Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno 6 n. 3 (2026)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

Metacognitive Engagement and Professional Agency in Teachers' Digital Education: Findings from a Qualitative Study

Flavia Capodanno [1] and Alessio Di Paolo [1] *

[1] Department of Humanities, Philosophy, and Education, University of Salerno; fcapodanno@unisa.it ; adipaolo@unisa.it
* Correspondence: fcapodanno@unisa.it; adipaolo@unisa.it

Abstract: Engagement in teachers' digital education cannot be inferred from participation, satisfaction, or tool use alone. This study examines metacognitive engagement as a transversal regulatory quality through which teachers become aware of learning, monitor limitations, formulate pedagogical criteria, and transfer digital competences into practice. A second-level qualitative analysis was conducted on a teacher education programme previously examined quantitatively. The analytic sample comprised 118 in-service teachers and 3,540 paired written rationales associated with fifteen SELFIE-inspired self-assessment items. Participant trajectories across pre-programme and post-programme responses were analysed through a longitudinal framework integrating awareness, planning, monitoring, evaluation, regulation, transfer, objects of reflection, levels of elaboration, and pedagogical functions of technology. Reanalysis of the paired quantitative data provided contextual evidence of perceived growth, with the overall mean increasing from 3.01 to 3.77. Qualitative findings showed substantial heterogeneity in reflective density. Expanded digital repertoires were common, whereas explicit metacognitive language was comparatively infrequent. The results indicate that score improvement and enthusiasm should not be treated as proxies for metacognitive engagement, which becomes visible when teachers articulate criteria, regulate action, recognise limits, and connect technological choices with student learning and professional agency.

Keywords: metacognitive engagement; teacher agency; digital education; teacher professional development; qualitative longitudinal analysis

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction[1]

Contemporary debates on engaging education increasingly require a distinction between visible participation and the deeper regulatory processes through which learners and professionals orient their activity, interpret feedback, revise strategies, and decide what should be carried forward into subsequent situations. Behavioural persistence, emotional involvement, and cognitive investment remain analytically valuable dimensions of engagement, yet none of them alone establishes whether an

[1] The paper is the result of scientific collaboration between the authors; however, the attribution of the paragraphs is as follows: Flavia Capodanno is the author of paragraphs 1, 6 and co-author of paragraphs 4, 5; Alessio Di Paolo is the author of paragraphs 2,3,5 and co-author of paragraphs 4 and 5.

individual understands the conditions of his or her own learning or can deliberately regulate participation in relation to an educational purpose (Fredricks et al., 2004; Thomas et al., 2022).
This distinction becomes particularly consequential in digital teacher education, where attendance, completion, positive evaluations, and the frequency of technology use may coexist with limited reflection on why a tool is selected, how its affordances are interpreted, what evidence is used to judge its pedagogical value, and when its use should be modified or discontinued.

The present study therefore conceptualises metacognitive engagement not as a fourth dimension placed alongside behavioural, emotional, and cognitive engagement, but as a transversal quality through which engagement becomes self-observing and regulative. In this sense, engagement acquires metacognitive depth when participants can identify what they know, distinguish mastery from familiarity, monitor uncertainty, plan subsequent action, evaluate the adequacy of a strategy, and transfer criteria rather than merely reproduce procedures. Recent work on digital learning has shown that metacognitive activity can differentiate modes of engagement that would otherwise appear similar at the level of observable behaviour, particularly when digital environments increase learner control and require ongoing monitoring of understanding (Omarchevska et al., 2026). Such a perspective also prevents the educational value of technology from being inferred from efficiency or novelty, since the use of generative systems may increase productivity while simultaneously reshaping cognition, metacognition, and epistemic agency in ways that require explicit pedagogical examination (Yan et al., 2025).
Professional agency provides the second conceptual axis of the study. Agency is treated neither as individual voluntarism nor as an abstract personal disposition, but as the capacity to interpret a situation, recognise available and constrained courses of action, make pedagogically defensible choices, and assume responsibility for their consequences within a situated professional environment. In teacher education, such agency develops through the interaction between perceptions, professional knowledge, institutional conditions, and opportunities for reflective action, with particular relevance for inclusive practice because educational decisions must remain responsive to differences among learners rather than being delegated to standardised routines or technological prescriptions (Aiello et al., 2016; Aiello, 2019, 2024; Sibilio & Aiello, 2018).

The relation between metacognition and agency is especially productive for interpreting professional learning. Self-regulation depends on the reciprocal organisation of metacognitive knowledge, a sense of agency, and representations of possible future selves, which together support goal selection, monitoring, behavioural adjustment, and transfer (Frazier et al., 2021). Agency itself can become an object of metacognitive judgement, since individuals monitor the degree to which an outcome is attributable to their own action and revise their sense of control as conditions change (Metcalfe & Greene, 2007; Metcalfe et al., 2010). From this standpoint, professional growth cannot be reduced to feeling more competent. It is more adequately observed when increased confidence is accompanied by a more differentiated capacity to formulate criteria, identify limits, anticipate consequences, and act on learning environments with a reasoned sense of responsibility (Broncano, 2014; Gao & Zhang, 2011).

These premises inform a second-level analysis of a thirty-hour in-service teacher education programme that integrated Artificial Intelligence, the production

of digital resources, and musical languages within a metadisciplinary design. The programme had previously been examined primarily through quantitative self-assessment data and an initial qualitative reading, with statistically significant increases reported across fifteen areas of perceived digital competence and with the final paper explicitly identifying the need for a more systematic analysis of the written rationales (Di Paolo et al., 2025). The present study addresses that analytical need by shifting the object of inquiry from the magnitude of perceived improvement to the meanings through which teachers narrated change, uncertainty, pedagogical intentionality, and the regulation of future practice.

