academic responsibilities or assessment grades. The questionnaire was administered following completion of the project activities, and its introductory information explicitly stated that the research analyses would be undertaken after grading had been completed. To protect participant confidentiality and reduce the possibility of identification, particularly within small course cohorts and among students undertaking individual

projects, direct enrolment identifiers, original respondent identifiers, and complete written responses were excluded from publicly accessible supplementary materials.

The application for ethical approval was submitted on 4 January 2026, and formal approval was granted by the responsible institutional ethics committee under Decision No. 24/2026 on 8 April 2026. Questionnaire responses were collected between 4 May and 14 September 2026, following the granting of ethical approval.

3. Results

3.1. Sample Characteristics, Coverage, and Score Distributions

Table 1 presents the characteristics of the respondents whose questionnaire records were successfully linked to administrative project information. The dataset comprised 132 female participants (88%) and 18 male participants (12%), with complete information available for this variable. Participants' ages were not collected through the questionnaire. All 80 respondents enrolled in Bachelor ICT in Education completed their project activities individually, whereas 68 of the 70 participants from the remaining courses undertook collaborative projects. Consequently, the distribution of project formats across courses did not permit a clearly differentiated examination of the independent effects of course context and working arrangement. The analytical model addressing RQ1 included 149 respondents distributed across 109 administrative project units. For RQ2, the final sample consisted of 122 respondents representing 94 project units, comprising 65 students from Bachelor ICT in Education, 31 from Master ICT in Education, 19 from ICT Didactics, and seven from Multimedia and Hypertext. The remaining 28 linked respondents who were not included in the RQ2 complete-case sample were distributed across these courses as follows: 15 from Bachelor ICT in Education, six from Master ICT in Education, six from ICT Didactics, and one from Multimedia and Hypertext. Among these 28 respondents, 27 did not provide sufficient substantive responses to calculate the digital pedagogical self-rating index, while one had a valid digital score but lacked information concerning practical teaching experience (Q5). All models within the RQ2 adjustment sequence were estimated using the same 122 complete cases to ensure comparability. These sample sizes reflect the eligibility criteria applied to each analysis and do not establish the representativeness of the analytical samples relative to the full population of invited students.

Calculation of the digital score required at least 17 substantive responses across the 22 questionnaire items. The primary digital analysis included 122 respondents with complete covariate information. Course designations correspond to the four administratively defined educational contexts, while the comparison between the 122 included respondents and the 28 excluded respondents is provided in Appendix A (Supplementary File S1: Statistical Tables).

Table 2 summarises the descriptive statistics and internal consistency estimates for the study measures. General and critical reflection demonstrated a weak positive correlation (r = 0.178), providing a rationale for examining them separately at the descriptive level without confirming that they represent distinct underlying latent factors. Although the digital pedagogical self-rating measure exhibited high internal consistency, its constituent items addressed six different competence areas. Teamwork quality and future-use intention scores were clustered towards the upper ends of their respective scales, indicating potential ceiling effects that limited their capacity to distinguish between respondents. These elevated ratings should not be interpreted as evidence of consistently effective teamwork or the subsequent application of acquired knowledge and skills.

All measures were evaluated on a five-point scale ranging from 1 to 5. Cronbach's alpha (α) was calculated using respondents with complete item-level data, comprising 89 participants for the digital pedagogical measure and the sample sizes reported in Table 2 for the remaining measures. Appendix A (Supplementary File S1: Statistical Tables) provides the supporting statistical results, including cluster-bootstrap confidence intervals for Cronbach's alpha, detailed definitions of the composite scores, and distributions of individual questionnaire items.

3.2. Perceived Project Design and Support in Relation to Project-Specific Reflection

Following adjustment for all specified covariates, perceived project design and support demonstrated a statistically significant positive association with project-specific reflection (β = 0.256, 95% CI [0.118, 0.393], p < 0.001; Table 3). The corresponding Holm-adjusted p-value for the primary analysis remained below 0.001, while the wild cluster test produced a p-value of 0.0004. Positive regression coefficients were also observed for general reflection (β = 0.363, 95% CI [0.199, 0.527]) and critical reflection (β = 0.211, 95% CI [0.061, 0.361]). The fully adjusted model accounted for 35.9% of the variation observed within the analytical sample (R^2 = 0.359). Furthermore, the coefficient associated with perceived project design and support remained positive throughout all 109 analyses in which individual project units were excluded sequentially, with estimates ranging from 0.215 to 0.275. These additional examinations indicate numerical stability within the available dataset rather than demonstrating that the findings would necessarily generalise beyond the observed sample.

G represents the number of administrative project units, while β indicates a standardised regression association. The arrows identify the direction of the regression relationships and should not be interpreted as evidence of causality. Confidence intervals were calculated using the CR1 cluster-robust estimator and the t(G−1) distribution, with degrees of freedom of 108, 93, 24, and 109 for the four models, respectively. The Holm correction was applied to two separate families of hypothesis tests, comprising the two primary analyses and the two supplementary analyses. The RQ1 and RQ2 models incorporated perceived project design/support, general and critical reflection as separate measures, four reported experience variables (Q8a–d), practical teaching experience (Q5), and course indicators. Project-specific reflection was additionally included as a predictor in the RQ2 model. The supplementary

teamwork model adjusted for perceived design/support, general reflection, critical reflection, and course membership. The future-use intention model included project-specific reflection, perceived design/support, general and critical reflection, Q8a–d, and course indicators. The residual degrees of freedom, calculated as n − k, were 137, 109, 43, and 138 for the respective models. Complete regression coefficient estimates for the primary and supplementary analyses are provided in Appendix A (Supplementary File S1: Statistical Tables).

3.3. Associations Between Project-Specific Reflection and Digital Pedagogical Self-Ratings

In the fully adjusted regression model, project-specific reflection demonstrated a positive association with digital pedagogical self-ratings (β = 0.268, 95% CI [0.075, 0.462], p = 0.007; Holm-adjusted p = 0.007; wild cluster p = 0.0060). The model incorporating project-specific reflection explained 39.5% of the observed variation (R^2 = 0.395), compared with 35.1% for the designated reduced model (R^2 = 0.351), representing an incremental increase of ΔR^2 = 0.044. Within the fixed-sample adjustment sequence illustrated in Fig. 1, controlling for general and critical reflection, perceived project design/support, and course membership produced a coefficient of β = 0.202, with a 95% confidence interval of [− 0.008, 0.412]. After incorporating the four reported experience variables (Q8a–d), the estimated coefficient increased to β = 0.270 (95% CI [0.079, 0.461]), while further adjustment for practical teaching experience (Q5) resulted in only a minor change. The exact covariance decomposition indicated that most of the 0.068 increase was associated with the adjustment terms for previous experience in technology-supported lesson planning (Q8b, + 0.050) and digital content development (Q8c, + 0.017). Under the specified adjustment conditions, these two experience variables exhibited negative conditional relationships with project-specific reflection but positive conditional relationships with digital pedagogical self-ratings. Although Q8 addressed experiences preceding the course, the information was collected retrospectively rather than through prospective baseline assessment. Consequently, the increase in the estimated coefficient following adjustment should be interpreted as a change in conditional covariance within the observed dataset rather than evidence of a causal effect of reflection or a specific underlying mechanism. The analysis does not establish that one adjustment specification is inherently preferable. Furthermore, comparisons within the fully adjusted model did not demonstrate that project-specific reflection was more strongly associated with digital pedagogical self-ratings than general or critical reflection (Appendix A, Supplementary File S1: Statistical Tables).

