

D2.1 Immersive STEAM Competence Contextualisation Guidelines

WP2 STEAM Contextualisation



Project Number 101249392



Co-funded by the
European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Document Information

Project Title	Immersive STEAM Education Academy
Project Acronym	imSTEAM
Project Number	101249392
Start Date	1/1/2026
Ending date	31/12/2028
Work Package	WP2 STEAM Contextualisation
Lead Organization	UNIPA
Dissemination Level	PU
Document Name	Immersive STEAM Competence Contextualisation Guidelines
Type	R
No Deliverable	2.1



Funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or EACEA. Neither the European Union nor the granting authority can be held responsible for them.

Disclaimer

The information, documentation and figures in this deliverable are written by the imSTEAM project consortium under EC grant agreement Project — 101249392 and do not necessarily reflect the views of the European Commission. The European Commission is not liable for any use that may be made of the information contained herein.

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation, or both. Reproduction is authorised provided the source is acknowledged.

Any dissemination of results reflects only the author's view, and the European Commission is not responsible for any use that may be made of the information it contains. The sole responsibility for the content of this publication lies with the authors.

Copyright message

This report, is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0); a copy is available here:
<https://creativecommons.org/licenses/by/4.0/>.



Document History

Version	Date	Submitted by	Reviewed by	Notes
VER1	13/6/2026	UPatras		
VER2	18/6/2026	UNIPA	All Partners	
VER3	20/7/2026	All Partners		
VER4	29/7/2026	UPatras	K. N. GOURGOULIATOS	



Table of Contents

Table of tables	6
Table of figures	9
Executive Summary	10
1. Introduction	12
1.1 Scope	12
1.2 Methodology	13
1.3 Audience	14
1.4 Structure	15
2. Desktop Survey	17
2.1 Method of Enquiry	17
2.2 Slovenia	20
2.3 Croatia	21
2.4 Italy	21
2.5 Belgium — Flemish Community	22
2.6 Greece	23
2.7 Cross-Validation across Systems	24
2.8 Conclusions	26
3. The PISA skills pool	28
3.1 Overall approach	28
3.2 Overall architecture	29
3.3 The Backward methodology	32
3.4 Pool metrics	33
3.5 Limitations	35
4. Workshop	37
4.1 Scope and goals	37
4.2 Methodology	37
4.3 Session 1	40
4.4 Session 2	43
4.5 Conclusions	45
5. Online Survey	46
5.1 Purpose, Instrument and Method	46
5.1.1 Identity and purpose of the study	46
5.1.2 The instrument	46
5.1.3 Deployment and fieldwork	48
5.1.4 Methodological decisions	48
5.1.5 Limitations	49
5.2 Respondent Profile	50



5.3	National Findings.....	51
	Italy.....	52
	Greece.....	54
	Belgium — Flemish Community	56
	Slovenia.....	57
	Croatia.....	58
5.4	Cross-National Analysis	59
	5.4.1 The training premise: emphasis, depth and difficulty.....	59
	5.4.2 The medium: difficulty and affordance.....	60
	5.4.3 Implementation: conditions and readiness.....	62
5.5	Conclusions	64
	National findings.....	64
	Cross-national findings	67
	What the survey establishes for the project	69
6.	Implementation guidelines.....	71
6.1	Scope.....	71
6.2	Guidelines.....	71
	6.2.1 Designing courses and microcourses.....	72
	6.2.2 Designing Train the trainer courses.....	74
	6.2.3 Teacher self-assessment.....	75
	6.2.4 National adaptation	77
	6.2.5 Scenario-based learning.....	78
6.3	Worked Examples	80
	6.3.1 STEAM Immersive Investigation	80
	6.3.2 Green Capability	82
7.	Conclusions.....	85
	7.1 What the deliverable establishes.....	85
	7.2 Boundaries of the deliverable	86
	7.3 Relationship to the work that follows	87
	Annex I The PISA skills Pool.....	88
	Annex II The ImSTEAM Teachers Capability framework	135
	Annex III AI Skills Supplement.....	151
	Annex IV Analysis of National Policies	155
	Annex V Online Survey Questionnaire	173
	Annex VI Online Survey Analysis	178
	References.....	184

Table of tables

Table 1 The various phases of the methodology from baseline to validation.....	14
Table 2 Dimensions of the guided questionnaire administered to partners.....	17
Table 3 Contributing partners and principal curriculum evidence, by national system.....	18
Table 4 Principal policy and curriculum documents reviewed, by national system.....	19
Table 5 Cross-system synthesis of STEAM provision against the PISA 2025 architecture.....	24
Table 6 Strands, layer-1 sources, top-level and skills pool.	34
Table 7 Summary of Strands covered by the Skills Pool and sub-abilities.	34
Table 8 Strands, items at the station and selection rule followed in the workshop.	39
Table 9 Prioritised subset selected by the workshop participants.....	40
Table 10 Kendall coefficient of concordance for the various choices of the participants.....	40
Table 11 Comparison of the selection of each set using the Jaccard similarity coefficient.....	41
Table 12 Reasoning theme and mentions by workshop participants.....	42
Table 13 Session 2 outcome by strand.	44
Table 14 Structure of the online survey instrument	47
Table 15 National versions, languages and valid responses (fieldwork 22 June – 20 July 2026).	48
Table 16 Composition of each national sample along the principal profile dimensions.	51
Table 17 Learning architecture.	72
Table 18 Credit and learning time.	72
Table 19 Worked example of teaching focus and learning objective.....	73
Table 20 Learning pathways.....	73
Table 21 Worked example of capability 2.1.....	75
Table 22 Worked example connecting skill, teaching focus and self-assessment prompt.	76
Table 23 Comparison between countries - status in the curriculum and adaptation move.....	78
Table 24 Capability 3.3 connecting the scenario element to its contribution.....	79
Table 25 Combination between capabilities and their contribution to the investigation.	80
Table 26 Student goal, STEAM and framework linking.	80
Table 27 Teacher's scaffolding move.	81
Table 28 Connection between capabilities and their contribution to investigation.....	82
Table 29 Linking between student goal, STEAM and framework.	83
Table 30 Skill 4.3.2 and teacher's scaffolding move.	83
Table 31 Sub-ability A1.1 - Recalling and applying scientific knowledge.	89
Table 32 Using and translating between forms of representation.	91
Table 33 Making and justifying scientific predictions.....	93
Table 34 Identifying, constructing, and evaluating models.....	95

Table 35 Recognising and developing explanatory hypotheses.....	97
Table 36 Explaining the societal implications of scientific knowledge.....	99
Table 37 Identifying the research question of a scientific study.	101
Table 38 Proposing an appropriate experimental design.	103
Table 39 Evaluating the fit of an experimental design to its question.....	105
Table 40 Interpreting data, drawing conclusions, and evaluating their relative merits.	107
Table 41 Searching, evaluating and communicating the merits of information sources.	109
Table 42 Distinguishing claims by type and warrant.....	111
Table 43 Constructing scientific arguments from data.....	113
Table 44 Critique standard flaws in sciences-related arguments.....	115
Table 45 Justifying decisions using scientific arguments on contemporary issues.....	117
Table 46 Environmental awareness.	119
Table 47 Environmental concern.	121
Table 48 Environmental agency.....	123
Table 49 Mathematical reasoning.	125
Table 50 Formulating situations mathematically.....	126
Table 51 Employing mathematical concepts, facts, procedures and reasoning.....	127
Table 52 Interpreting, applying and evaluating mathematical outcomes.	128
Table 53 Engaging with AI.....	129
Table 54 Managing AI.	131
Table 55 New A-strand capabilities.	132
Table 56 Integration map.	133
Table 57 Categories, anchoring reference, capabilities and skills.	136
Table 58 Capabilities, anchoring reference and skills.	136
Table 59 PISA 2025 Competency 1: using and translating between forms of representation.	140
Table 60 PISA 2025: Competency 1: making and justifying scientific predictions.	141
Table 61 PISA 2025, Competency 1: identifying, constructing and evaluating models.....	142
Table 62 PISA 2020, Competency 2: proposing an appropriate experimental design.....	143
Table 63 PISA 2025, Competency 2: interpreting data, drawing and evaluating conclusions.	144
Table 64 PISA 2025, Competency 3: searching, evaluating and communicating the merits of sources.....	145
Table 65 PISA 2025, Competency 3: constructing scientific arguments from data.	146
Table 66 PISA 2025, Competency 3: justifying decisions on contemporary issues.	147
Table 67 PISA 2025, environmental awareness.	148
Table 68 PISA 2025, environmental concern.....	149
Table 69 PISA 205, environmental agency.....	150
Table 70 Status within ImSTEAM.	152

Table 71 AI capability and anchoring reference.....	152
Table 72 Allit Framework, engaging with AI (domain level).	153
Table 73 Allit framework, managing AI (domain level).....	154
Table 74 Country findings for STEAM integration, assessment, PISA alignment and XR readiness.....	169
Table 75 Overall findings across the dimensions of STEAM integration, assessment, PISA Alignment and XR readiness.	171
Table 76 Gap analysis by country.	171
Table 77 Respondents profile collected so that needs can be understood across countries, subjects and settings.....	173
Table 78 Survey Q1.....	174
Table 79 Survey Q3.....	174
Table 80 Survey Q4.....	175
Table 81 Survey Q5.....	175
Table 82 Survey Q6.....	175
Table 83 Survey Q7.....	176
Table 84 Survey Q8.....	176
Table 85 Survey Q9.....	176
Table 86 Scale-mapping.	177
Table 87 Constraints on science teaching (Q3): mean severity and, in parentheses, top-two-box percentage, by country.....	179
Table 88 Self-assessed readiness for immersive teaching (Q8): mean and, in parentheses, top-two-box percentage, by country.	181
Table 89 Prior experience with immersive tools (Q9): percentage of respondents in each category, by country, with valid n	181
Table 90 Master descriptive table: item mean (top-two-box percentage) for sections Q1, Q2, Q4, Q5, Q6 and Q7, by country.....	182

Table of figures

Figure 1 The development of ImSTEAM Teacher Capability Framework methodology demonstrating the steps taken to develop it.....	13
Figure 2 The three-level mapping chain.	29
Figure 3 Student-side sources (layer 1) and teacher capability strands (layer 4 anchor)	30
Figure 4 Prioritisation of connections between student competencies, PISA sub-abilities and teacher capabilities.....	33
Figure 5 Distribution of the 292 valid responses across the five national versions.....	50
Figure 6 Reported severity of constraints on science teaching, by national system.....	52
Figure 7 Reported difficulty in the classroom (Q5) against anticipated benefit of the corresponding immersive affordance (Q6).....	60
Figure 8 Perceived benefit of immersive activities against the five affordances, by national system.	61
Figure 9 Self-assessed readiness for immersive teaching across five dimensions, by national system.	63
Figure 10 Prior experience with immersive tools, by national system.	63
Figure 11 Difference between depth of teaching (Q2) and curricular emphasis (Q1) for each competency, by country.	178
Figure 12 Difficulty in engaging students as scientists (Q4), by practice and country.	180
Figure 13 Reported difficulty (Q5) and anticipated immersive benefit (Q6) for each paired affordance, by country.....	180
Figure 14 Conditions required for immersive activities to be usable (Q7), by condition and country.	181

Executive Summary

This deliverable establishes the conceptual foundation on which the ImSTEAM intervention is built. Its purpose is to translate the competency architecture of the PISA 2025 Science Framework into a structured and usable account of the capabilities teachers require in order to develop scientific competence in their students, and to set out how that account is applied in the design of immersive STEAM education. The deliverable is addressed to two audiences: those who design teacher education and curricula, and the educators who deliver them.

The work proceeds in four stages, each reported in a chapter of this document. First, a desk-based review examined national curriculum provision in five partner countries — Slovenia, Croatia, Italy, the Flemish Community of Belgium, and Greece — examining each system in relation to the PISA 2025 architecture in order to establish an empirical baseline. Second, a systematic derivation translated the student-side competencies of PISA 2025 into a pool of teacher capabilities, reasoning from each student competency and sub-ability to the pedagogical content knowledge a teacher must hold for it to be attainable. Third, a consortium workshop prioritised the items in this pool, selected the subset to be taken forward into the project's pilot phase and elaborated each selected capability into the concrete teaching actions that enact it. Fourth, an online survey carried the work to a wider body of teachers across the five countries, validating the framework and gathering evidence on its localisation to national contexts and on the use of immersive technology to develop the capabilities.

The desktop review produced a consistent baseline across the five systems. The competences that immersive STEAM education seeks to develop are broadly present, but in latent form: they are seldom explicitly labelled and are rarely the subject of deliberate curriculum or instructional design, and not mapped to the PISA descriptors, so that existing good practice is neither transferable nor assessable. The marginalisation of the arts as a mode of scientific inquiry, the weakness of STEAM-specific assessment, and the implicit and therefore unusable character of PISA alignment recur across systems that otherwise differ markedly. This convergence indicates that the deficits are structural features of contemporary curriculum provision rather than national idiosyncrasies, and that they are therefore amenable to a common contextualisation response.

The principal output of the deliverable is the ImSTEAM Teacher Capability Framework, presented in Annex II. Derived entirely from the PISA 2025 Science Framework, it comprises four categories, eleven capabilities and forty-nine teacher skills, covering the three science competencies — explaining phenomena, scientific enquiry and data, and evidence, argument



and decision — together with the environmental science-identity dimension. Each capability is resolved into teacher skills that span the full teaching arc — what to teach, where students struggle, how to teach it, how to frame it, and how to assess it — and each skill carries a draft objective at a defined cognitive level, so that a capability functions as a specified unit of training rather than a general statement of desirable competence.

The framework was validated through an online survey of 292 teachers across the five national systems. The survey confirmed the empirical premise of the intervention, though not in the form of an unmet curricular ambition: teachers do not report a large gap between what their curricula emphasise and what they teach, and the case for the framework rests instead on difficulty, which the survey confirms — the scientific practices the framework addresses are reported as genuinely difficult to develop in students, consistently across every system and concentrated in the practices, such as forming and testing hypotheses and building and using models, that the framework is built around. The survey further established that the project's affordance-led treatment of immersive technology is sound: the classroom difficulties teachers report correspond, affordance by affordance, to the benefits they anticipate from immersive activities, and for every paired affordance the anticipated benefit exceeds the reported difficulty, in each of the five national contexts as well as in aggregate. Finally, it specified the conditions of use and the readiness against which the training must be pitched: teachers attach consistent and practical conditions to immersive activities — compatibility with class size, devices and lesson time — and judge themselves pedagogically prepared but technically unready, the great majority having little or no prior experience of the medium. National differences in constraint and requirement, most sharply the acute examination pressure and thin base of localised digital resource reported in Greece, inform the national-adaptation guidance, while unprompted responses pointing to mathematics and to the interpretation of artificial-intelligence models offer independent support for the strands the framework treats provisionally. The survey thus extends the empirical basis of the framework from the consortium to a wider body of teachers without altering its content.

The implementation guidelines set out how the framework is used. They identify its principal applications — the design of courses and microcourses, the preparation of trainers, teacher self-assessment, adaptation to a national curriculum, and scenario-based learning — and provide worked examples in which a single teacher skill is developed through an immersive resource. Throughout, the framework serves as the PISA-anchored foundation on which subsequent curriculum and scenario design depends, and immersive approaches are used where they develop a capability more effectively than conventional means rather than for their own sake. In this way the deliverable provides both the analytical foundation for the intervention that follows and the practical guidance required to put it to use.



1. Introduction

1.1 Scope

D2.1 defines the teacher capabilities required for immersive STEAM education and provides guidance on their use. Its scope follows from two commitments. The first is that the capabilities are derived from an established account of student competence rather than simply stated: the deliverable uses the PISA 2025 Science Framework as its sole reference framework and reasons from the student competencies that framework defines to the capabilities a teacher must hold in order to develop them. The second is that the resulting framework is designed to be practically usable: the deliverable does not stop at a statement of capabilities, but sets out how the resulting framework is applied in the design of teacher education and in classroom practice.

Within this remit, the deliverable covers the three scientific competencies of the PISA 2025 framework — explaining phenomena scientifically, constructing and evaluating scientific enquiry and interpreting data, and using evidence, argument and information to reach decisions — together with the environmental science-identity dimension that the framework defines alongside them. The capabilities are formulated throughout from the teacher's perspective: each capability specifies what a teacher must be able to do in order to develop a given ability in students. The guidance is addressed to two audiences whose needs differ — those who design teacher education and curricula, and the educators responsible for implementing them.