Four questions guided the analysis. The first concerned how teachers described changes in their digital competences across the paired pre-programme and post-programme rationales. The second examined which forms of awareness, planning, monitoring, evaluation, regulation, and transfer became visible in the written corpus. The third investigated whether increased self-assessment was accompanied by the articulation of pedagogical criteria and a more agentic relationship with digital technology. The fourth examined the conditions that appeared to facilitate or hinder movement from instrumental engagement, centred on acquiring applications and procedures, towards reflective and agentic engagement, centred on decisions, purposes, consequences, and student learning.

2. Presentation of a qualitative study

2.1. Research design and relation to the previous study

The research was designed as a qualitative longitudinal secondary analysis embedded within a previously reported mixed-method teacher education study. The quantitative component of the earlier work is not treated here as a new primary outcome. Instead, the original participant-level ratings are re-estimated from the archived dataset to establish an auditable empirical context for interpreting the qualitative trajectories. This separation between the earlier quantitative question and the present qualitative question is methodologically relevant because the current study asks not whether scores increased, but what kind of reflective, regulatory, and agentic meaning can be identified beneath or alongside those increases (Di Paolo et al., 2025).

The raw export used for the present analysis contained 118 complete cases, each with fifteen pre-programme ratings, fifteen post-programme ratings, and a written rationale associated with every rating. All denominators and descriptive statistics reported in the current paper were recalculated directly from that archived file rather than inherited from narrative summaries in the previous publication. Because the same participants contributed data at both time points, the quantitative contextual analysis was reconstructed using paired-sample procedures. This choice corresponds to the dependency structure of the data and replaces the independent-samples specification reported in the earlier article without altering the substantive role of the quantitative segment, which remains contextual rather than explanatory.

2.2. Participants

The final analytic sample comprised 118 in-service teachers from schools located across the Campania region of Italy. Participation in the professional-development programme was voluntary following recruitment through school

networks and institutional channels. The sample included 57 primary school teachers, 29 upper-secondary school teachers, 22 lower-secondary school teachers, and 10 early childhood education teachers. The qualification profile was heterogeneous, with 58 participants reporting a master's degree, 37 an upper-secondary diploma, seven a postgraduate specialisation course, seven a master-level postgraduate qualification, five a doctorate, and four a bachelor's degree. No claims of statistical representativeness are made, since the sample was self-selected and territorially bounded.

Identifying information contained in the administrative export was not used during analysis. A de-identified analytic copy was created and each participant was assigned a sequential code from P001 to P118. The participant, rather than the individual written rationale, constituted the primary interpretive unit, thereby preserving the longitudinal relation among the thirty rating-rationale combinations provided by each teacher and preventing repeated statements from the same person from being treated as independent qualitative observations.

Table 1. Characteristics of the analytic sample.

Dimension	Category	n	%
School level	Primary school	57	48.3%
School level	Upper-secondary school	29	24.6%
School level	Lower-secondary school	22	18.6%
School level	Early childhood education	10	8.5%
Highest qualification	Master's degree	58	49.2%
Highest qualification	Upper-secondary diploma	37	31.4%
Highest qualification	Postgraduate specialisation course	7	5.9%
Highest qualification	Postgraduate master qualification	7	5.9%
Highest qualification	Doctorate	5	4.2%
Highest qualification	Bachelor's degree	4	3.4%

2.3. Timing, training phases, and learning environment

The professional-development programme comprised thirty hours and was organised through a blended-learning format combining synchronous or in-person activities with online work on a dedicated platform. The archived response file spans 1 April to 29 July 2025, which constitutes the auditable data-collection window available for this secondary analysis. The programme was structured into three ten-hour components whose sequence was intended to move from conceptual and

operational familiarisation towards collaborative design and, finally, towards explicit reflection on learning and future practice.

The first component combined theoretical and practical activities concerning digital education, Artificial Intelligence, digital resource production, and the pedagogical use of musical languages. The second component involved co-design workshops in which teachers explored the construction and adaptation of digital artefacts and interdisciplinary activities. The third component was devoted to metacognitive reflection on choices, difficulties, resources, possible classroom use, and the transfer of learning to professional contexts. Music operated throughout the programme as both an expressive language and a metadisciplinary mediator, while Artificial Intelligence was treated as an object of pedagogical inquiry rather than as a self-justifying innovation (Di Paolo et al., 2025).

Table 2. Structure of the thirty-hour teacher education programme.

Phase	Duration	Main focus	Regulatory function
Theoretical and practical work	10 h	Digital education, Artificial Intelligence, digital resources, musical languages	Conceptual orientation and guided experimentation
Co-design workshops	10 h	Collaborative design of digital and interdisciplinary learning activities	Application, comparison of choices, peer exchange
Metacognitive reflection	10 h	Review of difficulties, resources, criteria, and possible transfer to practice	Monitoring, evaluation, regulation, and forward planning

2.4. Instrument

Data were collected through a questionnaire inspired by SELFIE, the European self-reflection framework designed to support reflection on the integration of digital technologies in educational organisations. In the present programme the instrument was adapted to individual teacher self-assessment and therefore should not be interpreted as a standardised psychometric scale or as a direct measure of objectively demonstrated digital competence. It comprised fifteen domains rated on a five-point Likert scale, each followed by an open-ended prompt asking participants to justify the selected rating (European Commission, 2025; Di Paolo et al., 2025).

The fifteen areas covered the preparation of digital lessons and resources, technology-supported classroom teaching, assessment and feedback, the identification of online resources, the creation of digital materials, virtual learning environments, school communication, adaptation to individual learner needs, creativity, student engagement, collaboration, interdisciplinary projects, competence assessment, timely feedback, and the use of digital technology to support students' reflection on learning. The parallel pre-programme and post-programme structure produced fifteen paired quantitative ratings and fifteen paired rationales for each teacher. The open rationale was not treated as a decorative supplement to the rating. It was the primary source for examining whether participants could provide reasons, identify criteria, describe uncertainty, and articulate regulatory intentions beyond a numerical judgement.

Table 3. SELFIE-inspired areas used in the paired self-assessment.