Figure 1. Relationship Between Project-Specific Reflection and Digital Pedagogical Practice/Capability Self-Ratings Across Sequentially Adjusted Models. The plotted points represent standardised regression coefficient estimates, while the horizontal error bars display 95% confidence intervals calculated using the CR1 cluster-robust covariance estimator, with clustering at the administrative project-unit level and inference based on t(93). Each model was estimated using the same fixed sample of 122 respondents distributed across 94 project units, with consistent standardisation applied within the analytical sample. Adjustment variables were introduced progressively, with each successive model retaining those included in the preceding specification. The fully adjusted model incorporated practical teaching

experience (Q5) in addition to four separately measured indicators of retrospectively reported pre-course experience (Q8a–d). These analyses represent retrospective comparisons of statistical associations under different adjustment specifications and should not be interpreted as demonstrating a causal progression or causal relationships.

The analysis examining differences in questionnaire item wording included 114 respondents across 88 administrative project units who satisfied the eligibility requirements for all three digital scoring approaches and had complete information for the fully adjusted model. The association between project-specific reflection and practice-oriented digital self-ratings was estimated at β = 0.306 (95% CI [0.114, 0.497]), compared with β = 0.189 (95% CI [− 0.038, 0.416]) for capability-oriented ratings and β = 0.278 (95% CI [0.082, 0.475]) for the overall digital index. Each outcome was standardised using its respective standard deviation calculated within the common analytical sample. However, the marginal confidence intervals were not designed to establish statistically significant differences between the estimated coefficients, and the wording-based item groups were not independently validated measurement scales. Additional sensitivity analyses using alternative scoring approaches continued to produce positive coefficient estimates. Restricting the analysis to respondents who completed all 22 digital items yielded β = 0.329 (95% CI [0.106, 0.551]; n = 88). In a separate sensitivity analysis excluding Q13c, the estimated coefficient was β = 0.246 (95% CI [0.056, 0.436]; n = 125). This analysis used a different analytical sample from the subsequent fixed-sample Q13c exclusion diagnostic (n = 122). Applying equal weighting across the six available DigCompEdu areas resulted in β = 0.260 (95% CI [0.049, 0.470]; n = 98). Because these alternative specifications modified both the composition of the analytical sample and the scoring procedures, their results could not independently identify the influence of questionnaire item content. Availability-weighted estimation produced β = 0.263 (95% CI [0.076, 0.450]) under the specified model of selection based on measured variables. When Bachelor ICT respondents were excluded, the estimated coefficient decreased to β = 0.166 (95% CI [− 0.115, 0.446]), based on 57 respondents distributed across 29 project units. Although the confidence intervals from all 94 analyses excluding one project unit at a time remained above zero, this numerical stability did not demonstrate that the association would generalise across different course contexts. The complete results of these sensitivity analyses are presented in Appendix A (Supplementary File S1: Statistical Tables).

An additional model incorporating adjustment for routine implementation continued to identify a positive association involving the original six-item project-specific reflection measure (β = 0.268, 95% CI [0.077, 0.459]). An alternative scoring procedure, which reverse-coded Q24a–b before combining them with Q24c–h, produced a smaller coefficient of β = 0.191 (95% CI [− 0.022, 0.405]). This revised scoring approach altered the conceptual content represented by the reflection measure. Its confidence interval included both zero and the coefficient obtained from the primary specification. Consequently, these findings illustrate uncertainty associated with defining the boundaries of the project-reflection construct rather than demonstrating the validity of either scoring method or providing evidence of an association in the opposite direction. As explained in Section 2.3, the six-item reflection measure was finalised after the investigator had access to the observed data and analytical results.

A further post hoc diagnostic examined the potential influence of overlapping questionnaire content while retaining the same 122 respondents and 94 administrative project units included in the primary RQ2 analysis. All model comparisons employed identical adjustment variables and the same CR1 cluster-robust inference procedure. Following the exclusion of Q13c, the coefficient for project-specific reflection was β = 0.256 (95% CI [0.064, 0.449]). Excluding both Q13c and Q15d resulted in β = 0.246 (95% CI [0.050, 0.442]), compared with β = 0.268 (95% CI [0.075, 0.462]) when the complete set of 22 digital items was retained. Item Q13c addresses reflection concerning one's own teaching practices, whereas Q15d concerns the facilitation of self-regulated learning, including reflective activities. Removing these items was intended to investigate the influence of explicitly overlapping questionnaire wording rather than eliminate measures representing an identical construct. When the coefficients were recalculated using the standard deviation of the original global digital outcome as a common reference, the corresponding estimates were 0.268 for the full 22-item measure, 0.257 after excluding Q13c, and 0.246 following the removal of both Q13c and Q15d. Although these reduced estimates remained consistent with a positive conditional association, the comparisons did not constitute formal statistical tests of differences between coefficients. Furthermore, they could not rule out wider conceptual overlap between questionnaire items or variation attributable to the shared self-report method (Supplementary File S1).

Across the six digital competence areas examined independently, the strongest evidence of positive associations after Holm adjustment was observed for digital resources (β = 0.241, adjusted p = 0.044), assessment (β = 0.255, adjusted p = 0.023), and facilitating learners' digital competence (β = 0.292, adjusted p = 0.025). The analytical samples differed across these area-specific models, ranging from 115 to 140 respondents. Accordingly, the estimated coefficients should not be interpreted as a comparative ranking of effects across competence areas or as evidence that participants experienced measurable improvements in digital pedagogical competence.

3.4. Associations Involving Teamwork, Future-Use Intentions, and Administrative Assessment Scores

The estimated association involving teamwork quality was characterised by considerable statistical uncertainty (β = 0.136, 95% CI [-0.107, 0.379]; n = 50, G = 25; wild cluster p = 0.3526). Furthermore, 32% of the observed teamwork quality scores reached the highest possible value on the measurement scale. This coefficient therefore represents the relationship identified within the analysed subgroup rather than the complete population of 68 respondents eligible for the collaborative-project assessment. A similar concentration towards the upper end of the response scale was observed for future-use intention. Specifically, 89 of the 150 respondents (59.3%) selected the maximum rating of five, whereas only nine participants provided ratings of three or below. In the fully adjusted supplementary model, the estimated coefficient linking project-specific reflection with future-use intention was β = 0.091 (95% CI [-0.121, 0.302]). The proportional-odds sensitivity analysis similarly produced an uncertain association between project-specific reflection and future-use intention (OR = 1.391, 95% CI [0.860, 2.250]). The limited

number of responses within the lower rating categories also constrained the assessment of the proportional-odds assumption.