The scope is defined by three boundaries. First, the deliverable establishes the framework and the principles governing its use; it does not itself constitute the courses, curricula or learning resources built from it, which are the object of subsequent work. Second, immersive technology is treated as one medium through which these capabilities may be developed and is considered only where it can develop those capabilities more effectively than conventional approaches; the deliverable is not a technical specification for immersive scenarios. Third, the framework presented here covers the scientific and environmental capabilities in full, while the mathematics strand is completed through the online validation reported later in this document and the treatment of artificial intelligence is retained as a supplement rather than incorporated into the operative framework. Both choices reflect a deliberate decision to include a capability only where it is adequately supported by both the source framework and the relevant pedagogy.

The operative output of the deliverable is the teacher-capability framework presented in Annex II. The fuller derivation from which that framework is drawn is retained in the annexes



for traceability, but it is the framework itself that the guidance applies and on which subsequent work builds.

1.2 Methodology

The methodology is governed by the principle of backward design: the method begins from what students are expected to achieve, as defined by the PISA 2025 Science Framework, and derives, from each student ability, the corresponding teacher capability required to develop that ability

The derivation follows a layered structure. Each capability is anchored across four established frameworks, and each layer answers a distinct question. The first defines the student outcome to be developed; the second identifies the pedagogical knowledge required to teach it; the third determines how immersive technology can support the development of the capability; and the fourth situates it within the European policy context. The value of the layered structure is that it renders the derivation traceable: every capability can be followed back through the chain to the student ability and the source from which it arises, and the contribution of each framework is made explicit rather than assumed.

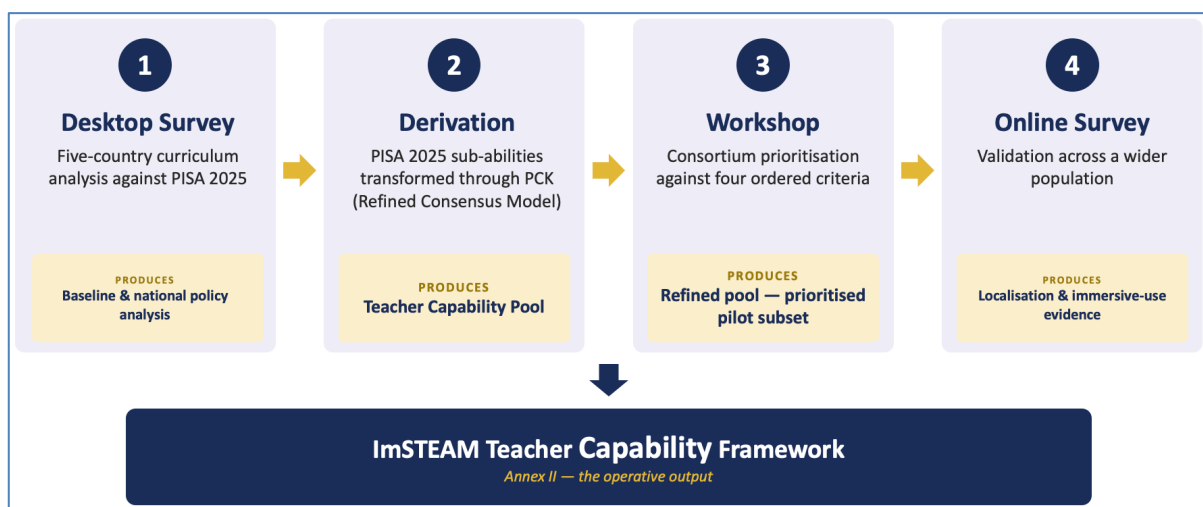


Figure 1 The development of ImSTEAM Teacher Capability Framework methodology demonstrating the steps taken to develop it.

The derivation is situated between two empirical stages and a separate selection process. It is preceded by a desktop survey, which establishes a baseline of existing curriculum provision against which the capabilities are to be contextualised and which supplies the evidence for later national adaptation. It is followed by a prioritisation process conducted through a consortium workshop. This reduces the exhaustive pool to the subset taken forward into the pilot and elaborates each selected capability into the teaching actions through which it is enacted. An online survey then extends the result to a wider population, gathers evidence on

localisation and on the use of immersive technology, and validates the framework without reopening it. The full sequence is summarised in the following table.

Table 1 The various phases of the methodology from baseline to validation.

Phase	Question addressed	Method or anchoring framework	Output
Baseline	What do national curricula already provide, and where are the gaps?	Common guided questionnaire; independent five-country desk analysis	Validated baseline (Ch. 2)
Derivation · Layer 1	What should students know and be able to do?	PISA 2025 Science Framework	Student competencies and sub-abilities
Derivation · Layer 2	What must a teacher be able to do to develop each ability?	Pedagogical Content Knowledge; Refined Consensus Model	Teacher-capability pool (Annex I)
Derivation · Layer 3	How can immersive technology support the development of the capability?	TPACK for immersive technology	Capabilities adapted for development through immersive approaches
Derivation · Layer 4	Where does the capability sit within EU policy?	DigCompEdu; GreenComp (environmental)	Policy-aligned capabilities
Selection	Which capabilities are taken forward into the pilot??	Consortium workshop; four ordered criteria; ranking and elaboration	Operative framework (Annex II, Ch. 4)
Validation	Is the framework validated, adaptable to national contexts, and suitable for immersive implementation?	Online survey of a wider population	Confirmed, localised framework (Ch. 5)

Read as a whole, the method converts an authoritative account of student competence into a teacher-capability framework that is pedagogically grounded, positioned for immersive delivery, aligned to European policy, prioritised for the pilot, and validated both within the consortium and among a wider group of respondents. The resulting framework is presented in Annex II, and the guidance governing its application follows in Chapter 6. The chapters between this introduction and that guidance report each phase of the method in turn, and should be read in relation to the structure outlined here.

1.3 Audience

This deliverable is intended for the groups who will use the teacher-capability framework. The formal description of the deliverable identifies curriculum developers and educators as the primary audience; within the wider intervention it additionally serves the principal target



groups — in-service teachers, pre-service teachers, and the trainers and content creators responsible for preparing these teachers — together with the educational designers and professional-development experts who support teachers and trainers, and the policy makers and educational leaders responsible for the systems within which they operate. Each of these audiences engages with the framework for a distinct purpose, and the deliverable is structured so that the guidance relevant to each audience may be readily identified.

The first audience comprises those responsible for the design of teacher education: curriculum and course developers, educational designers, professional-development experts, and the teacher educators who prepare other teachers. For this audience, the framework operates as a specification. Each capability corresponds to a discrete unit of training and is broken down into the teacher skills that constitute its structure, with the consequence that the framework may be translated directly into the design of courses and microcourses, the combination of multiple capabilities into longer professional-learning pathways, and the preparation of trainers. The implementation guidelines presented later in this document are organised principally around these applications.

The second audience comprises educators themselves — in-service and pre-service teachers, and the content creators who develop materials for teachers' use — who engage with the framework as practitioners rather than as designers. For a teacher, each capability constitutes a statement of the competence required to develop a given ability in students, and the teaching dimensions into which it is broken provide a structured means of determining where their practice is secure and where it requires further development. Employed in this manner, the framework serves as a diagnostic instrument in addition to a design tool.

The third audience comprises those responsible for the systems within which teaching is conducted: policy makers, educational leaders, and educational-policy designers at institutional and national levels. For this audience, the framework provides a common, PISA-anchored point of reference against which provision may be aligned, while the deliverable supplies a method for establishing where each capability is already present within a national curriculum, where it remains implicit, and where it is absent. The framework thus serves all of these audiences in common. The framework therefore serves all three audiences, but each uses it for a different purpose.

1.4 Structure

The document is organised to follow the sequence of the work it reports, advancing from the empirical baseline, through the derivation and prioritisation of the teacher capabilities, to the guidance governing their application. The chapters that follow this introduction are arranged accordingly.



- Chapter 2 reports the desk-based review, a five-country examination of national curriculum provision conducted against the architecture of the PISA 2025 Science Framework. Its purpose is to establish the empirical baseline for contextualising the capabilities within national settings, identifying where the relevant competences are already present, where they remain implicit, and where they are absent.
- Chapter 3 presents the derivation of teacher capabilities from the PISA 2025 framework. It sets out the method by which the student-side competencies defined by the framework are translated, competency by competency, into the capabilities teachers must possess to enable students to develop those competencies, and reports the resulting capability pool.
- Chapter 4 reports the consortium workshop at which the derived pool was prioritized, resulting in the subset taken forward into the pilot, and at which each selected capability was elaborated into the concrete teaching actions that enact it.
- Chapter 5 reports the online survey, which extends the work to a wider body of respondents and gathers further evidence on the localisation of the capabilities to national contexts and on the use of immersive technology to develop them.
- Chapter 6 sets out the implementation guidelines. It identifies the principal applications of the framework and provides worked examples in which individual teacher skills are developed through immersive resources.
- Chapter 7 brings together the main elements of the deliverable and states its relationship to the subsequent curriculum and scenario-design activities that build on it.

The annexes contain the supporting material. Annex I presents the full capability pool from which the framework is drawn; Annex II presents the operative teacher-capability framework, which constitutes the principal output of the deliverable; Annex III provides the supplementary treatment of artificial intelligence; and Annex IV presents the complete national-policy analysis underlying Chapter 2. Annex V reproduces, in English, the online survey questionnaire reported in Chapter 5, together with the mapping of its response scales onto the common metric used in the analysis; and Annex VI presents the full item-level statistical results of that survey. A list of references completes the document.

2. Desktop Survey

This chapter establishes the empirical baseline on which the contextualisation of immersive STEAM competences is grounded. It reports the findings of a five-country review of national curricula in science, mathematics, technology and the arts, examined in relation to the competency architecture of the PISA2025 Science Framework. The purpose of the exercise is not simply to describe or compare the five systems: it is to determine, system by system, the extent to which the competences that immersive STEAM education seeks to cultivate are already present, how they are expressed, where they remain implicit, and where they are absent altogether. The chapter begins with a statement of method and then presents five national analyses, each examining existing curriculum arrangements and the gaps that remain; it then triangulates these accounts to distinguish findings that recur across systems from those that are nationally specific; and it closes by consolidating the validated baseline.

The full national analysis, including the complete evidence for each country and source documentation, is presented in Annex IV.

2.1 Method of Enquiry

Evidence was collected through a common guided questionnaire administered to all consortium partners. The instrument was constructed as a shared analytical grid, so that each national system would be examined through the same set of dimensions and the resulting accounts would be comparable and consistently structured. Partners responded by conducting a structured, desk-based examination of their national curriculum corpora — programme frameworks, subject syllabi, cross-curricular guidance and assessment provisions — and by under each dimension of the grid with supporting documentary evidence. The dimensions were defined so as to map directly onto the constructs examined in this deliverable, and in particular onto the competency descriptors of the PISA 2025 Science Framework.

The grid directed partners to characterise their curricula along seven dimensions, summarised in Table 2.1. A standing instruction accompanied the instrument: partners were asked to interpret the arts dimension not as a self-contained aesthetic domain but in its relation to innovation in scientific reasoning, in keeping with the integrative rationale of the project. This instruction is consequential for the findings that follow, because it directs attention to whether the arts function as a mode of inquiry within science and mathematics rather than merely as a parallel subject.

Table 2 Dimensions of the guided questionnaire administered to partners.



Dimension	Focus of enquiry
Educational levels	Whether provision addressed lower- or upper-secondary education.
STEAM disciplines and integration	Whether disciplines were taught discretely or through interdisciplinary structures, with documentary evidence.
Learning objectives and competences	How objectives were formulated in relation to STEAM disciplines and to transversal competences, and whether stated explicitly or only integrated across subjects.
Assessment approaches	Whether STEAM-specific assessment existed, and the role of formative practice such as portfolios, process observation and feedback.
Alignment with PISA 2025	Points of convergence and divergence with the PISA assessment framework, with reference to scientific reasoning, mathematical literacy, problem solving and creative thinking.
National context	Reforms, policy priorities, governance arrangements and curriculum autonomy bearing on provision.
Implications for practice	Consequences for classroom practice and for teacher professional development.

Five national systems were examined — Slovenia, Croatia, Italy, the Flemish Community of Belgium, and Greece — each system was examined through the combined work of higher-education, school and intermediary partners. The contributing partners and the principal documents they analysed are set out in Table 2.2. The breadth of the corpus, and the fact that each system was analysed independently by partners working within it, are what allow the synthesis in Section 2.7 to treat recurrent findings as cross-validated rather than as artefacts of a single vantage point.

Table 3 Contributing partners and principal curriculum evidence, by national system.

System	Contributing partners	Principal evidence analysed
Slovenia	GoINNO Inštitut; Osnovna šola Solkan; University of Primorska (Faculty of Education)	Gymnasium and lower-secondary syllabi; Chemistry electives; Physics and Ecology projects; national sustainability and recovery-plan frameworks
Croatia	University of Rijeka (Faculty of Mathematics); DANTE; ADRIA	National Curriculum Framework; subject curricula for mathematics and ICT cross-curricular provision
Italy	I.C. 'Cassarà-Guida', Palermo; University of Palermo (UNIPA); CESIE	Nuove Indicazioni Nazionali (2025); national STEM guidelines (2023); upper-secondary guidelines for licei, technical and vocational tracks
Belgium (Flemish Community)	UC Limburg (UCLL); KSO Maaseik-Kinrooi; Mosa-RT; Schoolmakers	Secondary attainment targets and the sixteen key competences; documented practice of ICT, AI and XR implementation in schools
Greece	Academy of Athens; University of Patras (UPatras)	Skill Labs framework; Technology course curriculum; 2023 Mathematics (gymnasium) and Physics (lyceum) curricula

For completeness, and to document the evidential basis of the national analyses, the following table records the principal policy frameworks and curriculum documents examined in each system. The list is confined to the named instruments on which the findings of Sections 2.2 to 2.6 substantively rely; it is indicative of the corpus rather than an exhaustive list of every subject syllabus consulted.

Table 4 Principal policy and curriculum documents reviewed, by national system.

System	Policy and curriculum documents referenced
Slovenia	Gymnasium subject curricula — Mathematics (2008), Biology, Chemistry, Physics, Computer Science (upper secondary) Chemistry elective curricula — Chemistry Experiments; Chemistry in the Environment; Chemistry in Life Drawing in Geometry and Technology curriculum (2012) Physics and Ecology Projects curriculum Common Objectives framework (2023) Recovery and Resilience Plan — RRP (2021) Digital Education Action Plan — ANDI (2021–2027)
Croatia	National Curriculum Framework (2017) National Curriculum for High School Education (2017) Curriculum for Mathematics, primary and secondary schools — NN 7/2019 Curriculum for the Cross-Curricular Theme: Use of ICT — NN 7/2019
Italy	Nuove Indicazioni Nazionali — DM 221/2025 Linee guida per le discipline STEM — DM 184/2023 Indicazioni nazionali per i Licei (2010) Linee guida Istituti Tecnici (2010) Linee guida per l'istruzione professionale National Recovery and Resilience Plan — PNRR, Mission 4 Scuola Futura platform
Belgium (Flemish Community)	Secondary attainment targets — minimum goals and key competences (Onderwijsdoelen, 2025) Common Foundational Curriculum — GFL (2023) Educational Goals in Secondary Education (Onderwijsdoelen secundair onderwijs) Schoolmakers — Implementation of ICT, AI and XR in Education (2026)
Greece	New Curricula — Nea Programmata Spoudon (IEP, 2021–2023) Skill Labs Programme framework — 21st Century Skills Labs Technology Course curriculum (grades 7–9) Mathematics Curriculum 2023 — Gymnasium (National Gazette B, 235/2023) Physics Curriculum 2023 — Lyceum (National Gazette B, 2913/2023)

2.2 Slovenia

The Slovenian curriculum provides a strong disciplinary foundation for STEAM education, although interdisciplinary integration remains implicit rather than being organised under an explicit STEAM label.. Relevant content and guidance are distributed across programme frameworks, subject syllabi, cross-curricular documents and assessment guidance, and nowhere is STEAM presented as a standalone curriculum category. Despite this dispersion, the curriculum nevertheless contains substantial interdisciplinary provision. Subjects such as Spatial Modelling connect descriptive geometry, mathematics, informatics and information technology; Microbiology links biology, chemistry and laboratory work with sustainable development; and Agriculture connects the natural sciences with applied technologies including positioning systems and automated machinery. Environmental Education, framed as education for sustainable development, is the only thematic area explicitly designated as cross-curricular.

Two features distinguish the Slovenian provision as the most operationally mature in the corpus. In the Chemistry elective curricula, every operational objective is accompanied by a dedicated column identifying interdisciplinary links, making explicit the connections between chemistry and physics, biology, mathematics, the fine arts and computer science; transversal competences are articulated as ‘knowledge for lifelong learning’ and assessed through portfolios, peer and self-evaluation and continuous observation. The Drawing in Geometry and Technology curriculum advances the most complete integration model in the dataset, treating virtual three-dimensional space as a shared platform on which mathematical precision, technological tools and artistic creativity converge in the solution of engineering problems, with assessment that combines evaluation of both the final product and the process, peer assessment and standards-based criteria. At policy level, a sustainability-competence framework defined in operational terms — encompassing systems thinking, problem formulation, futures literacy and collective action — provides the most detailed vocabulary of environmental competence in the national corpus, while formative assessment is described as widely established in practice.