Code	Area of self-assessment
A1	Preparing lessons by modifying or creating digital resources
A2	Technology-supported classroom teaching with devices and digital resources
A3	Assessment, personal feedback, and support for students
A4	Identifying online platforms for digital educational resources
A5	Creating digital resources for teaching practice
A6	Using virtual learning environments with students
A7	Using digital technologies for school communication
A8	Adapting teaching to individual learner needs through digital technologies
A9	Encouraging student creativity through digital technologies
A10	Designing digital learning activities intended to engage students
A11	Facilitating collaboration among students through digital technologies
A12	Supporting interdisciplinary projects through digital technologies
A13	Assessing student competences through digital technologies
A14	Providing timely feedback through digital technologies
A15	Helping students reflect on their own learning through digital technologies

2.5. Data collection and construction of the qualitative corpus

The questionnaire data were collected digitally and exported to a spreadsheet in which every participant occupied a single row. The analytic corpus consisted of 3,540 short written texts, including 1,770 pre-programme rationales and 1,770 post-programme rationales. This structure made it possible to reconstruct fifteen pre-post pairs for every participant. All written data were analysed in Italian, the language in which they were produced. Excerpts reported in the present English-language manuscript were translated for publication after coding so that interpretive decisions remained grounded in the original wording.

Before coding, the corpus was cleaned conservatively. Spacing and obvious formatting artefacts were normalised, while spelling, syntax, emphatic capitalisation, and brief formulations were retained because their form could itself carry information about the degree of elaboration. No attempt was made to enrich or grammatically correct participants' responses. The numerical score remained linked to its corresponding rationale, but the qualitative coding was not determined by whether the rating increased, decreased, or remained stable. This separation reduced the risk of circular interpretation in which an increased score would automatically be read as evidence of reflective development.

Table 4. Structure and textual density of the qualitative corpus.

Indicator	Value	Interpretive note
Participants	118	Complete paired trajectories
Items per time point	15	Five-point self-assessment plus rationale
Pre-programme rationales	1,770	15 per participant
Post-programme rationales	1,770	15 per participant

Indicator	Value	Interpretive note
Total written rationales	3,540	Primary qualitative corpus
Median words per rationale, pre	5	Tokenised words
Median words per rationale, post	4	Tokenised words
Post rationales of one or two words	635	35.9% of post rationales
Participants repeating one formulation across all 15 post rationales	12	10.2% of participants

2.6. Qualitative longitudinal framework analysis

A longitudinal framework analysis was selected because the research questions required systematic comparison across participants and time while preserving the internal coherence of each individual trajectory. Framework approaches are particularly useful when analysis must combine theoretically informed categories with inductive refinement and when a matrix structure is needed to compare cases without fragmenting them into decontextualised quotations (Gale et al., 2013). Coding therefore operated at two connected levels. Individual rationales were the local units to which codes were attached, whereas the participant trajectory across the fifteen paired items was the higher-order interpretive unit from which patterns of continuity, discontinuity, and transformation were reconstructed.

The coding framework combined deductive sensitising concepts with inductive categories generated during close reading. Four analytical axes were used. The first addressed the regulatory process and included awareness, planning, monitoring, evaluation, regulation, and transfer. The second identified the object of reflection and distinguished the professional self, the digital tool, teaching practice, student learning, and the organisational context. The third addressed the level of elaboration and differentiated declarative, operational, reflective, strategic-regulatory, and agentic formulations. The fourth identified the educational functions attributed to digital technology, including motivation, personalisation, collaboration, creativity, feedback, self-assessment, and critical thinking. Codes on different axes were non-mutually exclusive because a single response could simultaneously express uncertainty, refer to student learning, articulate a future plan, and identify a pedagogical function.

Analysis proceeded through repeated familiarisation with complete participant trajectories, independent coding of an initial calibration portion of the corpus, comparison among researchers to clarify code boundaries, revision of the codebook, systematic charting in a participant-by-domain matrix, and the preparation of analytical memos. Particular attention was given to disconfirming cases in which numerical improvement coexisted with weak rationale, confidence coexisted with uncertainty, or stable ratings coexisted with more sophisticated pedagogical reasoning. This procedure treated disagreement during calibration as a resource for sharpening definitions rather than as a nuisance to be removed mechanically, while maintaining an explicit audit trail of coding decisions and category revisions (O'Connor & Joffe, 2020; Saldaña, 2021).

A computer-assisted lexical mapping was used only as an exploratory orientation device. Keyword-in-context searches and participant-level presence indicators were generated for references to expanded digital repertoires, further

practice or regulation, pedagogical purposes, confidence, contextual constraints, and explicit metacognitive language. These counts were then checked against surrounding text and were never used as substitutes for interpretive coding. Their function was to reveal areas requiring closer reading and to support negative-case analysis. Frequency was calculated at participant level, meaning that repeated use of the same marker by one teacher counted once within a category. Because categories could overlap, percentages were not expected to sum to one hundred.

Table 5. Multiaxial coding framework for the qualitative analysis.

Analytical axis	Categories	Analytical question
Regulatory process	Awareness; planning; monitoring; evaluation; regulation; transfer	What the teacher does with knowledge about learning and action
Object of reflection	Professional self; digital tool; teaching practice; student learning; organisational context	What the reflection is about
Level of elaboration	Declarative; operational; reflective; strategic-regulatory; agentic	How deeply the rationale articulates criteria and intentional action
Function of technology	Motivation; personalisation; collaboration; creativity; feedback; self-assessment; critical thinking	What educational purpose is attributed to digital technology

2.7. Quantitative contextual analysis and statistical sensitivity checks

The quantitative data were reanalysed only to establish the empirical context in which the qualitative trajectories were interpreted. For each participant, a composite descriptive score was calculated as the arithmetic mean of the fifteen items at each time point. Item-level and composite pre-post differences were examined with paired-samples t tests, ninety-five percent confidence intervals for the mean change, and standardised paired effect sizes expressed as Cohen's dz. Because the items used five response categories and the distributional assumptions of parametric procedures can be questioned for individual Likert-type data, Wilcoxon signed-rank tests were used as sensitivity analyses. The inferential conclusions were unchanged. This dual reporting is consistent with recommendations to distinguish the scale properties of individual Likert items from the broader inferential question and to report effect magnitude alongside statistical significance (de Winter & Dodou, 2010; Lakens, 2013; Sullivan & Artino, 2013).