None of the five supplementary analyses examining administrative assessment scores, after adjustment for course membership, provided clear evidence of convergence with the self-reported measures (all Holm-adjusted p = 1.000). The estimated association with digital pedagogical self-ratings was β = 0.075 (95% CI [-0.141, 0.292]), while the corresponding estimate for project-specific reflection was β = 0.009 (95% CI [-0.174, 0.192]). The uncertainty reflected in these estimates, together with differences in assessment criteria across courses, limits the extent to which administrative course assessment scores can serve as independent validation of questionnaire-based measures.

3.5. Qualitative Accounts and Integrated Interpretation of Findings

Both initial coding exercises examined the same inventory of 397 written entries and independently identified 393 responses as suitable for substantive coding. The remaining four entries consisted of two responses containing only a slash, one teamwork account providing contextual information without substantive coding content, and one learning insight that did not clearly relate to digital pedagogy. Following the application of the manuscript's eligibility criteria, 337 responses remained available for substantive qualitative coding. The two original coding frameworks comprised 52 and 57 substantive codes, respectively, together with two technical flags in the first framework and three in the second. Because the coding systems differed in their conceptual scope and level of detail, variations in the number of assigned codes or segmented text passages could not be treated as direct indicators of coding disagreement. The reconciliation process reviewed by the original investigator retained the initial coding decisions alongside the final interpretations, ensuring that the analytical process remained traceable.

Within the eligible written responses, distinguishing an explicitly stated educational purpose from a generally favourable evaluation of technology remained important. One participant account (E15) explained how using a chess clock could shift attention towards completing moves quickly rather than carefully considering them, while also acknowledging potential educational advantages associated with technology. This account therefore justified retaining both the possible pedagogical benefits and the limitations identified by the respondent, rather than assigning an exclusively positive or negative interpretation. Another response (E16) described restricting the use of digital resources to teachers' preparation and organisational activities to avoid increasing children's exposure to screens. This selective approach to technology use differed from the explicit rejection of digital technologies in classroom teaching expressed in E09.

Participants' responses concerning future professional development also demonstrated a distinction between identifying a learning objective and specifying how that objective would be achieved. One account (E17) suggested participating in seminars and practical training sessions to develop additional skills without identifying the particular competencies to be acquired. By comparison, another response

(E18) proposed extending an existing website, testing it with pupils, and subsequently modifying it in response to their feedback. The latter example explicitly connected a proposed activity with evaluation and further refinement. Nevertheless, these accounts describe the content and specificity of participants' intended plans rather than providing evidence that the proposed activities were implemented. Similarly, the absence of an explicit evaluation component in a brief written response cannot be interpreted as demonstrating that the respondent lacks the ability to evaluate educational activities.

Among the 145 eligible responses to Q26, 95 identified a specific professional development goal, while 78 described a concrete intended action. Of these, 57 responses included both a clearly defined goal and a specific proposed action. The remaining responses presented development objectives or intended actions with varying degrees of specificity. These classifications reflect the detail provided in participants' written statements rather than the quality, feasibility, or subsequent implementation of their professional development plans.

The qualitative accounts highlighted considerations relating to educational purpose and the potential continuation of digital pedagogical practices rather than expressing consistently positive attitudes towards technology. In E01, the respondent emphasised that digital technology becomes educationally meaningful when it contributes to achieving clearly identified learning objectives. Another participant (E02) proposed consulting accessibility guidelines when developing educational resources, while E03 identified the importance of verifying image licensing conditions. These statements demonstrate awareness of particular actions that participants intended to undertake. However, the actual accessibility of the resulting resources, compliance with relevant licensing requirements, and implementation of these intentions were not independently examined.

The possibility of reusing educational materials was discussed in relation to the initial effort required for their development and their potential availability for subsequent activities. For example, E04 explained that preparing certain resources requires considerable time, although the completed materials, such as the website developed during the course, may subsequently be used on multiple occasions. Similarly, E05 anticipated adapting previously created materials over several years. Neither response provided evidence concerning the actual frequency of subsequent reuse or measurable reductions in time or resources. Their relevance to educational sustainability therefore concerns participants' expressed intentions to maintain the appropriate use of educational materials over time, rather than independently verified durability, resource savings, or environmental benefits.

The teamwork accounts illustrated different aspects of collaborative engagement and responsibility. E06 connected mutual checking of completed work with the provision of feedback, whereas E07 described how insufficiently defined agreements contributed to duplicated effort before improved coordination was established. E08 similarly associated the distribution of responsibilities with communication needs that had not been adequately addressed. In E12, the respondent questioned whether responsibility should remain concentrated with one individual, while E11 emphasised personal accountability for completing

an assigned contribution to the website. The coding review also maintained a distinction between describing the allocation of tasks and explicitly evaluating whether that allocation was fair. For instance, E13 referred to distributing responsibilities according to participants' interests, whereas E14 specifically characterised the distribution as fair. Consequently, an account explaining how tasks were assigned was not automatically interpreted as evidence that the respondent considered the arrangement equitable.

Connecting the written accounts with their contributors' responses to the future-use intention item (Q20j) provided additional clarification concerning the different contexts in which participants anticipated applying their learning. Both E04 and E05 selected the maximum intention rating of 5/5 and described expectations of reusing or adapting educational resources. E10 reported an intention score of 4/5 while discussing potential applications for personal learning and examination preparation. In contrast, E09 selected the minimum rating of 1/5 and explicitly stated an intention not to incorporate digital technologies into classroom lessons. This combination was consistent with a low level of intended classroom use, although the response did not establish the reasons underlying that position. Considered together, these examples demonstrate why anticipated application in future studies or professional activities should not be equated directly with the transfer of digital practices into classroom teaching. The linked accounts clarify possible interpretations of the future-use measure without establishing the prevalence of particular intended applications across the wider sample. Table 4 brings together the quantitative findings and selected qualitative accounts, while identifying the limitations associated with interpreting these complementary forms of evidence.

The excerpt labels do not represent participant identifiers. All 12 original excerpts are linked to their respective contributors’ Q20j ratings, which are measured on a five-point scale (1–5). These excerpts were selected prior to conducting the linkage diagnostic. A further six excerpts illustrate distinctions within the coding process and are not included in the joint display. The illustrative presentation does not provide a basis for determining category frequencies or explaining any regression coefficient. The retrospective reconciliation process maintains the eligibility criteria established for the manuscript.

4. Discussion

4.1. Perceived Project Conditions and Reflection on Course-Based Work

Students’ perceptions of project design and support were associated with project-specific reflection, while the project-reflection measure contributed additional information about digital pedagogical self-ratings beyond general reflection, perceived learning conditions, course membership, and retrospectively reported experience. The diagnostic analysis focusing on practice-oriented questionnaire items provides a more specific interpretation of this relationship in terms of participants’ reported digital teaching practices. Nevertheless, the statistical models represent conditional associations between differently formulated self-evaluations, with estimates remaining sensitive to the measurement approaches and adjustment specifications employed. Supplementary written accounts further indicate that favourable

ratings of intended future technology use may concern the reuse of educational resources, adaptation of existing materials, or applications supporting personal study. Together, these findings distinguish the information represented by the different self-report measures without demonstrating independent abilities, improvements in professional competence, or the subsequent implementation of digital practices.