The principal limitation is structural rather than substantive. Because none of this good practice is labelled or consolidated at system level, it remains invisible and difficult to identify and transfer systematically: strong instances such as the interdisciplinary-links column or the peer-assessment model used in the Drawing in Geometry and Technology curriculum cannot be recognised, disseminated or assessed as STEAM provision. The environmental objectives embedded in the Physics and Ecology projects are not connected to the environmental competency of PISA 2025 or to a sustainability-competence framework; systems thinking is not addressed explicitly in that provision; and the arts, present in isolated cases such as Dance Composition and the chemistry–fine-arts connection are not systematically incorporated into



science or mathematics teaching at system level. Immersive technology is structurally absent despite its clear potential for molecular visualisation, virtual laboratories and ecological simulation, and formative assessment, although widely practised, is not codified, so its quality cannot be documented or evaluated systematically across the education system.

2.3 Croatia

The reviewed Croatian curriculum documents show a strong focus on interdisciplinary learning and the development of analytical reasoning, scientific thinking, and digital competence, primarily articulated in the National Curriculum Framework. This framework organises generic competences into three broad domains: ways of thinking; ways of working and using tools; and personal and social development. These domains encompass problem solving, decision making, critical thinking, metacognition, creativity, and innovation, and are intended to be developed across levels, areas, subjects, and cross-curricular themes. Assessment is conceptualised through the distinction between assessment for learning, assessment as learning, and assessment of learning; in mathematics, the emphasis is on conceptual understanding, mathematical communication, and problem solving rather than procedural recall. Arts-related assessment may include process-oriented and reflective dimensions within separate arts curricula, but the reviewed sources do not provide sufficient evidence to claim that arts-based assessment is systematically connected to STEM or STEAM learning.

However, the framework's ambition is not fully matched by explicit operationalisation at the STEAM level, and the main limitation is both terminological and structural. The reviewed documents do not use the concept of STEAM competences, nor do they define STEAM as a distinct competence framework. STEAM-relevant competences can be inferred from broader curricular aims, subject curricula, and cross-curricular themes, but their integration depends largely on interpretation by schools and teachers. The reviewed documents provide limited pedagogical guidance for interdisciplinary STEAM design and do not identify dedicated STEAM-specific assessment instruments. Similarly, the Arts dimension is not explicitly operationalised as an epistemological component of science, mathematics, technology, or interdisciplinary STEAM learning. No explicit instruments aligned with the PISA 2025 competency descriptors for STEAM competences are identified in the reviewed Croatian documents, making assessment one of the key areas for further development.

2.4 Italy

The Italian case is characterised by a pronounced tension between policy ambition and implementation. The most recent national framework articulates a pedagogical reorientation, described as a form of a renewed humanism, that treats STEAM not as a set of discrete technical subjects but as an integrated language in which scientific rigour meets creative



design and ethical reflection, with transversal skills presented as the means through which scientific content is conveyed. Its assessment provisions are among the most progressive in the corpus, privileging the application of knowledge in novel and complex contexts, evaluating process as well as product, and endorsing e-portfolios, self-assessment rubrics, systematic observation and the recognition of trial and error as a legitimate part of learning. The national STEM guidelines reinforce this orientation, positioning the interdisciplinary integration of theory and practice as the core of STEM education and invoking the framework of critical thinking, communication, collaboration and creativity, with authentic tasks and systematic observation proposed as the means of assessing competence acquisition. A national investment strand devoted to new skills and languages provides the policy framework within which such integration is encouraged across all school cycles.

In contrast to these progressive policy documents, the mainstream guidelines that continue to govern upper-secondary curricula and teaching present a markedly different picture. Objectives and competences are formulated discipline by discipline and are frequently fragmented; interdisciplinary approaches are not operationalised into concrete objectives; transversal competences are tied to individual subjects rather than pursued as a cross-disciplinary agenda; and no specific approach to the assessment of STEAM competences is articulated anywhere, the only STEAM-related skills subject to nationally standardised testing being those of mathematics. The arts are absent from both the STEM guidelines and the investment framework, even as the newest curriculum names them explicitly. Italy exhibits one of the most substantial gaps between progressive policy intent and operational assessment provision in the dataset: no dedicated STEAM assessment instruments were identified within mainstream provision, and the progressive intent of the newer documents is not yet matched by operational tools, with implementation dependent on school leadership and on structures for cross-departmental collaboration that are not guaranteed.

2.5 Belgium — Flemish Community

Of the systems examined, the Flemish system provides the most explicit structural treatment of transversal competences. Its attainment targets are organised not by subject but by sixteen key competences, two of which — competence in mathematics, science and technology, and digital competence and media literacy — contain the main substantive content relevant to STEAM., alongside cross-cutting competences in research, creativity, problem solving, critical thinking and collaboration. Culture and expression is itself one of the sixteen, thereby giving the arts an explicit structural presence, and the framework explicitly names systems thinking as a competence area, which is directly relevant to the environmental competency of PISA 2025. Within this competence-based architecture, school boards determine the subjects or interdisciplinary pathways through which the competences are realised.



A distinctive and methodologically valuable finding concerns the implementation of immersive technology rather than the content of any curriculum. The documented experience of integrating information technology, artificial intelligence and extended reality in Flemish schools shows that effective adoption typically unfolds over several school years and depends on the sustained development of a core team of early adopters before wider implementation. Success is conditioned by a supported vision and clear leadership, incremental implementation with defined milestones, pedagogical anchoring in genuine learning goals, professional development that creates space for critical dialogue, and a culture of experimentation supported by psychological safety; the corresponding threats are resource uncertainty, top-down imposition without buy-in, an excessive focus on the novelty of the technology or devices, and the underestimation of workload.

The gaps are accordingly structural rather than conceptual. Implementation is fully delegated to schools. This arrangement is grounded in the constitutional freedom of education that protects pedagogical autonomy and prevents the national prescription of specific teaching methods, produces considerable variability in the coherence and quality of STEAM competence development. The arts, although structurally present as culture and expression, are not connected to learning in science and technology; Assessment is not specified at national level, meaning that no system-wide evidence on the assessment of STEAM competences is available. Moreover, the time required for immersive technology to become sustainably embedded exceeds the duration of short-cycle interventions..

2.6 Greece

Greece comes closest among the systems examined to having a structured STEAM approach, through two institutional arrangements. The Skill Labs, introduced nationally and made compulsory in kindergarten, primary education and lower-secondary education, constitute the most structurally innovative provision in the dataset: organised as sequences of thematic workshops, they target critical thinking, communication, collaboration and creativity as explicit pedagogical objectives rather than as by-products of subject teaching, and they are assessed through descriptive and authentic evaluation, rubrics provided by the Ministry, the student portfolio and reflective journals for self- and peer-assessment. The Technology course introduces a STEAM content approach organised around inquiry into real-world problems through an engineering design process and articulates a nationally specific adaptation of the STEM acronym in which physics is specified in place of the broader category of science and language is added as an explicit element — a formulation that connects directly to the project's rationale for integrating the arts and language with scientific reasoning.

The 2023 curricula for mathematics and physics are among the most recently issued in the corpus and display close, if implicit, affinity with the PISA competency clusters. Mathematics formulates objectives as expected learning outcomes and integrates transversal



competences — reasoning, modelling, communication, reflection, critical understanding and data analysis — recurrently through its thematic fields, and adopts a formative orientation centred on assessment for learning. Physics aligns closely with the clusters concerned with explaining phenomena scientifically and with constructing and evaluating scientific enquiry and promotes two comparatively developed forms of assessment: laboratory reports modelled on scientific writing, and argumentative writing in which students develop and defend scientific claims.

Significant limitations nevertheless remain. The Skill Labs remain a time- limited addition to mainstream curriculum rather than a systemic redesign, their competence framework is not aligned with the PISA descriptors, and their implementation is concentrated in lower-secondary education, creating a discontinuity in students' progression into upper-secondary education. The implicit alignment evident in physics and mathematics is genuine, but it cannot yet be used systematically: teachers may be well aligned with the PISA constructs by virtue of their disciplinary formation yet may be unable to identify and articulate that alignment explicitly or design deliberately for the corresponding competences in their own students. Immersive technology, finally, is absent from the objectives of physics and mathematics despite the clear potential for virtual experimentation and three-dimensional modelling.

2.7 Cross-Validation across Systems

Considered together, the five national accounts converge on a small set of findings with sufficient consistency to give them evidential weight: because each system was analysed independently by partners working within it, the recurrence of a finding across systems provides cross-validation rather than mere repetition. Table 2.3 synthesises the position of each system on the dimensions that matter most for the contextualisation of immersive STEAM competences; the discussion that follows distinguishes what the systems share from the respects in which they differ.

Table 5 Cross-system synthesis of STEAM provision against the PISA 2025 architecture

Dimension	Slovenia	Croatia	Italy	Belgium (Fl.)	Greece
Explicit STEAM labelling	Absent	Absent (term unused)	STEM, not STEAM	Embedded, unnamed	Closest; still marginal
Locus of integration	Subject level	Policy level only	Policy vs. mainstream split	Competence architecture	Programme + subject level
Transversal competences	Operationally specific; unassessed	Three domains; no scaffolding	Referenced; not cross-disciplinary	Most explicit; devolved	Explicit; discontinuous
Arts as mode of inquiry	Isolated cases	Absent	Named, not enacted	Present, not linked	Via language extension

Dimension	Slovenia	Croatia	Italy	Belgium (Fl.)	Greece
STEAM-specific assessment	Advanced practice; uncodified	Defined; no instruments	Most severe gap	Unspecified nationally	Advanced; not PISA-aligned
Explicit PISA 2025 alignment	Implicit; unmapped	Implicit; no instruments	Implicit; unmapped	Implicit; systems-thinking hook	Implicit but close (C1/C2)
Immersive technology	Structurally absent	Not addressed	Not addressed	Multi-year maturation challenge	Absent from objectives

Five observations recur across every system.

1. First, STEAM is everywhere an emergent property of how individual subjects are taught in practice rather than a designed and labelled framework; in consequence it cannot be deliberately taught, assessed or certified, and its presence depends on the discretion of individual schools and educators.
2. Second, the arts are systematically marginalised as a means of generating and exploring knowledge: where they appear, they do so as aesthetic or cultural subjects, not as a mode of inquiry capable of enriching scientific and mathematical learning, so that is largely absent from the curricula analysed.
3. Third, STEAM-specific assessment is consistently the least developed dimension; formative-assessment principles are widely articulated, but no system describes instruments or criteria designed to assess integrative and interdisciplinary competences, and the most advanced practices that do exist remain subject-specific and disconnected from any shared framework.
4. Fourth, alignment with the PISA 2025 descriptors is implicit at best and, for that reason, unusable: teachers cannot deliberately design teaching to develop competences they cannot identify within their own curriculum, a deficit most acute for the epistemic-knowledge dimension and for the environmental competency.
5. Fifth, the integration of immersive technology is structurally nascent, and the evidence indicates that it cannot be treated as a short-term addition but requires a staged, multi-year trajectory.

The systems differ chiefly in where integration resides and how far it has been operationalised. In Slovenia and Greece the integrative substance is most developed at the level of subjects and programmes — the Slovenian interdisciplinary-links column and drawing curriculum, the Greek Skill Labs and the recent physics and mathematics curricula — whereas in Croatia and, in part, Italy it is asserted at policy level but left without subject-level

scaffolding. Belgium occupies a distinct position, having relocated integration into a competence-based architecture that is structurally explicit yet while leaving its implementation entirely to individual schools. Assessment maturity likewise varies: Slovenia and Greece exhibit genuinely advanced formative practice that is nonetheless uncoded or unaligned, whereas Italy presents an almost complete absence of STEAM assessment within mainstream provision. These differences refine the shared findings without undermining them; they indicate that the same underlying deficits surface at different points in different systems, but that the deficits themselves are shared.

2.8 Conclusions

The contextualisation analysis establishes a clear and consistent baseline. Across five independently analysed systems, the competences that immersive STEAM education seeks to develop are broadly present in latent form but are not explicitly labelled, deliberately designed, assessed against shared criteria, or mapped to the competency architecture of PISA 2025. The convergence of these findings — obtained from systems that differ markedly in governance, curriculum tradition and degree of reform — is the principal result of the chapter: it indicates that the deficits identified are structural features of contemporary curriculum structures and practice rather than the idiosyncrasies of any single context, and that they are therefore capable of being addressed through a common approach adapted to national contexts.

Three of these findings are foundational to the analysis that follows. The absence of explicit STEAM labelling makes existing good practice difficult to identify and transfer systematically, which is the precondition that any contextualisation must address before teaching can be deliberately designed and assessment developed for that competence. The marginalisation of the arts as a mode of scientific inquiry is both the most consistent gap and, given the strong isolated cases observed in Slovenia and Greece, the clearest opportunity for productive intervention. And the implicit and therefore difficult-to-use nature of the alignment with PISA— most acute for the epistemic and environmental dimensions — defines precisely what an explicit mapping at descriptor level must accomplish if teachers are to design for these competences deliberately. The divergences among systems, finally, are not an obstacle but a resource: the subject-level maturity of some systems and the structural explicitness of others supply concrete starting points from which a shared, transferable response can be constructed.

The analysis provides the following guidelines:

1. **Develop a Cross-Country Arts-in-STEAM Integration Model.** The near-universal absence of the Arts dimension from partner curricula is both the largest gap and the greatest opportunity. The course-design process should include the development of an Arts-in-STEAM integration model – grounded in the PISA 2025 framework,



LifeComp, and EntreComp. The model should show teachers in all five partner countries how design thinking, visual modelling, narrative argumentation, and speculative scenario building can enrich STEAM learning. The Slovenian Drawing in Geometry and Technology subject and the Italian 2026 curriculum's cross-curricular vision provide starting points;

2. **Prioritise STEAM Assessment Instrument Development.** Given the consistently substantial assessment gap across all five countries, the development of validated, adaptable STEAM formative assessment instruments should be treated as a priority output of the project's course-design and development process. These instruments should be: aligned with PISA 2025 competency descriptors; compatible with the Open Badge certification system; adaptable to different national assessment cultures (from the portfolio-based approach of Slovenia to the autonomy-based approach of Belgium); and available in partner languages through the ESEP platform. The Italian context – where no dedicated STEAM assessment instruments were identified within mainstream provision – should be the primary validation site.
3. **Design a Realistic Immersive Technology Implementation Pathway.** Building on the Schoolmakers implementation evidence, course implementations should design the ImSTEAM capacity-building programme as a multi-year, staged implementation pathway rather than a single training event. The first stage should focus on a core team of innovators within each partner school; the second on extending participation to a broader group of teachers; the third on institutional embedding and sustainability. This staged approach should be reflected in the structure of the micro-courses and in the design of the Skills Intelligence Hub, which should begin collecting data from stage one to inform the design of stage two.
4. **Establish a STEAM Skills Intelligence Loop.** The most significant system-level intervention that ImSTEAM can make across all five partner countries is the establishment of a functioning data loop connecting STEAM competence development in classrooms to curriculum improvement at institutional level. This loop requires PISA-aligned assessment instruments (to generate meaningful data), trained teachers (to collect it consistently), and analytical capacity (to interpret it and feed it back into teaching design). Course design should define the minimum viable data architecture for this loop in the first phase of the project, ensuring that all pilot activities generate data that can be meaningfully aggregated and analysed across partner countries.

3. The PISA skills pool

3.1 Overall approach

PISA defines what students should know in science.

If we know what students should know, can we derive from it what teachers need to know in order to teach it?

The Teacher Capability Pool is the foundational analytical output of Work Package 2. Its function is to set out, in a single structured inventory, the teacher capabilities that the **PISA 2025 Science Framework** implies: for every student competency and sub-ability that PISA establishes, the Pool identifies the corresponding capabilities a teacher must possess to enable students to attain that standard. The Pool is the product of a systematic derivation; it is not an adaptation of any existing taxonomy. Because no official taxonomy of teacher capabilities for immersive STEAM education exists, systematic derivation is the most methodologically defensible option.

The decision to construct a complete pool before specifying the pilot subset is deliberate. Selecting a smaller set at the outset — focusing only on capabilities that appear immediately feasible for the pilot— would define the project's view of teacher capability on partial evidence and could leave potentially relevant capability areas unexplored. Establishing the complete capability pool first ensures that coverage is demonstrable and that the subsequent selection of a pilot subset is a criteria-based decision taken against a known whole, rather than the by-product of early omission. The Pool is therefore larger than the project can pilot; that is by design, and it is what makes the derivation and selection process transparent and auditable.