The family of fifteen item-level paired tests was adjusted with the Holm procedure to limit inflation of the familywise error rate. Internal consistency across the fifteen items was estimated descriptively with Cronbach's alpha at each time point, while recognising that a high coefficient cannot establish unidimensionality and is not presented as validation of the SELFIE-inspired adaptation. The quantitative analysis was conducted on complete paired cases, of which there were 118, and no imputation was necessary.

Table 6. Paired quantitative context for the fifteen SELFIE-inspired items.

Item	Pre M	Post M	Mean change	dz	95% CI change	Holm p
A1	3.12	3.86	0.74	0.95	0.60 to 0.88	< .001
A2	3.31	3.85	0.53	0.73	0.40 to 0.67	< .001
A3	3.22	3.50	0.28	0.32	0.12 to 0.44	.001
A4	3.02	4.08	1.07	1.09	0.89 to 1.25	< .001
A5	3.09	3.95	0.86	0.89	0.68 to 1.03	< .001
A6	2.96	3.66	0.70	0.79	0.54 to 0.87	< .001
A7	3.39	3.86	0.47	0.52	0.31 to 0.64	< .001
A8	3.08	3.73	0.64	0.80	0.50 to 0.79	< .001
A9	3.13	3.79	0.66	0.76	0.50 to 0.82	< .001
A10	2.98	3.81	0.82	0.92	0.66 to 0.98	< .001
A11	2.79	3.73	0.94	1.01	0.77 to 1.11	< .001
A12	2.99	3.80	0.81	0.90	0.64 to 0.97	< .001
A13	2.69	3.69	1.00	1.00	0.82 to 1.18	< .001
A14	2.67	3.65	0.98	1.13	0.82 to 1.14	< .001
A15	2.64	3.58	0.93	0.91	0.75 to 1.12	< .001

For the fifteen-item composite, the mean increased from 3.01 to 3.77, with a mean paired change of 0.76 points, 95% CI [0.65, 0.87], $t(117) = 13.45$, $p < .001$, $dz = 1.24$. The Wilcoxon sensitivity analysis likewise indicated a systematic positive shift, $p < .001$. Cronbach's alpha was .955 at pre-programme assessment and .971 at post-programme assessment. These values describe the internal coherence of the response set in this sample and are not treated as evidence that the fifteen domains constitute a single latent construct.

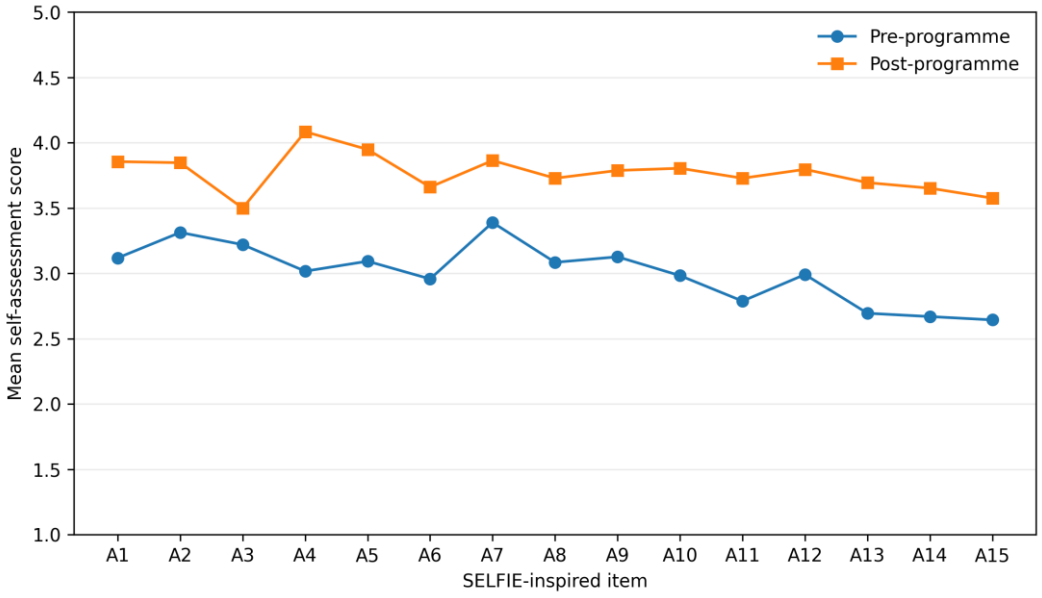


Figure 1. Pre-programme and post-programme mean self-assessment scores across the fifteen items. The figure is descriptive and the paired inferential analysis is reported in Table 6.

2.8. Trustworthiness, analytic transparency, and methodological safeguards

Several safeguards were incorporated to reduce overinterpretation. The first concerned analytic independence. Repeated rationales from the same teacher were nested within the participant trajectory and were not counted as independent qualitative cases. The second concerned separation between score direction and meaning, since coders examined the wording of rationales rather than assuming that an increased rating indicated metacognition. The third concerned disconfirming evidence, with deliberate attention to cases of numerical gain accompanied by formulaic responses and to cases of stable or uncertain ratings accompanied by sophisticated self-diagnosis. The fourth concerned the status of lexical counts, which were treated as navigational indicators rather than as equivalent to thematic prevalence.

Reporting was aligned with principles of transparent qualitative description concerning sampling, researcher procedure, data processing, analytic logic, and the relation between evidence and interpretation (O'Brien et al., 2014). Short participant quotations were selected only after trajectory-level analysis and are used to illustrate an interpretive claim rather than to stand in for it. All translations were prepared after the original Italian text had been coded. No names, schools, or other directly identifying details are reported.