The observed association between perceived project design/support and project-specific reflection is consistent with an understanding of reflection that recognises the influence of educational context (Boud and Walker, 1998). Clearly established learning objectives and success criteria may provide a basis for evaluating decisions, while challenging activities can encourage reconsideration of initial approaches. Similarly, feedback and instructional support may provide the evidence and resources necessary to review and refine previous decisions. This interpretation connects the principles of constructive alignment (Biggs, 1996) and formative feedback (Nicol and Macfarlane-Dick, 2006) with the deliberate and iterative reflective processes discussed by Nguyen et al. (2014). It also corresponds with the updated SQD framework, which positions reflection within a broader range of strategies designed to prepare prospective teachers for integrating digital technologies into educational practice (Tondeur et al., 2025). However, the findings do not establish which individual project condition contributes most strongly to reflection, since the design/support index combines students' evaluations of several educational conditions. Consequently, accounts of project-specific reflection should be understood in relation to the learning environment and project requirements experienced by participants.

An alternative explanation concerns the wording of the project-specific reflection items. Four questionnaire statements asked respondents whether their project experiences encouraged reconsideration of learners, pedagogical quality, professional identity, or previous experience. Agreement with these statements therefore involved attributing reflective benefits to the same projects whose design and support were being evaluated. Participants who perceived their projects favourably may have provided similarly positive ratings across both measures, while those reporting greater reflective engagement may also have evaluated the support received more positively. Although distinguishing general reflection from critical reflection acknowledges their conceptual differences (Kember et al., 2000), simultaneous statistical adjustment cannot separate these retrospective evaluations from reflective activity itself. Consequently, overlapping questionnaire content and shared evaluative tendencies remain plausible explanations for the observed associations (Podsakoff et al., 2024).

4.2 Additional Insights into Digital Pedagogical Self-Appraisal

Within the shared analytical sample of 114 respondents, project-specific reflection demonstrated a positive association with practice-oriented digital self-ratings ($\beta = 0.306$, 95% CI [0.114, 0.497]), whereas the estimate for capability-oriented ratings showed greater statistical uncertainty ($\beta = 0.189$, 95% CI [-0.038, 0.416]). These findings provide a more specific basis for interpreting the relationship in terms of participants' reported digital pedagogical practices. However, the marginal confidence intervals do not

demonstrate that the associations differ between the two item groups. Both subsets continue to employ response categories based on frequency, while also varying in their pedagogical focus rather than solely in questionnaire wording. Furthermore, the locally modified questionnaire items were not pilot-tested, and the two subsets have not been independently validated as separate measurement scales. The combined digital index represents the complete collection of digital questionnaire items; consequently, its association should not be interpreted as evidence specifically concerning perceived capability or objectively demonstrated professional performance (Blömeke et al., 2015; Quast et al., 2025).

The comparison with the research conducted by Tondeur and colleagues primarily concerns differences in the scope of the measures being examined. Their study investigated a combined framework of teacher preparation strategies, including reflection, whereas the present research considers reflection associated with course project work alongside broader reflective dispositions and students' perceptions of learning conditions (Tondeur et al., 2018). The additional proportion of explained variance ($\Delta R^2 = 0.044$) represents information captured within the particular analytical sample and model specification. It does not demonstrate that project-specific reflection is more influential than broader reflective orientations or establish which component of the updated SQD framework is effective (Tondeur et al., 2025). Accordingly, the contribution to teacher education research lies in explicitly identifying the focus of the reflection report, its questionnaire content, and the adjustment variables considered when examining its association with digital pedagogical self-appraisal.

Shared method variance and overlapping questionnaire content remain plausible alternative explanations for the findings associated with RQ2. Ratings collected simultaneously from the same respondents may reflect a generally favourable evaluation of the project or a tendency to agree with positively formulated statements (Podsakoff et al., 2024). Item Q13c addresses reflection concerning teaching practice, whereas Q15d incorporates reflection within the broader context of supporting self-regulated learning. When these explicitly reflection-related items were excluded from the analysis involving the same 122 respondents, the estimated project-reflection coefficient decreased from 0.268 to 0.246, with a 95% confidence interval of [0.050, 0.442]. Although this particular exclusion did not eliminate the observed association, neither the resulting estimate nor the wording-based diagnostic demonstrates discriminant validity or measures the extent of shared method variance. The alternative eight-item reflection specification and the analysis excluding the largest course also produced less precise estimates. Furthermore, administrative assessment scores were based on differing evaluation criteria, while the qualitative accounts originated from the same respondents. Neither source therefore provides an independent and consistently measured performance criterion capable of resolving these competing explanations.

Retrospectively reported experience requires cautious interpretation. Within the fixed RQ2 analytical sample, adjustment for previous experience increased the estimated association between project-specific reflection and digital pedagogical self-ratings. Most of this change was algebraically attributable to reported experience in technology-supported lesson planning and digital educational content development. Under the specified adjustment conditions, these experience measures were negatively

associated with project-specific reflection but positively associated with digital pedagogical self-ratings. However, this covariance pattern does not demonstrate that prior experience inhibits reflective learning or that statistical adjustment eliminates potential bias. Considering these experience measures separately acknowledges differences among preparation activities discussed within the updated SQD framework (Tondeur et al., 2025). Because participants reported their pre-course experiences retrospectively, alongside their current self-assessments, prior familiarity with digital technologies, opportunities for practical application, and confidence in completing the questionnaire remain possible influences on the observed associations.

4.3. Teamwork and Intended Future Use as Supplementary Evidence with Defined Limitations

The adjusted association involving teamwork was positive but characterised by considerable statistical uncertainty (β = 0.136, 95% CI [-0.107, 0.379]), while 32.0% of eligible respondents with observed teamwork scores reported the maximum composite rating. The confidence interval accommodates associations ranging from slightly negative to moderately positive, meaning that the findings neither demonstrate a clear relationship nor confirm its absence. An overall teamwork measure captures participants' perceptions of communication, coordination, and mutual support (Hoegl and Gemuenden, 2001), rather than directly assessing how students critically examine one another's pedagogical decisions. The qualitative accounts describing reciprocal checking, duplicated responsibilities, and subsequent renegotiation of tasks provide more specific illustrations of these collaborative processes. Theoretical perspectives concerning social interdependence and shared regulation of learning help explain why the quality of such interactions may be educationally relevant (Johnson and Johnson, 2009; Järvelä and Hadwin, 2013). Similarly, the differentiated approaches to collaboration examined by Admiraal et al. (2026) highlight the importance of considering how group members engage with alternative ideas. These theoretical connections support interpretation of the written accounts without attributing collaborative processes that were not directly assessed by the questionnaire. The distinction maintained during coding between task distribution and perceived fairness further illustrates this limitation: describing how responsibilities were allocated provides evidence of organisational arrangements, whereas evaluating their fairness requires an explicit statement from the respondent. Likewise, reporting that assistance was received and explaining what was learned through that assistance represent different forms of information.