Two distinct types of frameworks are used in the construction of the Pool, and the distinction matters. The first type — Level 1-source frameworks — define what students should know and be able to do. These are the starting points of the derivation: from each student ability they describe, the method asks what a teacher must know and be able to do in order to develop it. Three source frameworks are used

- **PISA 2025 Science** (OECD, 2023) is the primary source, providing the A-strand sub-abilities (Boxes 3–5) and the E-strand identity components (Box 11).
- **PISA 2022 Mathematics** (OECD, 2018) provides the M-strand, anchored in the four mathematical processes.
- **The AILit Framework** (OECD/European Commission, 2025, review draft) provides the AI strand, which is limited to the two domains that intersect with scientific reasoning.

The second type — Level 4-policy-alignment frameworks — serve a different and more limited function. DigCompEdu (Redecker, 2017) and GreenComp are not sources from which capabilities are derived. They are used, after the derivation is complete, to verify that the resulting capabilities are compatible with established European policy frameworks.



DigCompEdu provides this check for the science, mathematics and AI capabilities; GreenComp does so for the environmental capabilities. They locate the Pool in the European policy landscape; they do not generate its contents. The Pool is organised into four thematic sections called strands, each corresponding to one area of the STEAM remit and drawing on one source framework. The four strands are A (Science), E (Environment), M (Mathematics) and AI. Note that PISA 2025 Science alone feeds two strands — A and E — because the two parts of the framework address different aspects of scientific literacy: the A-strand concerns what students can do with scientific knowledge (explaining, enquiring, arguing), while the E-strand concerns how students relate to environmental science — their awareness, concern and sense of agency — which are dispositional rather than performative (explained in §3.2).

3.2 Overall architecture

The Pool is organised around a three-level mapping chain.

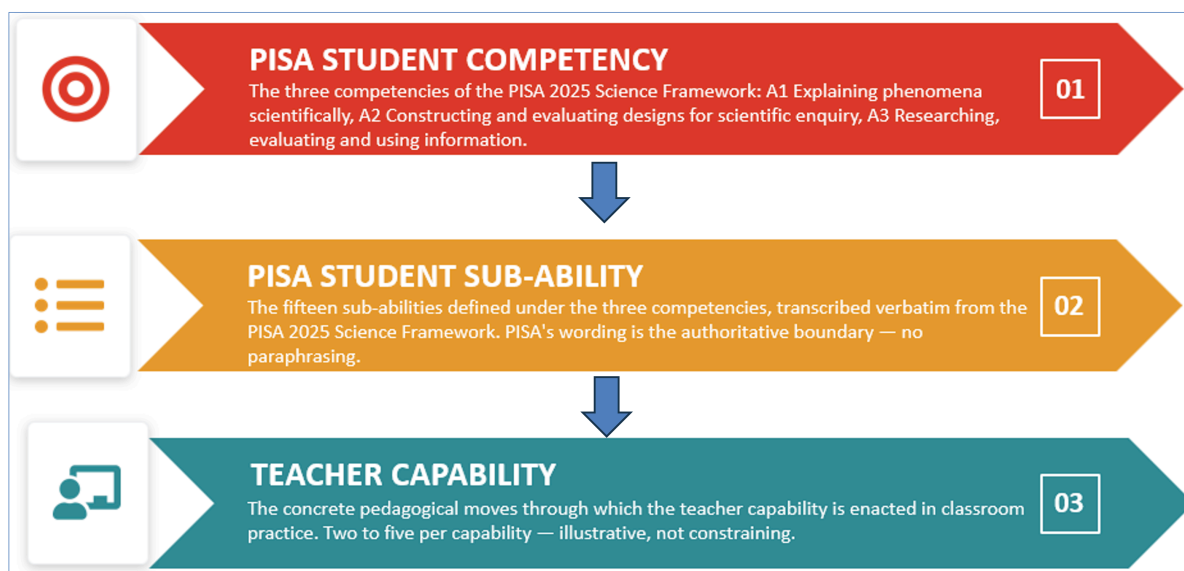


Figure 2 The three-level mapping chain.

Level 1 — student competencies. PISA 2025 organises scientific literacy around three top-level competencies: A1, explain phenomena scientifically; A2, construct and evaluate designs for scientific enquiry and interpret data critically; A3, research, evaluate and use scientific information for decisions and action. These three competencies are the entry point to the pool — they name the areas of student performance that the project ultimately seeks to improve. They cannot, however, serve as the practical units of analysis and derivation. Explaining phenomena scientifically covers too broad an area of teachers' professional knowledge and practice to ask meaningfully what a teacher must know to develop it: a teacher who can develop A1.1 (recall and application of scientific knowledge) and a teacher who can develop A1.6 (societal implications of science) may have almost no overlapping pedagogical knowledge requirements, yet both sit within A1. Working at Level 1 grain would produce either a trivially broad capability statement or a covert aggregation of several distinct ones.

Level 2 — student sub-abilities. Each top-level competency decomposes in PISA 2025 into a set of sub-abilities — the specific cognitive operations identified within each competency. A1 contains six sub-abilities (A1.1–A1.6), A2 four (A2.1–A2.4), and A3 five (A3.1–A3.5), for a total of fifteen across the science strand. The sub-ability is the operative unit of the entire derivation. It is specific enough that the five PCK questions — what content is central, what misconceptions arise, which teaching moves work, what understanding of the nature and production of scientific knowledge the teacher conveys, how students’ mastery can be demonstrated and assessed— have genuinely distinct answers for it. The sub-ability is also the unit at which the pilot selection was subsequently conducted: selecting at sub-ability grain means that selecting a sub-ability brings the complete associated set of teacher capabilities into the pilot, spanning the importance and positioning of the content within the curriculum. Selecting at a finer grain — choosing, say, the assessment capability of a sub-ability without its instructional capability — would produce microcourses that teach how to assess something they do not teach how to develop, which is pedagogically incoherent.

Level 3 — teacher capabilities. Each sub-ability yields up to five teacher capabilities, one for each component of the Refined Consensus Model of Pedagogical Content Knowledge that has a genuine answer for that sub-ability. This is the one-to-many relationship that constitutes the pool’s internal structure: one student competency maps to several sub-abilities; one sub-ability maps to several teacher capabilities; and the mapping in both directions is explicit, recorded, and traceable. Every teacher capability carries the identifier of the sub-ability it derives from and the PCK component it instantiates, so the full upward chain — from capability to sub-ability to competency to source framework — can be read in a single line.

A strand is the teacher-side counterpart of a single PISA student-side competency domain — the set of teacher capabilities derived by asking, of every item in that domain, what the teacher must be able to do to produce it in students.

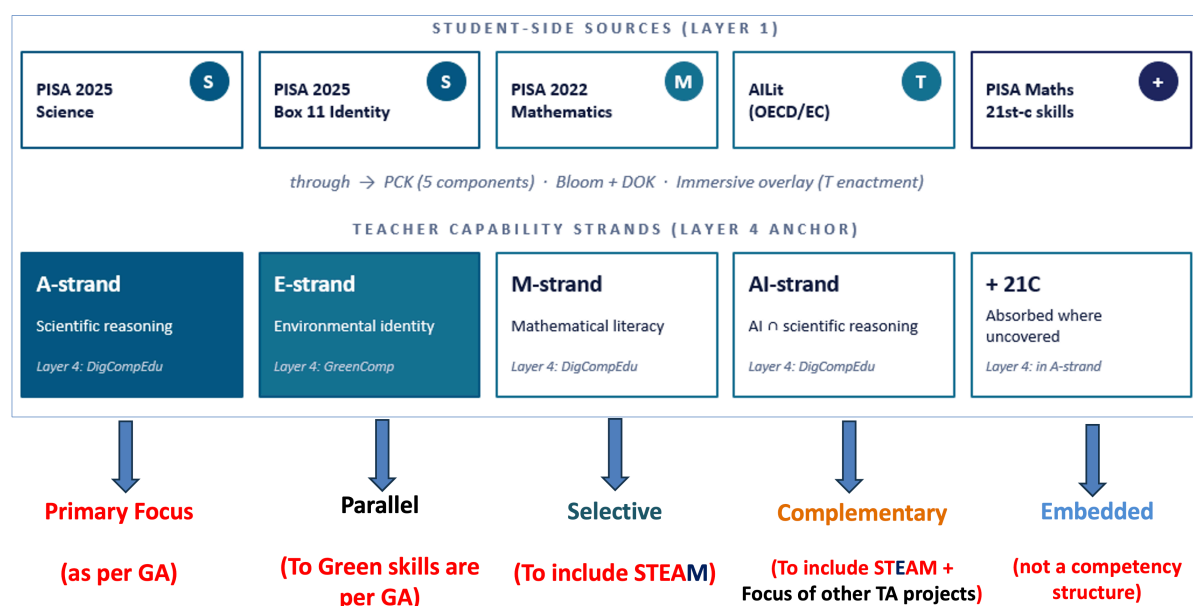


Figure 3 Student-side sources (layer 1) and teacher capability strands (layer 4 anchor)

The pool is organised into four strands, each anchored in a distinct Layer-1 source taxonomy.

A-strand (Science) — source: PISA 2025 Science, Boxes 3–5. The primary strand. Operative units: 15 sub-abilities (A1.1–A3.5). Every sub-ability is worked through all five PCK components — what to teach, where students struggle, how to teach it, what epistemic stance to convey, and how to assess it — producing up to five teacher capabilities per sub-ability and 75 in total.

E-strand (Environment) — source: PISA 2025 Science, Box 11. Unlike science competencies, environmental phenomena do not appear in PISA 2025 as a standalone competency with its own sub-ability decomposition. Instead, they appear as context categories and cross-cutting applications of the science competencies: a student may explain an environmental phenomenon (A1), design an enquiry into it (A2), or evaluate evidence about it (A3), but doing so draws on the same scientific competencies used for any other context. The only place where PISA treats the environment as an object in its own right is Box 11, which defines a Science Identity construct — not a competency, but a set of dispositions describing how students relate to environmental science: their awareness of environmental phenomena (E1), their concern about them (E2), and their sense of agency to act on them (E3). The E-strand uses these three identity components as its operative units — the functional equivalent of sub-abilities for this strand — and derives teacher capabilities from each. Inventing a sub-ability structure that the source does not provide would have manufactured detail it does not contain. The strand is accordingly compact: 13 capabilities. GreenComp replaces DigCompEdu as the policy-alignment check

M-strand (Mathematics) — source: PISA 2022 Mathematics. Operative units: four mathematical processes (M1 Mathematical Reasoning, M2 Formulating Situations Mathematically, M3 Employing Mathematical Concepts and Procedures, M4 Interpreting and Evaluating Mathematical Outcomes).

For each process, the derivation asked whether the corresponding teacher capability was already covered by the A-strand — that is, whether the pedagogical demands of teaching that mathematical ability are sufficiently similar to those of teaching the nearest scientific ability that a separate capability would add nothing. For M1, M2 and M4 the answer was no: the pedagogical work of developing mathematical reasoning, of teaching students to translate a real-world situation into mathematical terms, or of helping them interpret and evaluate mathematical results, is distinct enough from its scientific counterpart to warrant dedicated capabilities. M3 is treated more lightly because teaching procedural mathematical work — applying known concepts, facts and algorithms — falls within conventional mathematics teacher preparation and does not require a dedicated capability in this pool. Total: 11 capabilities.

PISA 2022 also identifies eight 21st-century skills — such as creativity and self-direction — that are not specific to mathematics and were therefore not incorporated into the M-strand. Two had no close counterpart in the existing pool and were added as new A-strand capabilities, identifiable by the .f suffix in their code (T-A1.5.f and T-A2.2.f). The remaining six overlap sufficiently with existing A-strand capabilities to be recorded as a second derivation source on those items rather than as new entries.

AI-strand — source: AILit Framework. AI literacy as a full competence domain is already the object of dedicated Erasmus+ Teacher Academy projects; ImSTEAM does not replicate that work. The strand is bounded to the intersection of AI literacy and the project's primary scope — scientific reasoning, modelling, and evaluation of evidence. The AILit Framework defines

four domains; only Engaging with AI (AI1) and Managing AI (AI3) carry teacher-pedagogical demands within that intersection, in total having six capabilities.”

3.3 The Backward methodology

PISA 2025 organises student scientific literacy under three competencies. Each competency is in turn decomposed into a set of student sub-abilities — the specific cognitive operations that PISA identifies and assesses within it. A1 (Explain phenomena scientifically) contains six sub-abilities; A2 (Construct and evaluate designs for scientific enquiry; interpret data critically) contains four; A3 (Research, evaluate and use scientific information for decisions and action) contains five. These fifteen sub-abilities constitute the operational units of the derivation. The competency level sets the broad domain; the sub-ability level provides sufficient specificity to ask a precise question: **what must a teacher know and be able to do in order to develop this particular student ability?**

That question is answered through Pedagogical Content Knowledge. Shulman (1986) introduced PCK to name the form of knowledge that is specific to teaching — distinct from knowing a subject and distinct from knowing how to teach in general. PCK integrates the two: the ways of representing and formulating a topic that make it comprehensible to a particular group of learners, combined with an understanding of what makes that topic specifically easy or difficult and of the preconceptions that students typically bring to the topic. The decisive property of PCK is that it is topic-specific: the knowledge a teacher needs to develop students' ability to construct scientific arguments from data is genuinely different from the knowledge needed to develop their ability to evaluate experimental designs, even though both sit within the domain of science teaching. A generic statement of desirable teacher competence cannot capture this specificity; only a derivation conducted at topic level can capture it.

The **Refined Consensus Model of PCK** (Carlson & Daehler, 2019), developed through an international summit of PCK researchers, provides the operative framework ImSTEAM uses. It identifies five components that together cover the principal questions a teacher must be able to answer in order to teach a specific topic well. For each student sub-ability, the method asks all five:

1. **Curricular saliency.** What content is genuinely central to this sub-ability, where does it sit in the curriculum sequence, and what deserves emphasis rather than incidental treatment?
2. **Knowledge of students' understanding.** What preconceptions, misconceptions and characteristic difficulties do students bring to this specific topic, and how do they surface in student work?
3. **Instructional strategies.** What topic-specific representations, analogies, tasks, sequences and teaching moves develop this sub-ability — not generic technique, but what works for this content?
4. **Orientations to teaching science.** What stance toward the nature of science does the teacher convey while teaching this topic, and how does that framing shape what students take science to be?



5. **Assessment knowledge.** How does the teacher elicit and judge whether students have genuinely developed this ability, as distinct from having produced a correct-looking answer?

For each sub-ability, the method asks all five PCK questions in sequence, producing one teacher capability for each question that yields a meaningful and distinct answer — so **between 1 and 5 teacher capabilities per sub-ability**. The five components are independent dimensions of teacher knowledge: a teacher who commands the instructional strategies for a given sub-ability holds a distinct dimension of professional knowledge from the teacher who understands the misconceptions students bring to it, or who knows how to design assessment tasks that distinguish genuine mastery from a correct-looking answer. Each component therefore gives rise to a distinct, independently specified capability. The resulting set of between 1 and 5 capabilities per sub-ability constitutes a comprehensive specification of the pedagogical demands that sub-ability places on a teacher: what must be known and done across every dimension of teaching it well. A microcourse built on one sub-ability accordingly claim systematic coverage of the principal pedagogical dimensions, because the procedure works through every dimension in turn and records a reasoned judgement for each one.



Figure 4 Prioritisation of connections between student competencies, PISA sub-abilities and teacher capabilities.

3.4 Pool metrics

Applying the methodology across all four strands produced the capability pool presented in full in Annex I. The counts below report what that derivation yielded, strand by strand. The asymmetry is deliberate: the A-strand is the largest because PISA Science is the primary anchor, interrogated to full PCK saturation across all fifteen sub-abilities; the Mathematics and AI strands are leaner, include capabilities only where a PCK question yields a meaningful and distinct response, so their lower per-item yield is by design rather than are

underdeveloped; and the Environment strand sits between, compact because it derives from a three-part identity construct rather than a parallel competency set.

Table 6 Strands, layer-1 sources, top-level and skills pool.

Strand	Layer-1 source	Top-level items	Skills pool
A-strand (Science)	PISA 2025 Science · Boxes 3–5	15 sub-abilities (A1.1–A3.5)	75
— of which 21C absorption (.f)	PISA 2022 Maths · 21C inventory	2 (T-A1.5.f, T-A2.2.f)	2
E-strand (Environment)	PISA 2025 Science · Box 11 · GreenComp · UNIPA WP2	3 identity components (E1–E3)	13
M-strand (Mathematics)	PISA 2022 Mathematics · 4 processes	4 processes (M1–M4)	11
AI-strand (AI literacy)	AILit Framework · 4 domains	2 domains in scope (AI1, AI3)	6
Total			105

The following table summarizes the Strands covered by the Skills Pool and the corresponding sub abilities.