The archived materials establish voluntary participation and provide the data necessary for the present secondary analysis, but they do not contain an institutional ethics approval or exemption identifier. The final submitted manuscript should therefore include the applicable institutional ethics determination, consent statement, and data-access conditions required by the journal and the authors' institution. No approval code is invented or inferred in the present draft.

3. Results

3.1. Quantitative context of perceived change

The paired reanalysis confirms a broad increase in self-assessed digital competence across the programme while also clarifying that the scale of change was not uniform across domains. The overall mean rose by 0.76 points on the five-point response scale, and every item retained statistical significance after Holm adjustment. The largest mean increases concerned identifying online platforms, assessing student competences, providing timely feedback, and facilitating collaboration. The smallest increase concerned assessment, personal feedback, and support as framed in A3, although even this item showed a positive paired difference.

The item most directly connected with the study's metacognitive interest was A15, which asked whether teachers could use digital technologies to help students reflect on their own learning. Its mean increased from 2.64 to 3.58. Seventy-eight teachers increased their rating, thirty-five retained the same score, and five reported a decrease. The standardised paired effect was substantial, $dz = 0.91$. This pattern establishes a context of perceived development but does not determine the reflective depth of that development. A higher rating may indicate greater familiarity, confidence, positive evaluation of the programme, or more sophisticated

regulation, and these possibilities can only be distinguished through the associated rationales.

3.2. Uneven reflective density and the limits of numerical improvement

The first qualitative result concerned the uneven density of the written corpus. The median response length decreased from five words before the programme to four words after it, and participant-level total word counts were also lower after the programme, with a Wilcoxon signed-rank sensitivity test indicating a systematic difference, $p < .001$. Among the 1,770 post-programme rationales, 635 contained only one or two tokenised words, corresponding to 35.9 percent of the final corpus. Twelve participants used the same formulation across all fifteen post-programme rationales. Expressions translated as "yes", "now I do", "quite", or "thanks to the course" communicated a positive orientation but did not expose criteria, monitoring processes, strategic choices, or evidence of transfer.

This divergence between numerical growth and argumentative elaboration is analytically important because it shows why reported engagement cannot be equated with metacognitive engagement. A teacher may assign a higher score and sincerely experience greater confidence while still providing no evidence of how competence is judged, which difficulties remain, or how a tool will be selected in relation to student learning. Conversely, a response containing uncertainty may carry greater metacognitive information when it distinguishes current understanding from what requires further practice. The qualitative analysis therefore treated the presence of limitation language as potentially productive rather than automatically negative.

3.3. Exploratory lexical mapping

The participant-level lexical mapping indicated that expanded digital repertoires were the most widespread marker, appearing in 83 of the 118 trajectories. References to regulation, practice, or the need for further development appeared in 63 trajectories, pedagogical orientation in 57, and confidence or self-efficacy in 51. Contextual constraints appeared in 30 trajectories, whereas explicit metacognitive language was identified in 17. These categories overlapped and were used only to orient close reading. The distribution suggests that the most common form of perceived change concerned what teachers could now access or do with technology, while explicit language concerning monitoring, awareness, and regulation remained comparatively less frequent.

Table 7. Participant-level markers identified during exploratory lexical mapping.

Marker	n	%	Operational indication
Expanded digital repertoire	83	70.3%	New applications, platforms, resources, or combinations of tools
Regulation or further practice	63	53.4%	Need for practice, further study, checking, or planned refinement
Pedagogical orientation	57	48.3%	Technology linked to a teaching or learning purpose
Confidence or self-efficacy	51	43.2%	Greater security, autonomy, or readiness to act

Marker	n	%	Operational indication
Contextual constraints	30	25.4%	Time, infrastructure, school conditions, uncertainty, or access barriers
Explicit metacognitive language	17	14.4%	Monitoring, reflection, awareness, error analysis, self-assessment, or regulation

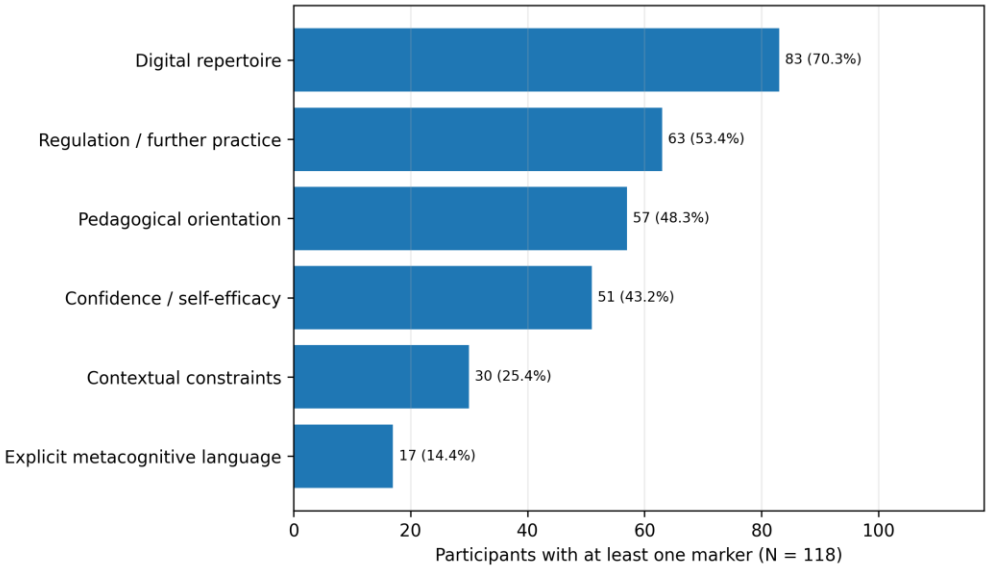


Figure 2. Participant-level presence of exploratory lexical markers. Categories are non-mutually exclusive and should not be read as definitive thematic prevalence.