The 73 discrepancies between participants' self-reported project formats and administrative classifications did not alter the project reference used for Q20 and Q24, as confirmed by the original investigator. Administrative records were used to determine the applicability of teamwork measures, given that 82 project units represented independently completed work. Consequently, the supplementary teamwork analysis was restricted to 50 respondents from collaborative projects who provided usable ratings, rather than examining differences between individual and group-based project formats. The original investigator also confirmed that the administratively identified collaborative groups completed separate projects, supporting the use of these project units for statistical clustering.

Ratings of intended future technology use were strongly concentrated at the highest response category (59.3%), while the adjusted relationship between project-specific reflection and future-use intention remained statistically uncertain. The linked qualitative examples provide further clarification of the different intentions represented by these ratings. Maximum scores were associated with expectations of reusing or modifying existing resources, a relatively favourable score accompanied plans for personal study, and a low score was linked to an expressed intention not to use digital technologies in classroom teaching. Consequently, Q20j encompasses a wider range of anticipated applications than the transfer of learning into classroom practice, understood as the generalisation and continued application of acquired knowledge or skills in another setting (Baldwin and Ford, 1988; Blume et al., 2010). The theory of planned behaviour also distinguishes an individual's intention to perform an action from the actual enactment of that behaviour (Ajzen, 1991). A combined measure incorporating perceived applicability, understanding, and satisfaction therefore cannot serve as an equivalent substitute for the direct question concerning future-use intention. Findings reported by Connolly et al. (2023), indicating limited transfer within a related inquiry- and problem-based teacher education context, provide a useful contrast to favourable intentions expressed at the conclusion of a course. The linked qualitative cases consequently identify possible areas for subsequent investigation rather than establishing the underlying reasons for participants' intended future practices.

4.4 The Contribution of a Sustainability-Oriented Interpretation

A sustainability-oriented perspective emphasises the basis on which digital pedagogical practices may be appropriately continued beyond the completion of an assessed educational activity. Sustainable assessment focuses on developing judgement that can inform future learning (Boud and Soler, 2016), while UNESCO's agenda for strengthening educators' capacities (UNESCO, 2020) and an ethically informed understanding of professional digital competence (Falloon, 2020) establish the relevance of this consideration within teacher education. In the present study, the written responses identify specific areas requiring such judgement, including educational purpose, accessibility, image licensing, and the potential adaptation of digital resources. Their contribution lies in clarifying what future evaluations might investigate, particularly how pedagogical decisions are reconsidered in response to different learners and educational contexts. However, these accounts do not establish participants' ability to maintain such evaluative judgement over time. The coding review further distinguishes between identifying a professional development objective, proposing an action intended to achieve that objective, and explicitly planning to assess its outcomes. For example, identifying a particular course or seminar may constitute a concrete intended action without clarifying the specific competence to be developed. Similarly, proposing to trial an educational resource in a classroom may identify an intended context of application without explaining how its pedagogical effectiveness would be evaluated. These distinctions concern the information explicitly provided in participants' written accounts rather than representing a measure of their professional competence.

This distinction is further illustrated by the linked qualitative examples. Participants E04 and E05 reported the highest future-use intention ratings while describing plans to reuse or adapt educational materials, whereas E02 and E03 expressed intentions to consider accessibility requirements and image licensing conditions. However, the continued availability or reuse of a digital resource does not independently establish its ongoing pedagogical suitability, accessibility, or compliance with relevant usage conditions (UNESCO, 2019; Fawns, 2022). From this perspective, the findings identify questions for subsequent investigation, including which educational resources were reused, what modifications were introduced, and how these decisions were justified in different educational contexts. The present study did not examine the subsequent implementation of these intentions, reductions in preparation workload, or measurable environmental benefits (Selwyn, 2021). Unlike intervention-based evaluations such as that conducted by Yao and Abd Halim (2026), this investigation contributes by identifying aspects of professional judgement and anticipated application that require further empirical examination.

4.5 Implications for Educational Practice and Evaluation

One practical implication is the importance of making the reasoning underlying the development of digital educational products accessible for examination. Project instructions could require students to identify a specific learning need, provide a rationale for their design decisions, consider a feasible alternative, and describe modifications introduced in response to feedback. This approach brings together the attentive and exploratory dimensions of reflection described by Nguyen et al. (2014) with the principles of constructive alignment and formative feedback (Biggs, 1996; Nicol and Macfarlane-Dick, 2006). The updated SQD framework positions these activities alongside instructional modelling, supported design processes, and authentic learning experiences (Tondeur et al., 2025). These recommendations represent teaching considerations informed by the existing literature and are consistent with the associations identified in the present study. However, determining their effectiveness would require systematic documentation of the instructional support provided, together with direct examination of students' reasoning processes and subsequent design revisions.

Plans for the future application of digital educational resources could also identify the intended educational environment, target learners, necessary modifications, and procedures for reviewing accessibility or usage rights. Intended applications supporting personal study should be distinguished from those involving classroom teaching, while actual implementation should be evaluated independently of stated intentions. The administrative assessment scores did not demonstrate clear adjusted associations with any of the five principal self-report measures, with all five Holm-adjusted p-values equal to 1.000. Existing course assessment rubrics incorporate reflective documentation and student engagement alongside criteria relating to the completed educational artefacts. Although the limited convergence does not establish that the self-report measures lack validity, administrative assessment scores cannot independently validate the digital index as an indicator of demonstrated pedagogical performance. Future evaluations could employ comparable assessment criteria for educational artefacts, examine students' explanations of design decisions and documented revisions,

and investigate how resources are subsequently adapted in authentic teaching contexts. Portfolios and reflective reports offer relevant methodological approaches for examining these processes (Haşlaman et al., 2024). Similarly, written professional development plans could distinguish the intended competence, the proposed practical action, and the evidence used to evaluate its implementation. These recommendations are informed by the study's findings and existing literature rather than representing educational interventions whose effectiveness was established through the present investigation.

4.6 Study Limitations

This investigation employed an exploratory, cross-sectional research design within a single institution and did not include a prospectively collected pre-project baseline, a comparison group, or subsequent follow-up measurements. Differences in course-group sizes also involved variations in academic programmes, prior experience, and project formats that could not be fully equalised through statistical adjustment for course membership. The locally modified measurement indices have not been demonstrated to possess measurement invariance or to represent separate underlying latent constructs. Furthermore, the use of a common reporting source, evaluatively framed questionnaire statements, and overlapping item content may have contributed to the observed covariance between measures. The original investigator's simultaneous responsibilities as course leader, assessor, and researcher may also have influenced participants' decisions to participate and the responses they provided. Although participation in the research was voluntary and declining consent had no implications for grades, reflective activities formed a compulsory component of the course. Information concerning instructional support was derived from the instructor's descriptions and course documentation rather than independently observed delivery. The qualitative accounts provide clarification of participants' intended meanings but do not constitute independent evidence of implemented practices or the depth of reflective engagement. The subsequent coding exercises employed different analytical frameworks, with their outputs reconciled and checked by the author. Accordingly, this procedure should not be regarded as independent double coding by human researchers or as providing an estimate of intercoder reliability. Similarly, the absence of a particular feature from a short written response cannot be taken as evidence that the participant lacks the corresponding ability or intention.