Table 7 Summary of Strands covered by the Skills Pool and sub-abilities.

A-strand · Science — 15 sub-abilities under 3 PISA 2025 Science competencies	
A1 · Explain phenomena scientifically (Box 3)	
A1.1	Recalling and applying scientific knowledge
A1.2	Using and translating between forms of representation
A1.3	Making and justifying scientific predictions
A1.4	Identifying, constructing, and evaluating models
A1.5	Recognising and developing explanatory hypotheses
A1.6	Explaining the societal implications of scientific knowledge
A2 · Construct and evaluate designs for enquiry; interpret data critically (Box 4)	
A2.1	Identifying the research question of a scientific study
A2.2	Proposing an appropriate experimental design
A2.3	Evaluating the fit of an experimental design to its question
A2.4	Interpreting data, drawing conclusions, and evaluating their relative merits
A3 · Research, evaluate and use scientific information for decisions and action (Box 5)	

A3.1	Searching, evaluating and communicating the merits of information sources
A3.2	Distinguishing claims by type and warrant
A3.3	Constructing scientific arguments from data
A3.4	Critiquing standard flaws in science-related arguments
A3.5	Justifying decisions using scientific arguments on contemporary issues
E-strand · Environment — 3 Science Identity components (Box 11)	
E1	Environmental Awareness
E2	Environmental Concern
E3	Environmental Agency
M-strand · Mathematics — 4 mathematical processes (PISA 2022)	
M1	Mathematical Reasoning
M2	Formulating Situations Mathematically
M3	Employing Mathematical Concepts, Facts, Procedures and Reasoning
M4	Interpreting, Applying and Evaluating Mathematical Outcomes
AI-strand · AI literacy — 2 AILit domains in scope	
AI1	Engaging with AI (in scientific reasoning contexts)
AI3	Managing AI (selectively, in scientific work)

3.5 Limitations

A small number of boundary conditions qualify the derivation and are recorded here.

The first concerns the source itself: the PISA 2025 Science Framework is a Second Draft, and the derivation is anchored on the version current at the time of the project's design. Each traceability chain will be checked and realigned against the final framework once it is published; because the chain is anchored at the level of competencies and sub-abilities rather than wording, the capability identifiers, PCK anchors and recorded fields are not expected to change when that update is made.

The second concerns the method's status: performing backward design at the level of an individual teacher capability is not a procedure adopted in its entirety from any single source but an adaptation specific to this project, and it is presented as a documented contribution traceable through the four-layer chain rather than as the application of an established pre-existing protocol.

The framework is also informed by the OECD Teaching Compass for taxonomy and policy alignment and is structurally compatible with it, but it does not claim to constitute an implementation of the Teaching Compass: the Compass articulates what teachers need



without specifying how to derive teacher capabilities from a student-outcome framework, so the relationship is one of structural compatibility rather than direct implementation.

Two further points concern coverage, and both are matters of design rather than gaps. The environmental treatment is deliberately compact and identity-based because that reflects how environmental science is actually structured in the source — as a sub-construct of science identity and a cross-cutting context — rather than as a parallel competency set mirroring the science competencies; inventing such a parallel structure would have manufactured detail the source does not contain. Similarly, the Mathematics and AI strands are populated selectively, only where a PCK question yields a meaningful and distinct pedagogical requirement for teachers, and the PCK dimensions left unpopulated in those strands are recorded as deliberate omissions rather than treated as missing entries. In both cases this methodological restraint is deliberate: the method specifies a capability where both the source and the pedagogy support one, and refrains from doing so where they do not.



4. Workshop

4.1 Scope and goals

The WP2 workshop convened the full consortium at the project kick-off meeting in Patras on 14 May 2026. Its remit followed from the state of the WP2 method at that point: the derivation chain had produced a pool of approximately one hundred teacher capabilities, designed to provide comprehensive coverage, which the ImSTEAM framework retains in full. The pilot builds a subset of that pool, and the purpose of the workshop was to determine which capabilities the subset would contain and to begin shaping them for downstream use.

The work was organised as three sessions, each with a distinct objective.

- **Session 1 — Prioritise.** The first session addressed which capabilities, among the full pool, matter most for the pilot. Proceeding from the premise that training programmes must be developed for a tractable number of capabilities, it produced the **pilot subset**: the ranked selection of capabilities the project carries into WP3 microcourse design.
- **Session 2 — Describe.** Taking the pilot subset as input, the second session expressed each selected capability as the concrete teaching actions that enact it, recording three to five Bloom-aligned **action keywords** per capability. This output provides the basis for the learning objectives and assessment items developed in WP3 and WP7.
- **Session 3 — Localise.** Drawing on the per-country desktop survey, the third session opened the question of which capabilities are present uniformly across the consortium and which call for country-specific shaping, conducted as a round-table discussion among the five national delegations. Its output is a **national contextualisation map** spanning Greece, Belgium, Italy, Slovenia and Croatia.

Prioritisation covered the science, environmental and mathematics strands (A1, A2, A3, E, M); the AI strand was carried in the wider pool as a research deliverable, with pilot selection focused on the five strands in scope. Selection proceeded against four ordered criteria — PISA leverage, immersive density, gap evidence and feasibility. Across the three sessions the workshop delivered three consolidated outputs: a ranked pilot subset of fifty-three teacher capabilities across thirteen microcourses, the action-keyword elaboration of each, and the five national localisation views.

Thirty (30) experts participated in the workshop.

4.2 Methodology

The prioritisation was conducted during a live, in-person consortium session, with the online survey — the third stage of the WP2 method — reserved for gathering input from a broader



population and providing external confirmation. Two properties of the task made this the right setting. First, the four selection criteria pull in different directions: an item strong on PISA leverage may be modest on feasibility, and reconciling such tensions requires real-time deliberation and reasoned argument, which a live session supports, and an asynchronous form does not. Second, work packages WP3 to WP5 build directly on the result, and their success depends on partners owning the selection. The survey accordingly does the work it is suited to — confirming the decision across a wide population — while the final, binding decision was made during the in-person workshop.

Participants worked in four groups of five to six, each composed to ensure diversity along two dimensions: by country, so that no two participants from the same country were assigned to the same table, and by stakeholder type, spanning academics, teacher trainers, associate partners and Community-of-Practice members. The mix means that each ranking produced by a group already embodies a consortium-level reconciliation of national perspectives rather than a single national position, and that each partner's local knowledge is distributed across several groups. The purely national perspective is recovered deliberately at the Session 3 localisation stage and in the national online survey.

Each group held a fixed table for the full ninety minutes; the strand sheets rotated between tables, in a structured adaptation of the world-café deliberation format. This produced four separately recorded group rankings of every strand within a single session, which is what allows the consolidation to distinguish an item every group ranked highly from one that only just survived, and to treat the two differently. Independence is the precondition for that aggregation to carry information: pooled judgements improve the collective answer only where the judgements are formed independently. Groups were therefore asked to read the rankings left by earlier groups, take them seriously, and then record their own — so that strong prior reasoning could inform a judgement while reducing, as far as possible, the influence of the order in which groups encountered the sheets. Disagreement that survives this treatment is retained as signal, marking the abilities on which national pedagogical traditions genuinely differ, which is exactly what the localisation stage needs to identify.

The same architecture — the four mixed groups, the fixed tables, the rotation — carried into Session 2. However, in this session the goal was not to rank items but to describe them.

Four A2 sheets were placed in the room, one for each of the three PISA science competencies and one for the environmental strandE), each accompanied by one or two A4 reference sheets listing the teacher capabilities grouped under every sub-ability. Each group read the teacher's capabilities, discussed each sub-ability together as a five-country team, and wrote down the codes of the sub-abilities they believed should remain in the pilot.

The unit of selection was the **student sub-ability**, not the individual teacher capability. Each sub-ability carries a comprehensive set of teacher capabilities covering all five PCK

dimensions examined spanning the full teaching arc — core content, student misconceptions, instructional strategy, framing and assessment — so selecting at this grain brings a complete teaching unit into the pilot, and a micro-course built around it can provide systematic coverage of the full pedagogical requirements associated with that ability. The teacher capabilities functioned as the decision aid: groups read down through the capabilities under a sub-ability to see what it looks like pedagogically and then decide whether the sub-ability as a whole was worth including in the pilot.

Every group applied the same four criteria, presented in a fixed order of weight that encodes the project's logic:

- **PISA leverage** — placed first because it carries the theory of change: teacher upskilling is justified insofar as it plausibly improves the student ability PISA measures.
- **Immersive density** — second, as the project's distinctive contribution: the capability genuinely benefits from VR, AR or MR, which grounds the case for immersive rather than conventional training.
- **Gap evidence** — establishing additionality: a real, identifiable national training gap, evidenced by the desktop survey and the partners' own context.
- **Feasibility** — operating as a constraint: whether the item can be built, delivered and assessed within the 36-month timeline.

The number of items to be retained from each strand was fixed in advance as a printed selection rule, so that the pilot's scale was set as a deliberate budget and each group faced genuine trade-offs — adding one item meant releasing another. The cut-offs reflect both delivery capacity and the relative weight of each competency in the science framework:

Table 8 Strands, items at the station and selection rule followed in the workshop.

Strand	Items at the station	Selection rule
A1 — <i>Explaining phenomena</i>	6 sub-abilities	Top 3 of 6
A2 — <i>Designing enquiry, interpreting data</i>	4 sub-abilities	Top 2 of 4
A3 — <i>Researching, evaluating, deciding</i>	5 sub-abilities	Top 3 of 5
E — <i>Environmental identity</i>	4–5 capabilities per component	Top 3 per component (E1, E2, E3)

Session 1 was confined to ranking. The pool was treated as fixed for the purposes of the ranking exercise: where a group found a capability unclear, redundant or differently classified, the issue was flagged with the facilitator and captured for the online validation survey. This procedure preserves the methodological traceability of any subsequent change to the pool. Each rotation concluded, where time allowed, with a five-minute report-back from each group surfacing the group's rankings, points of disagreement or unexpected findings, and country-specific considerations to carry into Session 2.

4.3 Session 1

Participants worked in **four mixed groups** (labelled A, B, C and D), deliberately composed of participants from different countries and institutions, rather than organised along national or institutional lines, so that each ranking reflected a consortium-wide judgement rather than a single national position. Each group occupied a fixed table; the strand sheets rotated between tables during two timed rotations so that every group ranked every strand.

The prioritised subset agreed during the plenary session is set out below, with the PISA/PCK descriptor for each selected item.

Table 9 Prioritised subset selected by the workshop participants.

Code	Selected teacher-capability target
A1.2	Using and translating between forms of representation
A1.3	Making and justifying scientific predictions
A1.4	Identifying, constructing and evaluating models
A2.2	Proposing an appropriate experimental design
A2.4	Interpreting data, drawing conclusions and evaluating their relative merits
A3.1	Searching, evaluating and communicating the merits of information sources
A3.3	Constructing scientific arguments from data
A3.5	Justifying decisions using scientific arguments on contemporary issues
E1.a	Curricular saliency — building cumulative socio-ecological systems awareness
E1.b	Students' understanding — anticipating linear-causal default reasoning
E1.d	Orientations — framing environmental phenomena as legitimate scientific content
E2.a	Curricular saliency — connecting environmental phenomena to lived experience
E2.b	Students' understanding — eco-anxiety versus constructive concern
E2.c	Instructional strategies — cultivating evidence-based hope through success cases
E3.b	Students' understanding — individualisation of structural environmental problems
E3.c	Instructional strategies — analysing action at different scales
E3.e	Assessment — authentic impact-analysis tasks with structural framing

Agreement between the four groups was measured with Kendall's coefficient of concordance (W), where 0 indicates no agreement and 1 indicates identical rankings.

Table 10 Kendall coefficient of concordance for the various choices of the participants.



Strand	A1	A2	A3	E1	E2	E3
Kendall's W	0.36	0.40	0.29	0.30	0.70	0.19

Concordance was highest for E2 ($W = 0.70$), where the consortium agreed strongly, and lowest for E3 ($W = 0.19$) and A3 ($W = 0.29$). The two least-concordant strands are also the two whose stations were ranked in the later time-block, so part of the lower agreement may reflect fatigue and time pressure rather than substantive disagreement. E3, however, carried a genuine, documented disagreement.

Comparing the complete set of selections each group made (seventeen picks per group) using the Jaccard similarity coefficient reveals that the four groups did not rank independently but fell into two recognisable blocs.

Table 11 Comparison of the selection of each set using the Jaccard similarity coefficient.

Jaccard	Group A	Group B	Group C	Group D
Group A	—	0.55	0.42	0.42
Group B	0.55	—	0.55	0.42
Group C	0.42	0.55	—	0.70
Group D	0.42	0.42	0.70	—

Groups C and D were near-identical (0.70 similarity; 14 of 17 picks shared), and groups A and B aligned closely (0.55). Cross-bloc similarity fell to about 0.42. The consortium therefore expressed two coherent prioritisation tendencies rather than four scattered opinions. The bloc split was not random: it tracked the PCK component each capability represents. Within the pool, the trailing letter of each code encodes a PCK anchor (a = curricular saliency, b = students' understanding, c = instructional strategies, d = orientations, e = assessment). Read against this structure, a single pedagogical disagreement recurs across strands:

- **Groups C and D** consistently prioritised the foundational, curriculum-grounding capabilities — A3.1 (source evaluation), E1.a and E3.a (curricular saliency).
- **Groups A and B** consistently set those aside as “too basic — teachers can already identify such curricula” and reached instead for higher-order instructional capabilities (E1.c, the action/instructional moves) and higher-Bloom sub-abilities (A1.5, developing explanatory hypotheses).

This is best read as one disagreement appearing in several places — **build from curricular grounding** (C/D) versus **assume the grounding and teach the harder pedagogical moves** (A/B) — rather than as multiple independent ties. It is the most important interpretive finding

of Session 1 and the principal distinction that should guide Session 3 localisation and WP3 scenario design

Five of the six strands produced a plenary verdict identical to the result generated by the Borda count, including all three A-strand ties. The single departure was E3. The Borda count placed E3.a (equity-relevant curricular content) in the pilot and excluded E3.c (analysis of action at scales). For E3, the plenary reversed this result, admitting E3.c and excluding E3.a — the bloc disagreement of Section 5.3 resolving live in favour of the A/B reading that E3.a was too foundational to warrant a dedicated immersive microcourse. Notably, the same argument was not accepted in the case of E1, where the equivalent foundational capability (E1.a) was retained. The plenary therefore considered each capability individually an enactment-heavy combination rather than by a blanket rule against curricular-saliency items. The final E3 trio (E3.b, E3.c, E3.e) spans students' understanding, instructional strategies and assessment — a deliberately enactment-heavy combination. It is recommended that the rationale for the E3.a/E3.c swap be recorded in the minutes while the reasoning remains fresh, since at present only the verdict, not its reasoning, is preserved in the workshop record.

Coding the free-text justifications against the four printed criteria, plus the themes that emerged spontaneously, produces a clear and somewhat unexpected picture of how groups actually reasoned.

Table 12 Reasoning theme and mentions by workshop participants.

Reasoning theme	Mentions	Where
Critical thinking	5	A2 (A, D); A3 (B, C, D)
Real-world relevance / life skills	3	A1 (C, D); A3 (D)
Bloom level / cognitive demand	2	A1 (A, D)
Pedagogical coherence / process logic	2	A1 (B); A2 (B)
Immersive density (VR/AR/MR)	1	A1 (A) only
Gap evidence	1	A1 (A) only
Feasibility	0	—

Groups based their reasoning primarily on cognitive demand, critical thinking and real-world relevance. The project's two most distinctive criteria — **immersive density** and **gap evidence** — were each cited only once, and feasibility not at all. This does not invalidate the rankings, but it has an important implication: the immersive-technology rationale, which is central to ImSTEAM's identity, did not arise spontaneously from partners' prioritisation. It will need to be developed and made explicit during WP3 scenario design rather than assumed to follow from the Session 1 selections.

4.4 Session 2

Session 2 took the prioritised subset settled at the close of Session 1 — the eleven microcourses and forty-nine teacher capabilities of the science and environmental strands — and elaborated each surviving item into the concrete teacher actions that enact it: what a teacher must be able to do in the classroom to enact or demonstrate that capability. The session therefore bridges the decision about what should be taught, made in Session 1, and the specification of how teachers should enact each selected capability.

The session retained the Session 1 arrangement: the same four mixed groups (A, B, C, D), each at a fixed table, with the strand sheets rotating between them, and each group writing only in its own space on the A2. The surviving codes were presented as fixed labels; groups worked beneath each one rather than re-opening the selection. For each surviving capability, groups recorded three to five action keywords describing the teacher action required, set at or below the cognitive level assigned to the capability and drawn from a Bloom verb bank supplied on the sheet. Participants were instructed to ensure that each keyword set spanned — the prior knowledge a teacher needs about where students start, the instructional moves that teach the capability, and the means of assessing it — so that the elaboration reproduced in miniature the PCK structure of the underlying pool entry.