3.4. From practical know-how to self-diagnosis

The dominant narrative of change concerned the acquisition of applications, platforms, and procedures. Many post-programme responses named resources that had not appeared in the corresponding pre-programme rationale and presented change through formulations equivalent to knowing more tools, being able to create resources, or feeling more prepared. This operational development is educationally relevant because a usable repertoire is a precondition for later strategic selection, yet it remains an initial form of engagement when the rationale does not indicate why one resource should be preferred to another or what evidence would justify adaptation.

A more reflective configuration appeared when participants differentiated what had been understood from what still required consolidation. P014 wrote that the procedure had been understood but that more practice was needed. P053 repeatedly indicated a need to deepen particular aspects rather than converting programme participation into an unqualified declaration of mastery. P065 stated that some uncertainty remained in the use of digital technologies. Such formulations were coded as emerging self-diagnosis because they make a boundary of competence visible and identify a future regulatory demand. Within the analytical framework, uncertainty therefore functions as evidence of metacognitive engagement when it is sufficiently specified to guide subsequent action.

This interpretation is consistent with a model of self-regulation in which awareness of current state and desired state supports strategy selection and future action, and with accounts of agency in which perceived control is monitored rather than assumed (Frazier et al., 2021; Metcalfe & Greene, 2007). The result also cautions against ranking participants by confidence alone. A highly confident but criterion-free statement may contain less regulatory information than a modest rating accompanied by an explicit account of what remains to be learned.

3.5. Pedagogical intentionality and the object of reflection

A second substantive theme concerned the movement from technical mastery towards pedagogical intentionality. Some teachers connected digital technology with motivation, participation, collaboration, personalisation, creativity, or inclusion and began to describe tools through the learning opportunities they might enable. P014, for example, linked future use with greater motivation and involvement and referred to support for students with special educational needs. Other participants connected digital activities with differentiated pathways, interdisciplinary work, or the creation of resources adapted to varied learning profiles.

These responses mark a change in the object of reflection. Technology is no longer considered only as something the teacher must learn to operate, but as a mediating element within a wider instructional relation involving students, purposes, tasks, and contextual constraints. The shift remains incomplete when engagement is equated with attractiveness or novelty, because motivational appeal does not by itself establish persistence, strategic depth, autonomy, or conceptual understanding. A more mature pedagogical criterion becomes visible when the teacher specifies what students should be enabled to monitor, create, compare, revise, or understand through the digital activity.

The distinction is particularly relevant for inclusive teacher education. Professional agency develops when perceived competence is translated into situated decisions that recognise learner variability and the consequences of pedagogical choices, rather than when teachers merely increase the number of available tools (Aiello et al., 2016; Aiello, 2019; Sibilio & Aiello, 2018). In this respect, the corpus suggests movement from technology as an object of professional acquisition towards technology as one component of an intentional learning design.

3.6. Advanced metacognitive and agentic configurations

The most elaborated evidence was concentrated in a smaller group of participant trajectories in which monitoring, feedback, error analysis, process evaluation, and transfer became explicit. P082 described the use of Google Forms, quizzes with immediate feedback, and self-assessment tools to help students monitor progress, identify recurring errors, and develop greater awareness of their learning trajectory. The significance of this response lies not in the named applications but in the regulatory sequence attributed to them, since technology is connected with feedback, diagnosis, monitoring, and the learner's capacity to act on information about performance.

P101 offered a different but related configuration. The programme was interpreted as an opportunity to compare prior professional practice with new possibilities offered by Artificial Intelligence, after which the reflection was extended to students' critical thinking and future construction of knowledge. Here the professional self, the technological environment, and student learning are brought

into the same explanatory frame. The response does not simply endorse Artificial Intelligence. It situates it within an evaluative process in which existing practice is reconsidered and future pedagogical action is justified through an educational criterion.

Across the corpus, these differences supported four provisional trajectories. The declarative-confirmatory configuration comprised positive but weakly elaborated rationales. The emergent-reflective configuration involved realistic recognition of uncertainties, limitations, or areas requiring further work. The strategic-regulatory configuration included intentions to experiment, criteria for selecting or adapting tools, and plans for monitoring their use. The transformative-agentic configuration connected professional reflection, instructional redesign, and the deliberate promotion of students' awareness, criticality, or self-regulation. These configurations are not proposed as linear stages or stable types. They function as interpretive patterns that distinguish qualitatively different relationships between engagement, metacognition, and professional agency.

Table 8. Provisional trajectory configurations derived from participant-level analysis.

Configuration	Defining feature	Illustrative evidence	Interpretive emphasis
Declarative-confirmatory	Positive endorsement with little or no articulated criterion	"Yes"; "Now I do"; "Thanks to the course"	Presence without visible regulation
Emergent-reflective	Recognition of uncertainty, limits, or need for consolidation	"I understood the procedure, but I need more practice" (P014)	Awareness and self-diagnosis
Strategic-regulatory	Planned experimentation, selection criteria, adaptation, or monitoring	Intentions to test tools in relation to learner needs and review their effects	Regulation and conditional decision making
Transformative-agentic	Professional reflection linked with instructional redesign and student self-regulation	Monitoring progress and recurring errors (P082); comparing prior practice with AI-supported possibilities (P101)	Agency oriented towards learner participation and epistemic responsibility

4. Discussion

The findings support a more demanding interpretation of engagement in teacher digital education. Participation in a programme, increased self-assessment, satisfaction with new tools, and the expansion of a digital repertoire can all be meaningful outcomes, yet none is sufficient to establish that engagement has become metacognitive. Metacognitive engagement becomes analytically visible when teachers can make the grounds of their judgement available for examination, identify what remains uncertain, monitor the adequacy of a strategy, formulate a subsequent action, and transfer reflection from their own learning to the learning conditions they design for students. This formulation is consistent with work that treats metacognitive engagement as a mode capable of explaining durable learning beyond passive or merely active interaction with digital media (Omarchevska et al., 2026).