The availability of digital self-rating scores was determined by the number of questionnaire items completed, meaning that individual respondents could contribute different combinations of items to the analysis. Although the wording-based diagnostic maintained a consistent analytical sample, it altered the pedagogical content of the item groups alongside their wording and therefore cannot be considered an experiment isolating wording effects. The reasons underlying previously missing questionnaire responses could not be determined. Coverage of teamwork measures was selective, while the concentration of responses at the upper end of the scales reduced their capacity to distinguish between respondents. Moreover, analytical specifications and diagnostic procedures were developed with awareness of the observed data and statistical model outcomes. Taken together, these limitations indicate the need for prospective investigations conducted across multiple educational contexts,

incorporating comparable evidence of pedagogical reasoning, the quality of developed educational artefacts, and their subsequent application in practice.

5. Conclusions

This exploratory study examined the relationships between project-based reflection, perceived learning conditions, and self-assessed digital pedagogical practices in teacher education. Project-specific reflection was positively associated with the combined digital self-rating index and with practice-oriented self-ratings examined through the wording-based diagnostic. These conditional associations provided additional information beyond general reflective orientations and perceived project conditions. However, the findings do not establish improvements in digital pedagogical competence or demonstrate statistically significant differences between associations involving practice-oriented and capability-oriented questionnaire items. The qualitative findings further distinguished participants' intentions to reuse existing educational resources, adapt digital materials, and apply technologies to personal learning or future teaching activities. From an educational sustainability perspective, these findings highlight the importance of examining how digital educational resources are evaluated, modified, and applied beyond the completion of assessed coursework. Future research should therefore investigate demonstrated pedagogical practices, the reasoning informing resource adaptation, and the continued suitability of digital materials across different educational contexts. The present study did not assess sustained implementation, changes in professional competence, or measurable environmental benefits.

Declarations

Author Contributions: Conceptualisation, S.S.D. and G.N.; methodology, S.S.D. and G.N.; formal analysis, S.S.D. and G.N.; validation, S.S.D. and G.N.; writing-original draft preparation, S.S.D.; writing-review and editing, S.S.D. and G.N.; visualisation, S.S.D. and G.N.; project administration, S.S.D. Both authors have reviewed and approved the final version of the manuscript.

Funding: This research received no external funding. The preparation and development of the manuscript were completed without financial support from public, commercial or not-for-profit funding organisations.

Institutional Review Board Statement: The original empirical investigation was conducted in an academic context involving Alexander College, Cyprus, and Canterbury Christ Church University (CCCU), UK. The ethics application was submitted on 4 January 2026, and formal ethical approval was granted by the responsible institutional ethics committee under Decision No. 24/2026 on 8 April 2026. Questionnaire responses were collected between 4 May and 14 September 2026, following the granting of ethical approval. The present manuscript examines previously collected empirical data and did not involve additional participant recruitment or new empirical data collection.

Informed Consent Statement: The original empirical investigation reported obtaining participants' informed consent for the research use of anonymised or pseudonymised questionnaire responses through Question 1 (Q1). Separate permission was requested through Question 2 (Q2) for the publication

of anonymised quotations from open-ended responses. Participation in the research was voluntary, and declining consent did not affect participants' academic assessments, course obligations, or grades. Although reflective activities formed part of the course requirements, the inclusion of participants' responses in the research was subject to their consent. The present manuscript uses previously collected data, subject to verification that the original consent conditions permit the present analysis and publication.

Consent for Publication: The original study obtained separate permission for the publication of anonymised extracts from participants' written responses. The present manuscript does not disclose directly identifiable participant information. Any reproduction or adaptation of the original qualitative evidence is subject to the applicable consent conditions, licensing and permissions. Both authors must review and approve the final manuscript before submission.

Data Availability Statement: The present manuscript is based on previously collected empirical data from the original investigation, which were reused and reanalysed with permission from the original investigators. Individual-level questionnaire responses, complete written responses, and participant-linked qualitative coding records are not publicly available due to participant confidentiality, privacy considerations, and applicable data-protection requirements. Supporting statistical results are provided in Appendix A (Supplementary File S1: Statistical Tables). Additional research documentation is retained in a restricted-access research repository, subject to the access arrangements established by the original investigators.

Supplementary Materials: The following supplementary material accompanies this manuscript: Appendix A (Supplementary File S1: Statistical Tables), containing supporting statistical results associated with the quantitative analyses reported in the study. Individual-level questionnaire responses, complete qualitative responses, participant-linked coding documentation, and internal analytical scripts are not included in the publicly available supplementary materials.

Acknowledgements: The authors acknowledge the contributions of the original investigators responsible for the empirical research, questionnaire development, data collection, statistical analyses, and supporting research documentation upon which the present manuscript draws. Their work provided an important methodological and empirical foundation for the present investigation. The authors also acknowledge the permission granted to reuse and reanalyse the original research data. The original investigation reported the use of InstaText for proofreading, DeepL Pro for preliminary translation, and ChatGPT for assistance with analytical code preparation. These disclosures relate exclusively to the original investigation and do not necessarily describe the tools used in preparing the present manuscript.

Conflicts of Interest: The authors declare no competing financial or non-financial interests that could have influenced the preparation, interpretation or presentation of this manuscript, subject to confirmation by both authors before submission.

Code Availability: The analytical scripts, computational records, and qualitative coding documentation associated with the underlying investigation are not included in the supplementary materials accompanying this manuscript. These materials are retained in a restricted-access research repository and remain subject to the applicable institutional data-protection and access requirements. Appendix A (Supplementary File S1: Statistical Tables) contains the supporting statistical results designated for submission.

Clinical Trial Registration: Not applicable. The investigation concerns educational practices, reflective learning experiences and self-reported digital pedagogical measures and does not involve clinical trials, medical interventions or clinical research procedures.