The sequence followed at the kick-off departs in one respect from the pre-circulated agenda, recorded here for traceability. The agenda had framed the second session as a plenary consolidation step; in the event, the consolidation was completed live at the close of Session 1, where the consortium agreed the final subset before the break. The Session 2 slot was accordingly given to the action-keyword elaboration described here, and presenting the survivors as fixed labels served the additional purpose of confirming the Session 1 verdict.

Because the sheets rotated within a single ninety-minute session, coverage varied across strands, and that variation is the principal determinant of how strongly each part of the result is evidenced. The table records which groups contributed to each strand, the granularity at which they recorded, and the resulting evidence strength.

The A2 competency strand was supported by the strongest evidence, with all four groups contributing per- capability notes; A3 and A1 were elaborated per skill by three and two groups respectively. The environmental strand was reached by three groups, which recorded mainly at the level of the three components (Awareness, Concern, Agency) rather than the individual capabilities beneath them, so the environmental action material stands at component level. The surviving items themselves are settled regardless of coverage, since their identity was fixed in Session 1; coverage governs the volume of qualitative material attached to each, not whether it is in the pilot.

The recorded material was richer than a simple list of keywords, taking three overlapping forms: short teacher-action phrases (the intended output, for example connecting a research question to measurable variables, or scaffolding student-led design); fuller instructional-method notes (for example co-creating assessment rubrics with students, or creating an



environment in which failure is treated as a legitimate part of enquiry and learning for student enquiry); and concrete classroom examples (for example using contemporary public-health messaging to connect data to behaviour, or documented waste-management successes to build evidence-based environmental hope). Several groups tagged entries with a cognitive level, applied helpfully but unevenly across groups. The mixed register and partial Bloom tagging are what make a distillation step the bridge between the raw sheets and a usable framework layer.

The output accordingly serves two purposes, kept distinct so that each is fit for its use. The first is a distilled action layer: a small, normalised, Bloom-consistent set of teacher-action keywords per capability, derived from the raw input and incorporated into the teacher-capability framework. The second is a lesson-seed annex: the raw notes, methods and examples preserved with codes normalised, retained as direct input to WP3 scenario design. Three rules govern the distillation into the framework layer:

- **Pull to Bloom.** Every keyword is held at or below its capability's cognitive level; an entry recorded above its anchor is pulled down to it.
- **Include only where recorded per skill.** The action-keyword layer enters the framework where it was captured per capability — that is, for the A-strand — so that per-capability keywords rest on evidence rather than inference.
- **Environmental layer carried from the pool.** Since the environmental input was recorded per component, the environmental action layer is carried from the validated Pool v1.5 capability descriptions, which hold per-capability detail the workshop sheets recorded at the level of the broader components.

Each capability's action layer additionally carries a provenance marker recording how strongly it is evidenced — strong for A2, good for A3, indicative for A1, and pool-carried for the environmental strand — so that an entry supported by provisional or limited evidence is visibly flagged for the post-Patras validation rather than presented as equal in weight to a strongly evidenced one.

Session 2 produced a teacher-action layer for all seventeen surviving science and environmental items. The table records the outcome by strand.

Table 13 Session 2 outcome by strand.

Strand	Capabilities elaborated	Evidence	Action layer in the framework
A1 — Explaining phenomena	A1.2, A1.3, A1.4	Indicative (2 groups)	Distilled per skill; flagged for validation
A2 — Designing enquiry, interpreting data	A2.2, A2.4	Strong (4 groups)	Distilled per skill
A3 — Researching, evaluating, deciding	A3.1, A3.3, A3.5	Good (3 groups)	Distilled per skill
E — Environmental identity	E1·a/b/d, E2·a/b/c, E3·b/c/e	Component-level (3 groups)	Carried from the Skills Pool

Across the sheets, each capability was elaborated across three layers of teacher action— the prior knowledge a teacher needs about where students start, the instructional moves through which the capability is developed, and the means of assessing it. Representative keywords recorded by the groups include connecting a research question to measurable variables and scaffolding student-led design (instructional); co-creating assessment rubrics with students and assessing students’ reasoning rather than focusing solely on the final answer (assessment); and grounding abstract content in concrete cases — contemporary public-health messaging to link data to behaviour, documented waste-management successes to build evidence-based environmental agency.

4.5 Conclusions

Session 1 delivered a defensible, multi-country pilot subset of **eleven microcourses and forty-nine teacher capabilities**, with the mathematics strand to be added (to thirteen / fifty-three) on completion of its online validation. The consolidation showed particularly strong agreement on the project’s central priorities— the project’s central A-strand anchors (A1.4, A3.3) and the strongest environmental capability (E2.a) commanded unanimous or near-unanimous support — and the pre-committed tie-breaking rule resolved every tie produced by the formal ranking procedure without recourse to a coordinator casting vote, with the sole exception of the substantive decision on E3 taken openly in plenary.

Three findings carry forward:

- **A real, structured disagreement exists** between a “foundational grounding” tendency (Groups C/D) and a “higher-order pedagogy” tendency (Groups A/B). This axis, not the individual ties, is the substantive output to manage in Session 3 localisation and course design.
- **The immersive rationale must be supplied deliberately.** Partners prioritised on cognitive demand and relevance, not on immersive density; WP3 cannot assume the immersive case follows automatically from these selections.
- **A pool-integrity item is open:** the absence of an assessment capability in E1 and E2, to be resolved in the course capability audit.

Session 2 provided the operational tier beneath each capability: where the pool entry states what a teacher must know and be able to do, the distilled keywords specify how that capability is enacted through concrete classroom practices, organised under prior knowledge, instruction and assessment. Only the A-strand material recorded at capability level is incorporated directly, each capability carrying a provenance marker (strong, good, or indicative) that travels with it into the framework; the environmental layer is carried from Pool at the same per-capability grain. The keywords are held at or below each capability’s Bloom level, so the framework’s cognitive ceiling is preserved end to end. From there the layer may also support two downstream uses— as the seed for WP3 learning objectives and assessment items, and, in its raw form, as the lesson-seed annex supplied to scenario design.



5. Online Survey

This chapter reports the final stage of the method set out in Chapter 1: the online survey through which the work of the preceding chapters is carried to a wider population of teachers. The chapter opens with the identity, instrument and method of the survey, including the decisions governing how its results are read (Section 5.1); it then presents the profile of the respondents (Section 5.2), the findings for each national system (Section 5.3), and the cross-national analysis (Section 5.4), before stating what the survey establishes for the framework and for the work that builds on it (Section 5.5).

5.1 Purpose, Instrument and Method

5.1.1 Identity and purpose of the study

The survey occupies the validation phase of the WP2 methodology summarised in Table 1. The phases that precede it are deliberately internal to the consortium: the desk review was conducted by the partners, the derivation was performed against the source frameworks, and the prioritisation was decided in a live workshop whose participants were drawn from the partnership and its immediate networks. The survey is the point at which the method turns outward. Its function is to extend the result to a population of practising teachers who took no part in producing it, and to do so without reopening the framework: the capabilities themselves were fixed by the derivation and the workshop, and the survey gathers evidence around them rather than votes upon them.

Within this function the survey pursues three specific purposes, and the design of the instrument follows directly from them. The first is confirmation: to establish whether the competency areas from which the framework is derived are recognised by teachers as present in their curricula and as genuinely difficult to teach in depth, which is the empirical premise on which a training intervention rests. The second is localisation: to gather, country by country, the evidence on curriculum emphasis, teaching constraints and usability conditions that the national-adaptation guidance of Chapter 6 requires, thereby recovering the purely national perspective that the workshop's mixed-country groups deliberately set aside. The third concerns the medium: to test whether the pedagogical affordances through which the project treats immersive technology correspond to the difficulties teachers actually report, and to establish the baseline of teacher readiness against which the training must be pitched. The survey is thus not an opinion poll on the framework but a structured examination of the conditions under which the framework will be used.

5.1.2 The instrument

The questionnaire was developed from a common English master and comprises a respondent profile followed by ten question sections, summarised in Table 14. Its architecture embodies two designed pairings. Sections Q1 and Q2 present the same eight competency areas twice — first asking how strongly the national curriculum emphasises each area, then



how deeply the respondent is actually able to teach it — so that the gap between curricular aspiration and classroom enactment can be observed directly rather than inferred. Sections Q5 and Q6 apply the same logic to the medium: Q5 asks how often teaching is impeded by each of five sources of difficulty — phenomena too small or large to see, processes too slow or fast to observe, invisible mechanisms, experiments too dangerous or costly to run, and real places or situations that cannot be brought into the classroom — and Q6 asks how much benefit immersive activities would bring against the five affordances that answer them. The correspondence between the two lists is exact by construction, and it is this correspondence that allows the survey to test the project's central premise about the medium against teachers' reported practice. The remaining sections establish the constraints under which teaching operates (Q3), the difficulty of engaging students in scientific practices (Q4), the conditions an immersive activity must satisfy to be usable (Q7), and the respondent's self-assessed readiness and prior experience (Q8, Q9), with open questions capturing what the closed items do not.

Table 14 Structure of the online survey instrument

Section	Focus	Items	Response format
Profile	Country, subject(s) taught, level, years of science teaching, age group, gender, school type, geographic area	8	Categorical
Q1	Emphasis of the national curriculum on the eight competency areas derived from the PISA 2025 framework	8	Five-point intensity scale
Q2	Depth to which each of the same eight areas is actually taught	8	Five-point frequency scale
Q3	Constraints limiting science teaching (class size, laboratory provision, devices and connectivity, language availability of materials, examination pressure, curriculum load, professional development, time, institutional support)	10	Five-point intensity scale
Q4	Difficulty of engaging students in thinking and working as scientists across eight practices	8	Five-point difficulty scale
Q5	Sources of difficulty in conveying scientific content (scale, speed, invisibility, danger or cost, inaccessibility of real settings)	5	Five-point frequency scale
Q6	Potential benefit of immersive activities against the five corresponding affordances, with an open extension	5 + 1 open	Five-point intensity scale; free text
Q7	Conditions an immersive activity must satisfy to be usable in class (time, class size and devices, readiness for use,	6	Five-point importance scale

	curriculum and examination alignment, teacher training, language availability)		
Q8	Self-assessed readiness for immersive teaching (knowledge of tools, integration into lesson design, technical skill, classroom management, assessment through immersive activities)	5	Five-point preparedness scale
Q9	Prior experience with immersive tools	1	Single choice, five categories
Q10	Open comments	1	Free text

5.1.3 Deployment and fieldwork

The instrument was deployed as five national versions — Italian, Greek, Dutch, Slovene and Croatian — each translated from the English master and adapted to the terminology of the national education system, so that items referring to curriculum documents, school levels and examination arrangements read naturally in each context. The survey was administered online through a standard form platform and distributed through the partners' national networks of schools, teacher communities and professional contacts. Fieldwork ran from 22 June to 20 July 2026. Participation was anonymous and voluntary; the first item of every version presented the study information and recorded informed consent, and no response was accepted without it. All 292 submissions received across the five versions carried consent and were retained. Table 15 reports the distribution of responses by national version after the reassignment described below.

Table 15 National versions, languages and valid responses (fieldwork 22 June – 20 July 2026).

National version	Language	Valid responses	Share of total
Italy	Italian	202	69.2%
Greece	Greek	48	16.4%
Belgium — Flemish Community	Dutch	21	7.2%
Slovenia	Slovene	15	5.1%
Croatia	Croatian	6	2.1%
Total		292	100%

5.1.4 Methodological decisions

Four decisions govern the treatment of the data, and they are stated here so that every figure in this chapter is read as intended. First, the five language versions are harmonised onto

common numeric scales: each closed item uses a five-point response scale whose anchors were translated equivalently across versions, and responses are mapped onto a common 1–5 metric, with the anchor wording of each section documented alongside the questionnaire in Annex V. Second, the country is the unit of analysis. The national samples differ greatly in size — Italy alone contributes some seven in ten responses — and any figure pooled across countries would in substance be a figure about Italy. Pooled results are therefore reported only as descriptive context, are never weighted, and are never presented as a consortium-wide average; every comparative claim in this chapter is made at the level of national distributions set side by side. Third, and following from the same principle, the smallest samples are read as indicative rather than representative: the Slovenian ($n = 15$) and especially the Croatian ($n = 6$) results are treated as national soundings whose value lies in the direction and pattern they suggest, and percentages derived from them are not given an interpretive weight their size cannot carry. Fourth, one respondent who completed the Italian version but declared Belgium as their country is reassigned to the Belgian sample, since the meaning of the harmonised responses does not depend on the language of completion; the counts in Table 15 reflect this reassignment. Item non-response was negligible — fewer than one per cent of closed-item cells across the dataset — and is handled by computing every statistic on the valid responses to that item. The two open questions were analysed thematically, and their coded content is used in Sections 5.3 and 5.4 to interpret the closed results.

5.1.5 Limitations

The evidential status of the survey is bounded by its design, and four limitations are recorded so that the findings are not asked to bear more than they can. First, the sample is self-selecting: respondents were reached through the partners' professional networks and chose to participate, so the sample is a convenience sample of teachers within the partnership's reach rather than a probability sample of the five teaching populations, and teachers already interested in science pedagogy or in technology may be overrepresented. Second, the national samples are unequal, and the reading rules stated above are the consequence: the survey supports comparison of national patterns, not precise estimation of national parameters. Third, no inferential statistics are computed, and this is a matter of design rather than omission: the survey's function in the method is confirmatory and localising, the sampling frame would not support population inference, and significance testing on samples of this structure would manufacture a precision the data do not possess. Fourth, all measures are self-reported perceptions gathered at a single point at the close of the school year; they record how teachers see their curricula, constraints and readiness, which is precisely the



information a training intervention requires, but they are not observational measures of classroom practice. None of these limitations compromises the survey's purpose. Confirmation, localisation and the testing of the immersive premise all concern the perceptions and conditions of the teachers the project will train, and it is exactly these that the instrument measures.

5.2 Respondent Profile

The survey drew 292 valid responses across the five national versions. Their distribution by country is shown in Figure 5, and the composition of each national sample along the profile dimensions is summarised in Table 16. The two must be read together, because the distribution is markedly uneven: Italy contributes 202 of the 292 responses, some seven in ten, while Croatia contributes six. This imbalance is the reason for the reading rules set out in Section 5.1, and it governs the interpretation throughout: the country is the unit of analysis, and the Slovenian and Croatian returns in particular are read as national soundings rather than as representative samples of their teaching populations.

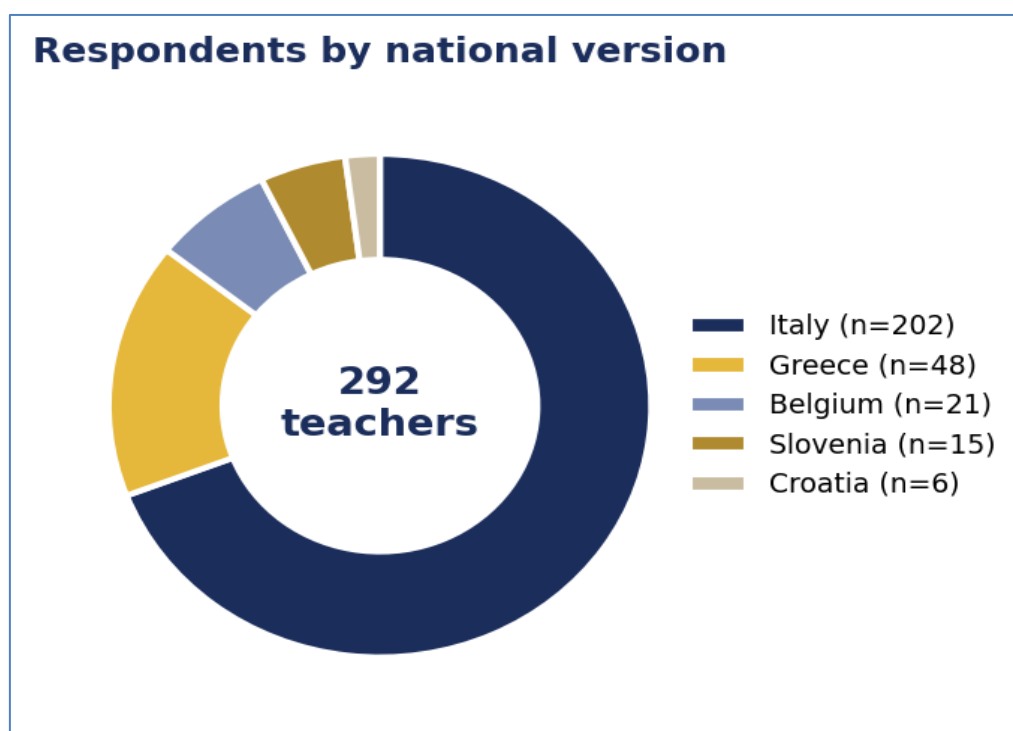


Figure 5 Distribution of the 292 valid responses across the five national versions.