The discrepancy between score growth and short post-programme rationales is therefore not a peripheral linguistic feature. It exposes the difference between endorsement and regulation. An increase from a lower to a higher self-assessment may truthfully represent improved confidence, but without an articulated criterion it remains impossible to determine whether the teacher can discriminate among tasks, tools, and educational purposes. This finding also qualifies the interpretation of self-efficacy. Perceived capability may support action, experimentation, and persistence, but professional agency requires that the sense of capability be joined to situated judgement and responsibility. The integration of agency and metacognition is therefore more informative than either construct considered in isolation (Frazier et al., 2021; Gao & Zhang, 2011).

The responses that acknowledged uncertainty deserve particular attention. In professional-development research, post-intervention uncertainty can easily be read as evidence of limited success. The trajectory analysis suggests a different possibility. When uncertainty is specific enough to identify a boundary of knowledge and a next step, it may indicate that participants have moved from undifferentiated familiarity towards a more accurate representation of their competence. Metacognitive judgements of agency are not valuable because they maximise confidence but because they help individuals distinguish what they control from what they do not yet control and adapt action accordingly (Metcalfe & Greene, 2007; Metcalfe et al., 2010).

The transition from tool-centred discourse to pedagogical intentionality likewise has implications for digital education. The educational problem is not whether teachers can name or operate a growing number of applications, but whether they can establish a relationship between technological affordances, learning goals, learner variability, evidence of progress, and the ethical consequences of design decisions. This is especially significant in the presence of generative Artificial Intelligence, where gains in speed and production can coexist with changes in cognitive effort, epistemic dependence, and the distribution of agency between human and computational actors (Yan et al., 2025). A teacher who can critically direct technology according to an educational purpose occupies a different position from one who merely adopts it because it is available or engaging.

The results also strengthen an agentic interpretation of teacher professional development. Teacher agency is not manifested by unbounded autonomy, but by the capacity to read contextual possibilities and constraints, make intentional choices, and act in ways that preserve educational responsibility and learner participation. The corpus contains evidence of both sides of this relation. Some teachers described new confidence and plans for experimentation, while others referred to infrastructure, practice needs, or uncertainty. Such constraints are not external noise to be removed from the analysis. They are part of the situation within which agency is exercised and therefore part of what professional learning must enable teachers to interpret (Aiello et al., 2016; Aiello, 2019, 2024; Sibilio & Aiello, 2018).

The use of self-assessment can consequently be reconsidered. A numerical self-rating becomes formative when it is accompanied by prompts that require the participant to justify a judgement, name evidence, identify a difficulty, plan an adjustment, or specify conditions for transfer. In professional-development programmes, this function could be strengthened through reflective prompts that

distinguish confidence from evidence, peer dialogue focused on competing design criteria, design journals that document changes of mind, portfolios containing successive versions of digital artefacts, and feedback oriented towards process rather than only product. Such devices would make metacognitive engagement more observable while simultaneously providing opportunities for it to develop.

The present analysis also contributes to the methodological discussion of qualitative data generated through questionnaires. Short rationales are often treated as isolated comments, yet their interpretive value changes when they are reconstructed as longitudinal participant trajectories. A phrase such as "I need more practice" may be trivial when detached from context, but it becomes analytically meaningful when it follows a pre-programme declaration of unfamiliarity and accompanies a specific change in rating or a newly articulated pedagogical purpose. The participant-by-item matrix therefore preserved temporal and thematic relations that would have been lost through simple frequency counting or by treating all 3,540 responses as independent units.

5. Methodological cautions and limitations

The study should be interpreted within several methodological boundaries. The sample was voluntary and restricted to one Italian region, which limits transferability to differently organised teacher populations. The intervention had no randomised comparison group, and the quantitative changes are self-reported. Accordingly, neither score increases nor qualitative shifts can be attributed causally to the programme alone. The numerical findings are used to contextualise participants' narratives and should not be read as direct measures of demonstrated classroom competence.

The SELFIE-inspired instrument was adapted for individual teacher self-assessment and was not used as a validated unidimensional scale. The high internal-consistency coefficients observed in this dataset may reflect substantial covariance among the fifteen areas, but they do not establish construct validity or justify treating the instrument as a diagnostic test. Reporting the paired composite and item-level estimates is therefore descriptive and contextual. The sensitivity analyses based on nonparametric procedures reduce dependence on distributional assumptions but cannot repair limitations originating in measurement design (Sullivan & Artino, 2013).

The written rationales varied considerably in length and many post-programme responses were extremely brief or repetitive. Absence of metacognitive language cannot be equated with absence of metacognition, since a participant may engage in sophisticated monitoring without expressing it in a short form field. The inverse is also true, because the presence of terms such as reflection or awareness does not by itself demonstrate effective regulation. For this reason, lexical indicators were triangulated with close reading and participant-level trajectories, and the provisional configurations are presented as interpretive patterns rather than psychometric classifications.

The secondary nature of the analysis also requires explicit publication transparency. The same training dataset underpinned the earlier study, but the present paper pursues a distinct question, uses the written rationales as its primary empirical object, reconstructs the unit of analysis at participant level, and re-estimates the paired quantitative context only to support interpretation (Di Paolo et

al., 2025). This distinction should be made explicit in any submission so that readers can identify what evidence is reused, what analysis is new, and how the present contribution extends rather than duplicates the previous report.

No classroom observations, learning analytics, or follow-up measures were available in the present dataset. The study can therefore identify narrated regulation but cannot establish whether that regulation was enacted or sustained in professional practice. A subsequent phase should combine stimulated-recall interviews with digital artefacts and, where feasible, classroom observation. A maximum-variation sample of approximately twenty teachers selected from contrasting pre-post trajectories would permit examination of how reported awareness is translated into planning, adaptation, feedback, and student-facing forms of metacognitive support.