References

1. Admiraal W, Saab N, Guo P (2026) Online collaborative project-based learning in higher education: Students' motivation, collaborative learning, and perceived outcomes. *Active Learning in Higher Education*. Advance online publication. https://doi.org/10.1177/14697874261419448
2. Ajzen I (1991) The theory of planned behavior. Organ Behav Hum Decis Process 50:179–211. https://doi.org/10.1016/0749-5978(91)90020-T
3. Baldwin TT, Ford JK (1988) Transfer of training: A review and directions for future research. Pers Psychol 41:63–105. https://doi.org/10.1111/j.1744-6570.1988.tb00632.x
4. Bernsteiner A, Haagen-Schützenhöfer C, Schubatzky T (2025) Teacher education in the age of digitality: Conclusions from a design-based research project. Eur J Educ 60 Article e12904. https://doi.org/10.1111/ejed.12904
5. Biggs J (1996) Enhancing teaching through constructive alignment. High Educ 32:347–364. https://doi.org/10.1007/BF00138871
6. Blömeke S, Gustafsson J-E, Shavelson RJ (2015) Beyond dichotomies: Competence viewed as a continuum. Z Psychol 223:3–13. https://doi.org/10.1027/2151-2604/a000194
7. Blume BD, Ford JK, Baldwin TT, Huang JL (2010) Transfer of training: A meta-analytic review. J Manag 36:1065–1105. https://doi.org/10.1177/0149206309352880
8. Blumenfeld PC, Soloway E, Marx RW, Krajcik JS, Guzdial M, Palincsar A (1991) Motivating project-based learning: Sustaining the doing, supporting the learning. Educational Psychol 26:369–398. https://doi.org/10.1080/00461520.1991.9653139
9. Boud D, Soler R (2016) Sustainable assessment revisited. Assess Evaluation High Educ 41:400–413. https://doi.org/10.1080/02602938.2015.1018133
10. Boud D, Walker D (1998) Promoting reflection in professional courses: The challenge of context. Stud High Educ 23:191–206. https://doi.org/10.1080/03075079812331380384
11. Colás-Bravo P, Conde-Jiménez J, Reyes-de-Cózar S (2021) Sustainability and digital teaching competence in higher education. Sustainability 13(22). Article 12354. https://doi.org/10.3390/su132212354

12. Connolly C, Logue PA, Calderón A (2023) Teaching about curriculum and assessment through inquiry and problem-based learning methodologies: An initial teacher education cross-institutional study. Ir Educational Stud 42:443–460. https://doi.org/10.1080/03323315.2021.2019083
13. Falloon G (2020) From digital literacy to digital competence: The teacher digital competency (TDC) framework. Education Tech Research Dev 68:2449–2472. https://doi.org/10.1007/s11423-020-09767-4
14. Fawns T (2022) An entangled pedagogy: Looking beyond the pedagogy–technology dichotomy. Postdigital Sci Educ 4:711–728. https://doi.org/10.1007/s42438-022-00302-7
15. Fetters MD, Curry LA, Creswell JW (2013) Achieving integration in mixed methods designs—Principles and practices. Health Serv Res 48:2134–2156. https://doi.org/10.1111/1475-6773.12117
16. Guo P, Saab N, Post LS, Admiraal W (2020) A review of project-based learning in higher education: Student outcomes and measures. Int J Educational Res 102:101586. https://doi.org/10.1016/j.ijer.2020.101586
17. Haşlaman T, Uslu A, N., Mumcu F (2024) Development and in-depth investigation of pre-service teachers' digital competencies based on DigCompEdu: A case study. Qual Quant 58:961–986. https://doi.org/10.1007/s11135-023-01674-z
18. Hatton N, Smith D (1995) Reflection in teacher education: Towards definition and implementation. Teach Teacher Educ 11:33–49. https://doi.org/10.1016/0742-051X(94)00012-U
19. Hoegl M, Gemuenden HG (2001) Teamwork quality and the success of innovative projects: A theoretical concept and empirical evidence. Organ Sci 12:435–449. https://doi.org/10.1287/orsc.12.4.435.10635
20. Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70. https://www.jstor.org/stable/4615733
21. Hsieh H-F, Shannon SE (2005) Three approaches to qualitative content analysis. Qual Health Res 15:1277–1288. https://doi.org/10.1177/1049732305276687
22. Järvelä S, Hadwin AF (2013) New frontiers: Regulating learning in CSCL. Educational Psychol 48:25–39. https://doi.org/10.1080/00461520.2012.748006
23. Johnson DW, Johnson RT (2009) An educational psychology success story: Social interdependence theory and cooperative learning. Educational Researcher 38:365–379. https://doi.org/10.3102/0013189X09339057
24. Kaldi S, Zafeiri S (2023) Identifying reflective modes in pre-service teachers' written reflections on the implementation of project-based learning during the school practicum. Front Educ 7:1072090. https://doi.org/10.3389/feduc.2022.1072090
25. Kember D, Leung DYP, Jones A, Loke AY, McKay J, Sinclair K, Tse H, Webb C, Wong FKY, Wong M et al (2000) Development of a questionnaire to measure the level of reflective thinking. Assess Evaluation High Educ 25:381–395. https://doi.org/10.1080/713611442
26. MacKinnon JG, Nielsen MØ, Webb MD (2023) Cluster-robust inference: A guide to empirical practice. J Econ 232:272–299. https://doi.org/10.1016/j.jeconom.2022.04.001

27. Momdjian L, Manegre M, Gutiérrez-Colón M (2025) A study of preservice teachers' digital competence development: Exploring the role of direct instruction, integrated practice, and modeling. Eval Program Plan 109:102538. https://doi.org/10.1016/j.evalprogplan.2025.102538
28. Nguyen QD, Fernandez N, Karsenti T, Charlin B (2014) What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. Med Educ 48:1176–1189. https://doi.org/10.1111/medu.12583
29. Nicol DJ, Macfarlane-Dick D (2006) Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Stud High Educ 31:199–218. https://doi.org/10.1080/03075070600572090
30. Podsakoff PM, Podsakoff NP, Williams LJ, Huang C, Yang J (2024) Common method bias: It's bad, it's complex, it's widespread, and it's not easy to fix. Annual Rev Organizational Psychol Organizational Behav 11:17–61. https://doi.org/10.1146/annurev-orgpsych-110721-040030
31. Quast J, Rubach C, Porsch R (2025) Professional digital competence beliefs of student teachers, pre-service teachers and teachers: Validating an instrument based on the DigCompEdu framework. Eur J Teacher Educ 48:698–721. https://doi.org/10.1080/02619768.2023.2251663
32. Redecker C (2017) In: Punie Y (ed) European framework for the digital competence of educators: DigCompEdu. Publications Office of the European Union. https://doi.org/10.2760/159770
33. Selwyn N (ed) (2021) Ed-tech within limits: Anticipating educational technology in times of environmental crisis. *E-Learning and Digital Media, 18*, 496–510. https://doi.org/10.1177/20427530211022951
34. Sjølie E, Rødsjø E, Skancke L, Carvajal P, Solhaug S, Jensen MR, Kaalvik H (2026) The empirical base of student project collaboration in higher education: A mapping review. Rev Educ Res. Advance online publication https://doi.org/10.3102/00346543261435632
35. Smestad B, Hatlevik OE, Johannesen M, Øgrim L (2023) Examining dimensions of teachers' digital competence: A systematic review pre- and during COVID-19. Heliyon 9 Article e16677. https://doi.org/10.1016/j.heliyon.2023.e16677
36. Tiainen O, Lutovac S, Korkeamäki R-L (2024) Rethinking approaches to reflection in initial teacher education. Educational Res 66:245–262. https://doi.org/10.1080/00131881.2024.2346088
37. Tondeur J, Aesaert K, Prestridge S, Consuegra E (2018) A multilevel analysis of what matters in the training of pre-service teacher's ICT competencies. Comput Educ 122:32–42. https://doi.org/10.1016/j.compedu.2018.03.002
38. Tondeur J, Trevisan O, Howard SK, van Braak J (2025) Preparing preservice teachers to teach with digital technologies: An update of effective SQD-strategies. Comput Educ 232:105262. https://doi.org/10.1016/j.compedu.2025.105262
39. Tondeur J, van Braak J, Sang G, Voogt J, Fisser P, Ottenbreit-Leftwich A (2012) Preparing pre-service teachers to integrate technology in education: A synthesis of qualitative evidence. Comput Educ 59:134–144. https://doi.org/10.1016/j.compedu.2011.10.009