The respondents are, across all five systems, an experienced body of teachers. A majority in four systems, and half of the Croatian respondents, report more than a decade of science teaching, rising to some seven in ten in Greece and Belgium, and the least experienced band — three years or fewer — is small everywhere. The sample therefore reflects the judgement

of established practitioners rather than of teachers new to the profession, which strengthens the standing of the perceptions it records: respondents are reporting on curricula they have taught for many years and on constraints they have long worked within. The subjects taught span the sciences the framework addresses — physics, chemistry, biology and the natural sciences — together with mathematics and, in the Flemish and Croatian returns, integrated STEM and technology, consistent with the way science is organised in each system.

The samples differ in composition in ways that matter for the national reading. The level taught varies with national school structure: the Italian and Greek samples are weighted towards upper-secondary teaching, whereas the Belgian, Slovenian and Croatian samples include a substantial share of lower-secondary teachers, reflecting both the structure of these systems and the networks through which the survey was distributed. Gender composition also varies, from a male-majority Greek sample to strongly female-majority Slovenian and Croatian ones; the Italian sample, the largest, is close to seven-tenths female. School type is predominantly public across the consortium, with the notable exception of the Flemish sample, in which subsidised free schools — the characteristic form of Belgian secondary provision — predominate. The urban–rural balance is broadly urban, most completely so in the small Croatian sample, whose six respondents all teach in urban schools. These compositional differences are not weaknesses to be corrected but features of the national contexts the survey set out to capture, and they are carried forward into the national findings of Section 5.3, where each sample's profile qualifies the reading of its results.

Table 16 Composition of each national sample along the principal profile dimensions.

National system	Valid n	Female	Upper secondary	11+ years' experience	Urban schools
Italy	202	69%	65%	58%	76%
Greece	48	38%	65%	71%	79%
Belgium — Flemish Community	21	38%	33%	71%	38%
Slovenia	15	80%	40%	53%	60%
Croatia	6	83%	33%	50%	100%

5.3 National Findings

This section reads the survey country by country. Each national system is given a portrait that draws together its returns across all sections of the instrument — the emphasis its curriculum places on the competencies and the depth to which they are actually taught, the constraints under which its teachers work, the difficulty they report in engaging students as scientists, the sources of that difficulty in the classroom, the benefit they anticipate from immersive

activities, the conditions they attach to using them, and their own readiness — and closes with the content their open responses identify as where immersion would help most. The purpose is localisation: to surface the national particularities that the mixed-country groups of the consortium workshop deliberately set aside, and that the national-adaptation guidance of Chapter 6 must accommodate. The portraits are read subject to the reading rules of Section 5.1. The Italian and Greek returns rest on substantial samples; the Belgian sample is modest; and the Slovenian and especially the Croatian portraits are drawn from small samples and are presented as indicative national soundings rather than as representative measurements. All ratings are reported on the common 1–5 metric described in Section 5.1, and the shorthand "top-two-box" denotes the proportion of respondents choosing one of the two highest categories on an item. Constraints are the most nationally distinctive part of the evidence, and Figure 6 sets them side by side across the five systems; each portrait refers to it.

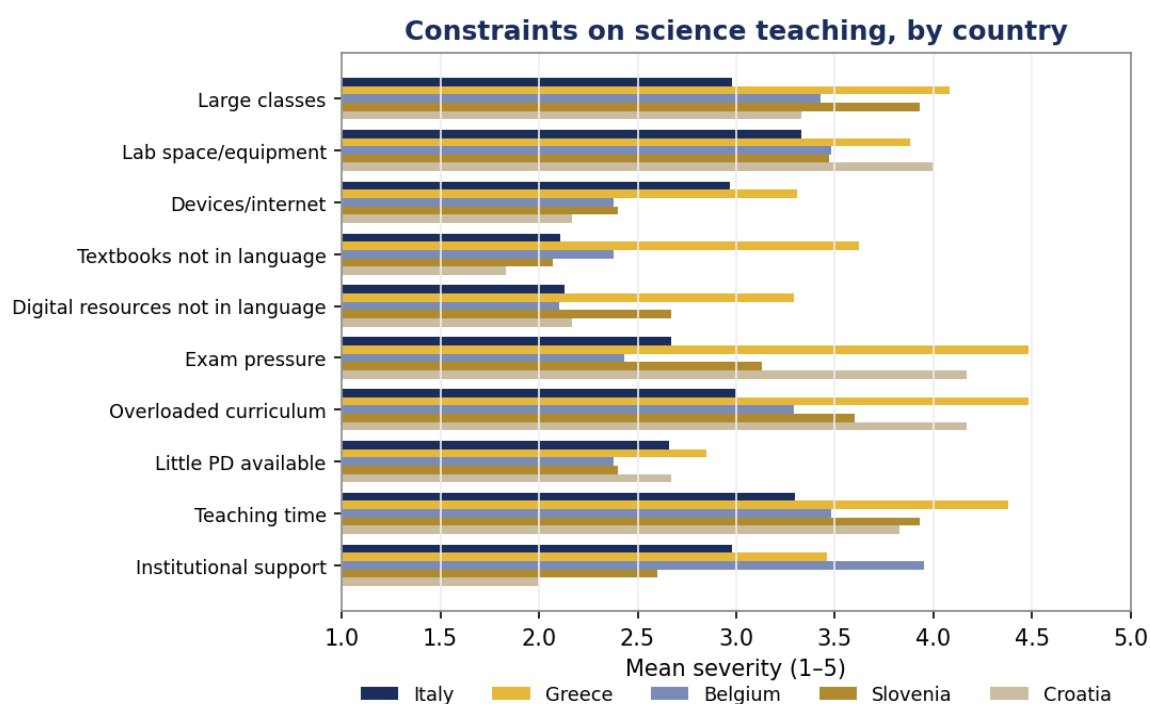


Figure 6 Reported severity of constraints on science teaching, by national system.

Italy

The Italian sample is by far the largest, at 202 valid responses, and the most experienced: some three in five respondents report more than a decade of science teaching. It is weighted towards upper-secondary teaching, is close to seven-tenths female, and is drawn entirely from public schools, predominantly in urban settings. Its size means that its returns are the

most stable of the consortium and the least sensitive to individual responses, and it is the only sample large enough to read distributions rather than merely central tendencies.

The relationship between curricular emphasis and depth of teaching is, in the Italian returns, the reverse of the shortfall the intervention might be expected to reveal. Across all eight competency areas, respondents report teaching each to slightly greater depth than they judge the national curriculum to emphasise it: making sense of data, the highest-rated area on both measures, rises from an emphasis mean of 3.36 to a depth mean of 3.59, and the same modest upward movement holds for prediction, argument and the environmental area. Experimental design is the lowest-rated competency on both measures and the one the fewest teachers report teaching in depth, a pattern consistent with the practical constraints reported below. The reading this supports is that Italian teachers do not experience a large gap between curricular ambition and their own practice in these competencies; the case for training therefore rests not on a perceived shortfall in what is taught but on the difficulty of teaching it and on readiness to use new means, both examined below.

The Italian constraint profile, shown in Figure 6, is the most moderate in the consortium: no constraint reaches a mean severity of 3.4. The two felt most strongly are the availability of laboratory space and equipment (3.33) and of teaching time (3.30) — the practical limits on hands-on science rather than the systemic pressures that dominate elsewhere. Examination pressure (2.67) and the availability of professional development (2.66) are felt only moderately, and the unavailability of up-to-date materials and digital resources in Italian is the least of the sample's concerns (2.11 and 2.13, each rated "not at all" a constraint by more than four in ten respondents), unsurprisingly given the scale of the Italian-language educational-resource base. The difficulties teachers report in engaging students as scientists are moderate and evenly distributed, most pronounced in helping students form and test hypotheses (3.30) and build and use models (3.29), and least in supporting everyday evidence-based decisions (2.98); on most of these items a majority of respondents cluster on the mid-point, indicating difficulty that is real but not acute. Among the classroom sources of that difficulty, the inaccessibility of real places and situations (3.43), invisible mechanisms (3.36) and experiments too dangerous or costly to run (3.32) are felt most strongly — precisely the difficulties the immersive affordances are designed to address.

Against this, the anticipated benefit of immersive activities is high and consistent across all five affordances, ranging from 3.74 for conducting otherwise dangerous experiments safely to 4.00 for visiting inaccessible real places, the latter rated in the top two categories by three-quarters of respondents. The conditions teachers attach to use are exacting and clear: compatibility with class size and available devices (4.16) and with lesson time (4.05) are the

most strongly required, and availability in Italian, though the least pressing constraint on current teaching, is nonetheless a firm condition of use (3.93). Readiness is the evident weak point. The two most technical readiness items — knowledge of immersive tools (2.81) and technical skill in using them (2.82) — are the only readiness items below the scale mid-point, while the pedagogical readiness items, classroom management (3.16) and assessment (3.14), sit above it. The pattern is telling: Italian teachers judge themselves pedagogically equipped but technically unprepared. This is confirmed by prior experience, where 89 of the 202 respondents have never used immersive tools and a further 50 have tried them only once or twice, so that some seven in ten stand at the beginning of the learning curve the training must climb.

The Italian open responses are the richest of any system and locate the perceived value with unusual precision. The content teachers most frequently name is the microscopic and the abstract: atomic and molecular structure, electron configuration, the chemistry of matter, cell and molecular biology, DNA and genetics, together with the phenomena of physics — thermodynamics, electricity and magnetism, energy conservation, free fall — that students cannot directly observe. Several respondents extend the case to the inaccessible in space and time, proposing immersive visits to the interior of the human body, the ocean, the interior of rocks, the universe, and the long timescales of fossil formation, mineral growth, evolutionary change and the growth of populations. The open material also carries a consistent set of qualifications that the design work should heed: that immersive activity should complement rather than replace practical and manual work, whose development is already under pressure from limited time; that its spectacle must not displace its scientific content, which places the teacher's framing role at the centre; and, raised repeatedly, that classes are too large for such work, with more than one respondent judging a maximum of ten to fifteen students necessary for it to be effective.

Greece

The Greek sample comprises 48 responses, is experienced — seven in ten report more than a decade of teaching — and is weighted towards upper-secondary and the physical sciences, with physics, chemistry and biology the subjects most represented. It is the one male-majority sample in the consortium and is drawn entirely from public schools. Its defining feature is the intensity of the systemic constraints it reports, and in this it stands apart from every other system.

The emphasis–depth relationship in Greece resembles the Italian pattern but is more pronounced: across every competency area, depth of teaching is rated above curricular emphasis. Emphasis ratings are the lowest of any system, with the modelling and experiment-



design areas rated below the mid-point (2.62 and 2.58), while depth ratings recover to the mid-3 range. The reading is not that Greek teachers teach these competencies especially deeply — they do not, in absolute terms — but that they perceive the national curriculum as placing comparatively little explicit emphasis on them, a perception that must be read together with the constraint profile that follows.

That profile is the survey's single most striking national result. Examination pressure (4.48), curriculum overload (4.48) and shortage of teaching time (4.38) all reach mean severities approaching the top of the scale, each rated in the top two categories by around nine in ten respondents, and large class sizes (4.08) follow closely. Greece is also the only system in which the unavailability of up-to-date materials and digital resources in the national language registers as a substantial constraint (3.62 and 3.29), where in every other country these items are among the least pressing. This national signature — a demanding examination system compressing an overloaded curriculum into insufficient time, worked by teachers with a comparatively thin base of localised digital resource — is the most important localisation finding of the survey, and it bears directly on the national-adaptation and language provisions of Chapter 6. Consistent with this pressure, Greek teachers report the greatest difficulty of any system in engaging students as scientists: forming and testing hypotheses is rated 4.06, difficult or very difficult for eight in ten respondents, and every practice in the section is rated above 3.4. The classroom sources they emphasise are invisible mechanisms and inaccessible real places (each 3.67).

The anticipated benefit of immersive activities is correspondingly the highest among the larger samples, every affordance rated above 4.1 and above 4.2 for visiting real places and manipulating variables live. The conditions attached to use are shaped by the constraint profile: alignment with the curriculum and examinations (4.21) ranks far higher here than in any other system, and availability in Greek (4.06) is a firm requirement — both direct consequences of the pressures the sample reports. Self-assessed readiness is the highest in the consortium, though still only moderate, and is concentrated in classroom management (3.77) rather than in the technical items. The open responses reinforce the picture: teachers name microscopic biological processes, chemical reactions, the nervous system and brain, quantum physics, and the depiction of force and motion in settings such as space as the content immersion would serve, and more than one connects the case explicitly to the difficulty of the abstract micro-world. Running through the open material is a recurring concern, stated most fully by one respondent, that the examination-driven, mechanistic solving of problems bearing little relation to reality leaves no room for the genuine understanding of causes and consequences that immersive approaches would support — a

concern that frames the Greek appetite for the method as much as any feature of the closed responses does.

Belgium — Flemish Community

The Flemish sample comprises 21 responses and differs from the others in composition as well as in its returns. It is weighted towards lower-secondary teaching, drawn substantially from the subsidised free schools characteristic of Belgian provision rather than from the public sector that dominates elsewhere, less urban than the other samples, and includes integrated STEM and technology alongside the individual sciences. It is experienced, with seven in ten reporting more than a decade of teaching.

The emphasis–depth relationship in the Flemish returns is mixed rather than uniformly inverted. Making sense of data is rated highly on both measures, and the depth of modelling is reported well above its emphasis, but for several competencies — prediction most clearly — teachers report teaching to less depth than the minimum attainment targets emphasise, the one clear instance in the survey of the expected shortfall. The constraint profile is the most distinctive in the consortium and points to organisational rather than curricular or resource pressures. Institutional and administrative support (3.95) is the single most strongly felt constraint on any item in any system, while the pressures that dominate Greece — examination pressure (2.43) and language availability (2.38 and 2.10) — are among the lightest here. Class size (3.43), teaching time and laboratory provision (each 3.48) occupy the middle of the profile. Difficulty in engaging students is moderate, most pronounced in using science to address real problems (3.71) and explaining phenomena scientifically (3.57), and notably low in reading and interpreting data (2.86). The classroom sources of difficulty emphasised are processes too slow or fast to observe (3.29) and invisible mechanisms (3.19).

Anticipated benefit from immersive activities is high, led by visiting inaccessible real places (4.14), and the conditions attached to use are among the most demanding in the survey on the practical items: readiness for use without heavy preparation (4.14, top-two-box 95%), fit with lesson time and with class size and devices (each 4.14 or above), and alignment with the curriculum (4.24) are all strongly required, whereas the inclusion of teacher training (3.19) is required less strongly than anywhere else — consistent with a sample that already reports comparatively strong readiness. Readiness is indeed among the stronger in the consortium, led by classroom management (3.62) and integration into lesson design (3.29), though knowledge of tools (2.86) remains the weakest item as it does everywhere. Prior experience is the most favourable of any system: eight of the 21 respondents use immersive tools at least occasionally, though a majority still stand at or near the start. The open responses, though few, are specific — the particle or atomic model, the internal structure of the body and organs,

volcanic eruptions, cellular chemistry including respiration and photosynthesis, and heat balance in physics — and one respondent frames the affordance in a way none other does, valuing the capacity for a whole class to observe the same thing at the same moment and to see calculation and observation side by side.

Slovenia

The Slovenian sample comprises 15 responses and is read as indicative. It is strongly female, split between lower- and upper-secondary teaching, drawn entirely from public schools, and spans chemistry, biology and physics together with a computing-and-informatics respondent. Its modest size means the portrait rests on patterns rather than precise values, and individual responses carry appreciable weight.

As in Italy and Greece, depth of teaching is reported at or above curricular emphasis across the competencies, with making sense of data again the highest-rated on both measures. The constraint profile resembles the moderate Italian pattern more than the acute Greek one: large class sizes and shortage of teaching time (each 3.93) are the most strongly felt pressures, followed by curriculum overload (3.60), while examination pressure (3.13) is felt more mildly and the language-availability items are among the least pressing anywhere, textbooks in Slovene registering essentially no top-two-box responses. Slovenian teachers report the lowest frequency of any system for the classroom sources of difficulty captured by the instrument — no source exceeds 3.20 — suggesting a teaching context in which the specific obstacles immersive activities address are encountered somewhat less often than elsewhere. Difficulty in engaging students is nonetheless moderate and is highest for building and using models (3.60).