6. Conclusions

This second-level analysis shows that the educational meaning of engagement cannot be inferred from presence, enthusiasm, or score improvement. Across the 118 teacher trajectories, perceived digital competence increased substantially, but the written rationales revealed a more differentiated picture in which practical acquisition, confidence, uncertainty, pedagogical intentionality, and explicit regulation appeared in different combinations. The analytical value of metacognitive engagement lies precisely in this differentiation, since it identifies the point at which participation becomes capable of observing and directing itself.

For teacher digital education, the distinction has both conceptual and practical consequences. Training becomes more educationally consequential when teachers are supported not only to acquire applications and procedures, but also to formulate criteria for their use, monitor the limitations of their own knowledge, test the consequences of design choices, and transfer reflection towards student learning. Professional agency emerges when competence is converted into reasoned action within real constraints and when technology remains subordinate to pedagogical judgement rather than defining the direction of practice.

The most advanced trajectories in the corpus suggest that this movement is possible. Digital tools became reflective surfaces through which some teachers reconsidered prior practice, designed feedback and self-assessment opportunities, and imagined conditions in which students could monitor errors, progress, and knowledge construction. The next research phase should test whether these narrated forms of awareness are sustained in enacted practice through stimulated-recall interviews, digital artefact analysis, and observational evidence. Such work would make it possible to examine not simply whether teachers use digital technologies more often, but whether they increasingly direct that use with metacognitive, pedagogical, and epistemic responsibility.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

References

Aiello, P. (2019). Teacher Education e Induction Period. Agentività del docente e sostenibilità di modelli formativi. *Nuova Secondaria, 36*(10), 58–61.

Aiello, P. (2024). Autodeterminazione, agentività e orientamento: verso una valorizzazione delle differenze di genere. In A. Di Stasi (Ed.), *Dalla non discriminazione alle pari opportunità: un itinerario di confronto, ricerca e sperimentazione di buone prassi a UNISA... e oltre* (pp. 47–56). Ledizioni.

Aiello, P., Sharma, U., & Sibilio, M. (2016). La centralità delle percezioni del docente nell'agire didattico inclusivo: perché una formazione docente in chiave semplessa? *Giornale Italiano della Ricerca Educativa, 9*(16), 17–28.

Broncano, F. (2014). Daring to believe: Metacognition, epistemic agency and reflective knowledge. In A. Fairweather (Ed.), *Virtue epistemology naturalized: Bridges between virtue epistemology and philosophy of science* (pp. 49–66). Springer. https://doi.org/10.1007/978-3-319-04672-3_4

de Winter, J. C. F., & Dodou, D. (2010). Five-point Likert items: t test versus Mann–Whitney–Wilcoxon. *Practical Assessment, Research & Evaluation, 15*(11), Article 11. https://doi.org/10.7275/bj1p-ts64

Di Paolo, A., Capodanno, F., La Manna, N., & Viola, I. (2025). La formazione degli insegnanti tra algoritmi e armonie in una prospettiva metadisciplinare inclusiva. *Journal of Inclusive Methodology and Technology in Learning and Teaching, 5*(3), 1–21.

European Commission. (2025). *SELFIE, Self-reflection on Effective Learning by Fostering the Use of Innovative Educational Technologies*. https://education.ec.europa.eu/selfie

Frazier, L. D., Schwartz, B. L., & Metcalfe, J. (2021). The MAPS model of self-regulation: Integrating metacognition, agency, and possible selves. *Metacognition and Learning, 16*(2), 297–318. https://doi.org/10.1007/s11409-020-09255-3

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology, 13*(1), 117. https://doi.org/10.1186/1471-2288-13-117

Gao, X., & Zhang, L. J. (2011). Joining forces for synergy: Agency and metacognition as interrelated theoretical perspectives on learner autonomy. In G. Murray, X. Gao, & T. Lamb (Eds.), *Identity, motivation and autonomy in language learning* (pp. 25–41). Multilingual Matters. https://doi.org/10.21832/9781847693747-004

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t tests and ANOVAs. *Frontiers in Psychology, 4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Metcalfe, J., Eich, T. S., & Castel, A. D. (2010). Metacognition of agency across the lifespan. *Cognition, 116*(2), 267–282. https://doi.org/10.1016/j.cognition.2010.05.009

Metcalfe, J., & Greene, M. J. (2007). Metacognition of agency. *Journal of Experimental Psychology: General, 136*(2), 184–199. https://doi.org/10.1037/0096-3445.136.2.184

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine, 89*(9), 1245–1251. https://doi.org/10.1097/ACM.0000000000000388

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods, 19*, 1–13. https://doi.org/10.1177/1609406919899220

Omarchevska, Y., Schuetze, B. A., Richter, J., Schanze, S., Ulrich, N., & Scheiter, K. (2026). From passive to metacognitive: The role of mode of engagement in enabling lasting learning using digital media. *Educational Psychology Review, 38*(1), Article 23. https://doi.org/10.1007/s10648-026-10120-z

Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). SAGE.

Sibilio, M., & Aiello, P. (Eds.). (2018). *Lo sviluppo professionale dei docenti. Ragionare di agentività per una scuola inclusiva.* EdiSES.

Sullivan, G. M., & Artino, A. R., Jr. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education, 5*(4), 541–542. https://doi.org/10.4300/JGME-5-4-18

Thomas, A. K., Wulff, A. N., Landinez, D., & Bulevich, J. B. (2022). Thinking about thinking about thinking … & feeling: A model for metacognitive and meta-affective processes in task engagement. *Wiley Interdisciplinary Reviews: Cognitive Science, 13*(6), e1618. https://doi.org/10.1002/wcs.1618

Yan, L., Pammer-Schindler, V., Mills, C., Nguyen, A., & Gašević, D. (2025). Beyond efficiency: Empirical insights on generative AI's impact on cognition, metacognition and epistemic agency in learning. *British Journal of Educational Technology, 56*(5), 1675–1685. https://doi.org/10.1111/bjet.70000