40. UNESCO (2019) *Recommendation on open educational resources (OER).* https://www.unesco.org/en/legal-affairs/recommendation-open-educational-resources-oer
41. UNESCO (2020) *Education for sustainable development: A roadmap.* https://doi.org/10.54675/YFRE1448
42. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative (2007) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. PLoS Med 4:e296. https://doi.org/10.1371/journal.pmed.0040296
43. Yao N, Abd Halim ND (2026) Enhancing STEM teachers' technology integration competencies to support education for sustainable development. Sustainability 18 Article 2520. https://doi.org/10.3390/su18052520
44. Young K, Doğança Küçük Z, Delahunty T, Dempsey M, Maglaperidze N (2026) Evaluating the impact of a digital competency-based placement model on STEM pre-service teachers' digital competence and teaching experiences. Teach Teacher Educ 175:105466. https://doi.org/10.1016/j.tate.2026.105466
45. Zhang L, Yang C, Zheng Y (2026) Digital competence for sustainable education of pre-service teachers: A systematic literature review (2014–2024). Front Psychol 16:1710983. https://doi.org/10.3389/fpsyg.2025.1710983
46. Žerovnik A (2026) *Supplementary files for project reflection and digital pedagogical self-ratings in teacher education* [Data set and supplementary materials]. Zenodo. https://doi.org/10.5281/zenodo.22875121

Tables

Tables 1 to 4 are available in the Supplementary Files section.

Figures

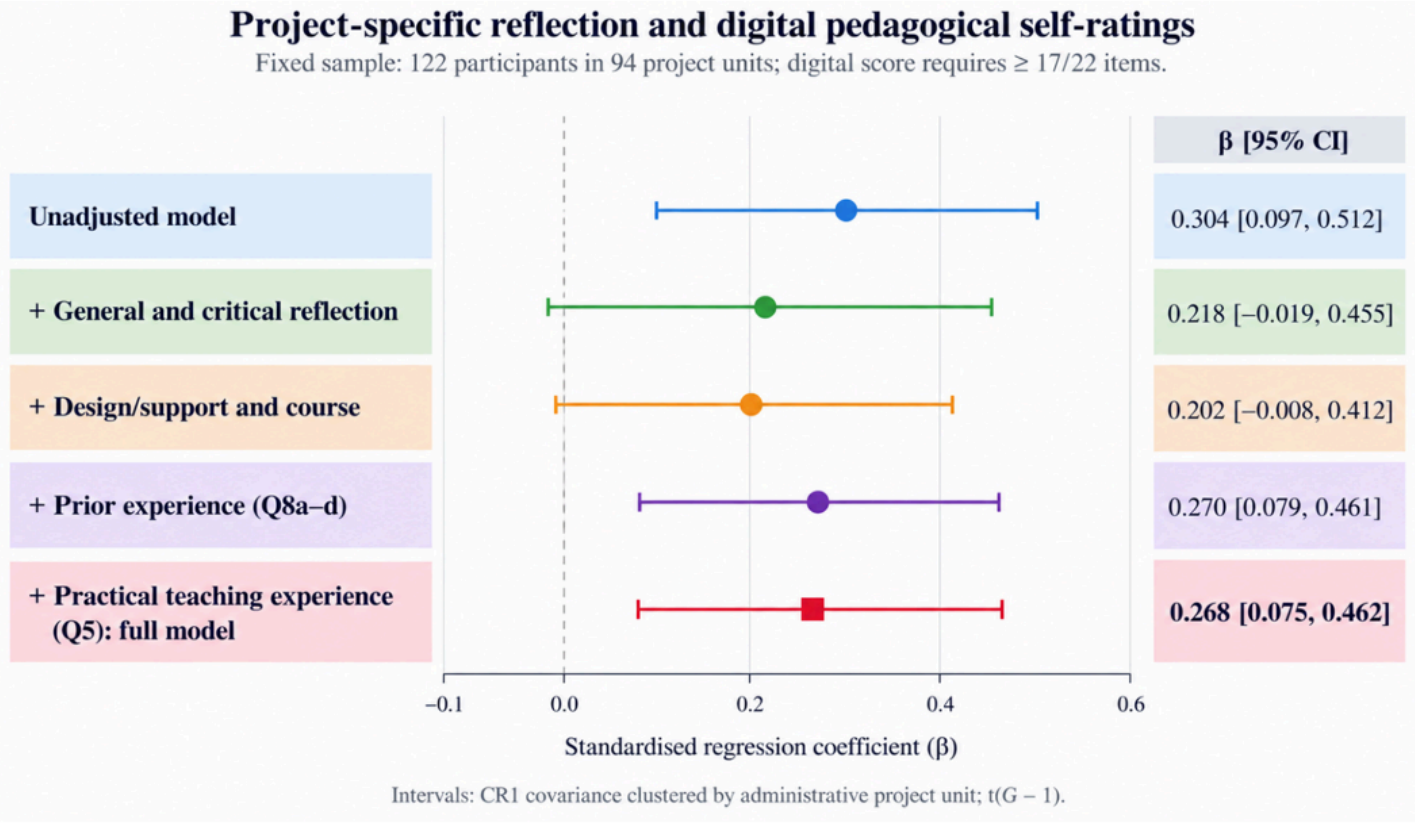


Figure 1

Relationship Between Project-Specific Reflection and Digital Pedagogical Practice/Capability Self-Ratings Across Sequentially Adjusted Models. The plotted points represent standardised regression coefficient estimates, while the horizontal error bars display 95% confidence intervals calculated using the CR1 cluster-robust covariance estimator, with clustering at the administrative project-unit level and inference based on t(93). Each model was estimated using the same fixed sample of 122 respondents distributed across 94 project units, with consistent standardisation applied within the analytical sample. Adjustment variables were introduced progressively, with each successive model retaining those included in the preceding specification. The fully adjusted model incorporated practical teaching experience (Q5) in addition to four separately measured indicators of retrospectively reported pre-course experience (Q8a–d). These analyses represent retrospective comparisons of statistical associations under different adjustment specifications and should not be interpreted as demonstrating a causal progression or causal relationships.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- S2.StatisticalTables.xlsx
- Appendices.docx

- Tables.docx