Even against this lighter profile of reported difficulty, the anticipated benefit of immersive activities is high, four of the five affordances rated 4.07 and each rated in the top two categories by four in five respondents — an appetite that does not depend on the difficulties being acute. Among the conditions attached to use, readiness for use without heavy preparation (4.13) and, distinctively, the inclusion of teacher training (4.00) rank highest, the latter more strongly required than in the more confident Flemish sample. Readiness is moderate, led by classroom management (3.53), with the technical and assessment items clustered just below the mid-point. What most distinguishes the Slovenian returns is the open material, which is the most technically specific of any system and reaches beyond the science strands into the project's supplementary concerns. Alongside fields and field densities, electromagnetic waves, forces and energy, the particle world and the sub-microscopic level of chemistry, respondents named the interpretation of machine-learning models — and specifically the difficulty of explaining opaque, "black-box" models — as content where

immersive or visual approaches would help. One respondent observed that everything to do with fields, and with wave phenomena that propagate in all directions, seems especially well suited to augmented reality. This is a small but pointed signal that the abstract content teachers most want to make visible extends to the computational, and it is consistent with the treatment of the AI strand elsewhere in this deliverable.

Croatia

The Croatian return comprises six responses and is reported as a national sounding rather than as a measurement. No percentage drawn from it should be given interpretive weight; its value lies in the direction it indicates and in the consistency of that direction with the wider evidence. The six respondents are predominantly female, teach across secondary and the upper years of primary, span mathematics, biology, chemistry and — in one case — artificial intelligence, and all teach in urban schools.

Within these limits the pattern is clear. Depth of teaching is reported at or above emphasis, as in the other systems, with making sense of data the highest-rated area. The constraints reported most strongly — examination pressure and curriculum overload (each 4.17) and laboratory provision (4.00) — echo the Greek profile more than the moderate Italian one, while the language-availability and institutional-support items sit at the bottom of the profile. Difficulty in engaging students is moderate, highest for forming and testing hypotheses (3.50), and the classroom sources emphasised are dangerous or costly experiments and inaccessible real places (each 3.83). The most notable feature of the sounding is the relationship between appetite and preparation. The anticipated benefit of immersive activities is the highest recorded anywhere in the survey — every one of the five affordances rated in the top two categories by every respondent, with means from 4.33 to 4.67 — while readiness is the lowest of any system, no item reaching the mid-point and knowledge of tools rated 2.50. The gap between the value teachers see in the method and their preparedness to use it is therefore wider here than anywhere else, which sharpens rather than weakens the case for training. The conditions attached to use are demanding in proportion to that appetite, with fit to class size and devices rated 4.83. The single substantive open response is the most developed piece of open material in the survey and is explicitly mathematics-anchored: it identifies three-dimensional solids, spatial sections and the graphs of functions, including trigonometric functions, as content students find very hard to visualise and even to sketch, and argues that immersive activity in mathematics would require classes of no more than ten with a workstation for each student, against the twenty-four to twenty-eight the respondent currently teaches, if students are to be active participants rather than observers. It is at once a confirmation of the immersive approach beyond the science strands and a precise



statement of the enabling conditions the approach demands. The Croatian sounding should be revisited as the pilot proceeds and a larger national response becomes available.

5.4 Cross-National Analysis

Where the previous section read each system on its own terms, this section reads the five together, setting the national returns side by side to establish what holds across the consortium and where the systems diverge. It is organised around the three questions the survey was designed to answer, stated in Section 5.1. The first concerns the training premise: whether the perceived relationship between what curricula ask and what teachers deliver supports the case for the intervention. The second concerns the medium and is the analytical centre of the chapter: whether the difficulties teachers actually report in the classroom correspond to the pedagogical affordances through which the project treats immersive technology. The third concerns implementation: what the conditions teachers attach to use, and their own readiness, imply for the design of the training and for its national adaptation. Throughout, the cross-national reading is conducted under the rules of Section 5.1 — national distributions are compared, no consortium-wide average is formed, and the small Slovenian and Croatian samples are read for direction rather than magnitude.

5.4.1 The training premise: emphasis, depth and difficulty

The instrument approached the training premise first through the paired measures of curricular emphasis and depth of teaching. The expectation that these might reveal a shortfall — competencies emphasised in the curriculum but not delivered in depth — is not borne out by the returns, and this is reported as a finding rather than concealed. In four of the five systems the relationship runs the other way: teachers report teaching the eight competencies to slightly greater depth, on average, than they judge their curriculum to emphasise them, the difference ranging from around a fifth of a scale point in Italy to around a third in Greece and Slovenia. Only the Flemish sample shows the expected direction, and there only weakly and for some competencies rather than all. The competency for which depth most consistently lags is the designing of experiments, the lowest-rated area on both measures in most systems. The reading this supports is clear and should be stated plainly: across the consortium, teachers do not perceive a large gap between what their curricula ask of these competencies and what they themselves deliver, and the case for the intervention does not rest on such a gap.

It rests instead on difficulty. The measures of how hard teachers find it to engage students as scientists (Q4) and of the classroom sources of that difficulty (Q5) tell a consistent cross-



national story. In every system the scientific practices are reported as moderately to highly difficult to develop in students, most acutely in Greece, where forming and testing hypotheses reaches 4.06, and least so in Italy, but nowhere trivially; and in every system the difficulty is concentrated in the same practices — forming and testing hypotheses and building and using models — the practices that depend most on making unobservable structure and process available to students. The classroom sources teachers emphasise are likewise consistent: invisible mechanisms, inaccessible real places and situations, and experiments too dangerous or costly to run recur across the systems as the difficulties felt most strongly. It is this shared structure of difficulty, not a perceived curricular shortfall, that establishes the need the training addresses, and it leads directly to the question of the medium.

5.4.2 The medium: difficulty and affordance

The central design of the instrument pairs each classroom source of difficulty (Q5) with the immersive affordance intended to answer it (Q6): phenomena too small, large, slow or fast to observe and invisible mechanisms with the affordance of seeing otherwise invisible processes; experiments too dangerous or costly with conducting them safely; the difficulty of grasping abstract structure with entering a model from within; inaccessible real places with visiting them; and static representation with manipulating variables and observing effects in real time. The correspondence between the two lists is exact by construction, which allows the survey to test the project's central premise about the medium directly. Figure 7 sets the pooled reported difficulty against the pooled anticipated benefit for each paired affordance.

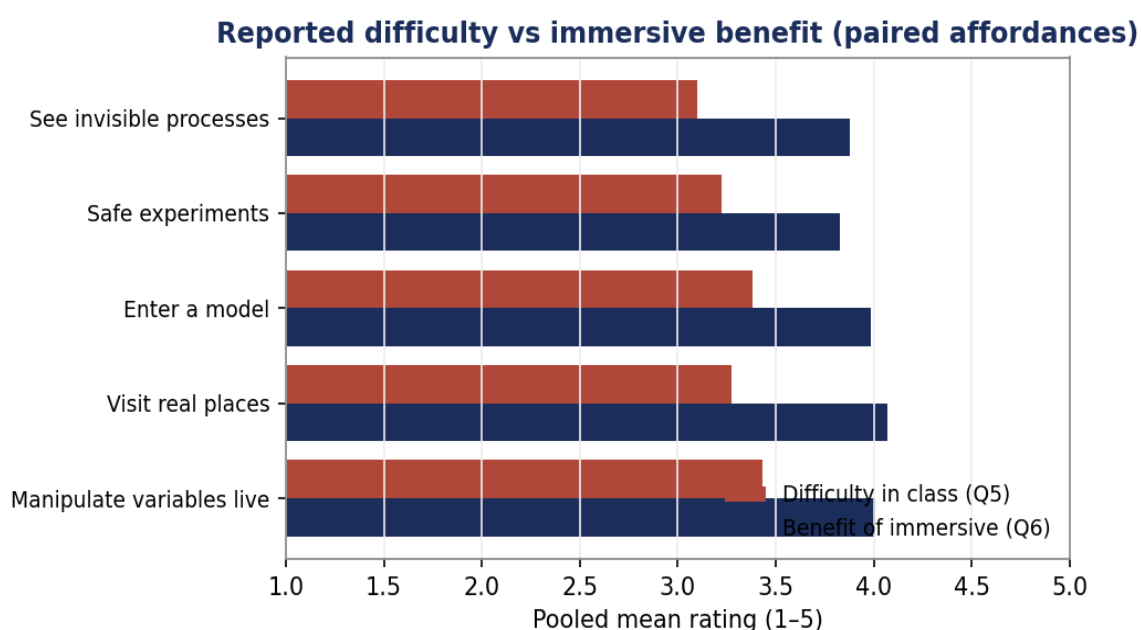


Figure 7 Reported difficulty in the classroom (Q5) against anticipated benefit of the corresponding immersive affordance (Q6).

The result is the clearest confirmation the survey provides. For every one of the five paired affordances the anticipated benefit stands above the reported difficulty: teachers see the immersive affordance as more valuable than the corresponding difficulty is frequent. The two affordances rated most beneficial — visiting inaccessible real places and manipulating variables in real time, each pooled above 4.0 — are precisely those answering the difficulties teachers report most strongly, and the affordance of entering a model, which answers the difficulty of abstract structure that Q4 identifies as among the hardest to teach, is rated close behind. The relationship holds not only in the pooled figures but within every national sample: the overall benefit ratings range from 3.84 in Belgium to 4.53 in Croatia, and no country rates any single affordance below 3.57. The premise on which the project treats immersive technology — that its value lies in specific pedagogical affordances matched to specific teaching difficulties, rather than in the technology as such — is therefore supported by the teachers who would use it, and supported in the same direction across five different national contexts.

This cross-national consistency is shown in Figure 8, which reports the perceived benefit of each affordance by country.

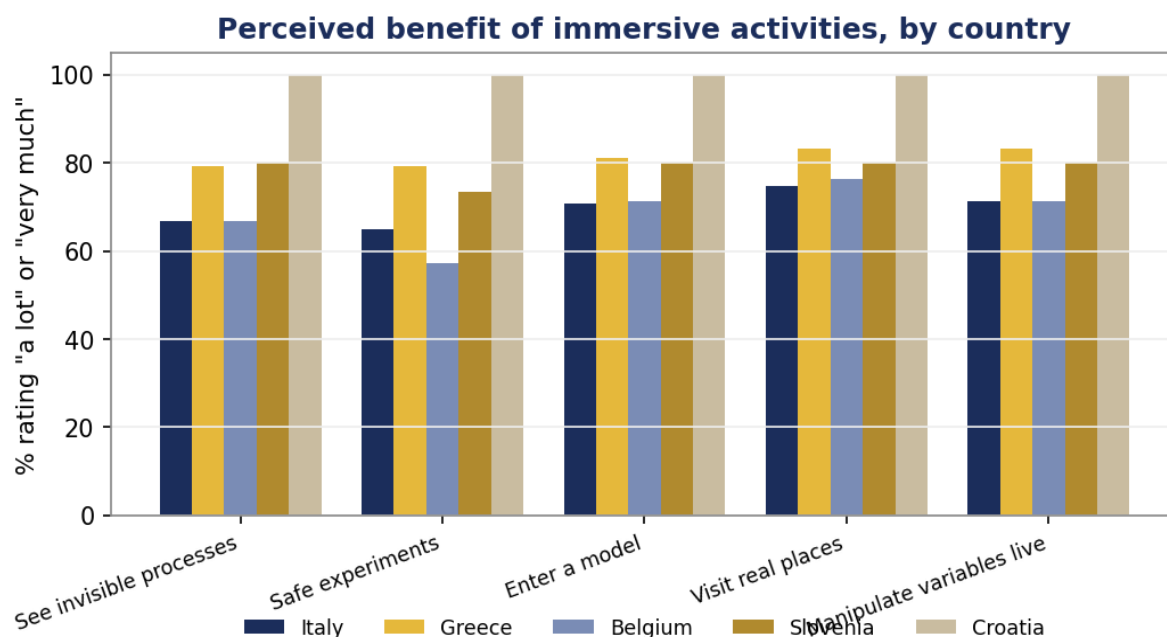


Figure 8 Perceived benefit of immersive activities against the five affordances, by national system.

The striking feature is the agreement: despite the wide differences in the systems' constraints, curricula and readiness documented elsewhere in this chapter, the shape of the benefit profile is common to all five, with visiting real places and manipulating variables live the most valued affordances in most systems and no affordance falling out of favour in any. The

differences that do appear are of level rather than of pattern — the Greek and Croatian samples rate every affordance more highly than the others — and are consistent with the more acute difficulties those systems report. The affordance-led treatment of immersive technology is thus not a construction that holds only in the aggregate; it is recognised, in the same form, by teachers working in materially different conditions.

5.4.3 Implementation: conditions and readiness

The survey's third purpose was to establish the conditions under which the framework will be used, and here the cross-national reading converges on a set of requirements that the design of the training and its national adaptation must satisfy. The conditions teachers attach to using immersive activities (Q7) are, in their ranking, remarkably consistent across the consortium. Compatibility with class size and available devices is the most strongly required condition overall and the first or second requirement in every system; fit with the available lesson time and readiness for use without heavy preparation follow closely. These are practical, operational requirements rather than aspirational ones, and they align exactly with the constraints teachers report most strongly — class size, laboratory provision and time — so that the conditions of use are, in effect, the constraints restated as design requirements. Two conditions vary in a way that carries directly into national adaptation: alignment with the curriculum and examinations is required far more strongly in the examination-pressured systems, Greece and Croatia, than elsewhere, and availability in the national language, though a firm requirement everywhere, is most pressing in Greece, the one system that also reports language availability as a substantial present constraint.

If the conditions of use are consistent, so is the readiness gap, and it is the most important implementation finding of the survey. Figure 9 reports self-assessed readiness by country across the five dimensions. In every system the profile has the same shape: the pedagogical dimensions — classroom management and, to a lesser degree, assessment — sit highest, at or above the scale mid-point, while the two technical dimensions, knowledge of immersive tools and technical skill in using them, sit lowest, below or at the mid-point in every country. Knowledge of tools is the single weakest readiness item in the consortium, rated below the mid-point in four of the five systems and lowest of all in Croatia.

The consistent interpretation is that teachers across the consortium judge themselves pedagogically prepared to teach with immersive activities but technically unprepared to operate them — a diagnosis with a direct consequence for the training, which must build technical fluency from a low base while assuming, and drawing on, existing pedagogical competence rather than teaching it afresh.



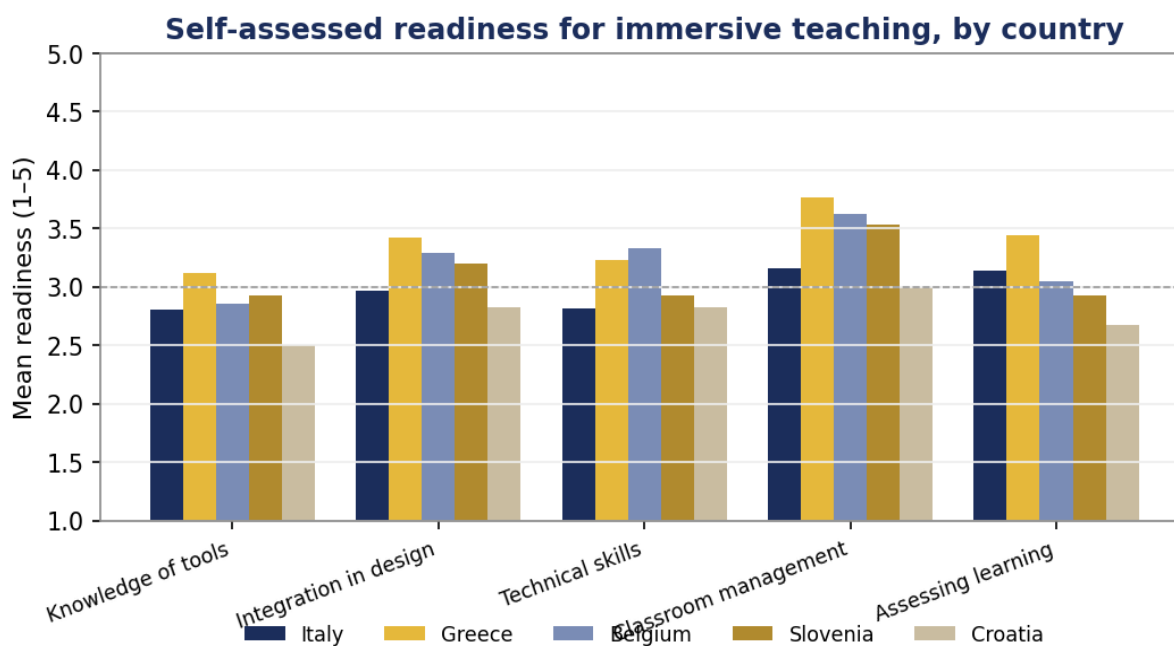


Figure 9 Self-assessed readiness for immersive teaching across five dimensions, by national system.

Prior experience, shown in Figure 10, confirms the readiness gap from a different direction and sets the baseline from which the training begins.

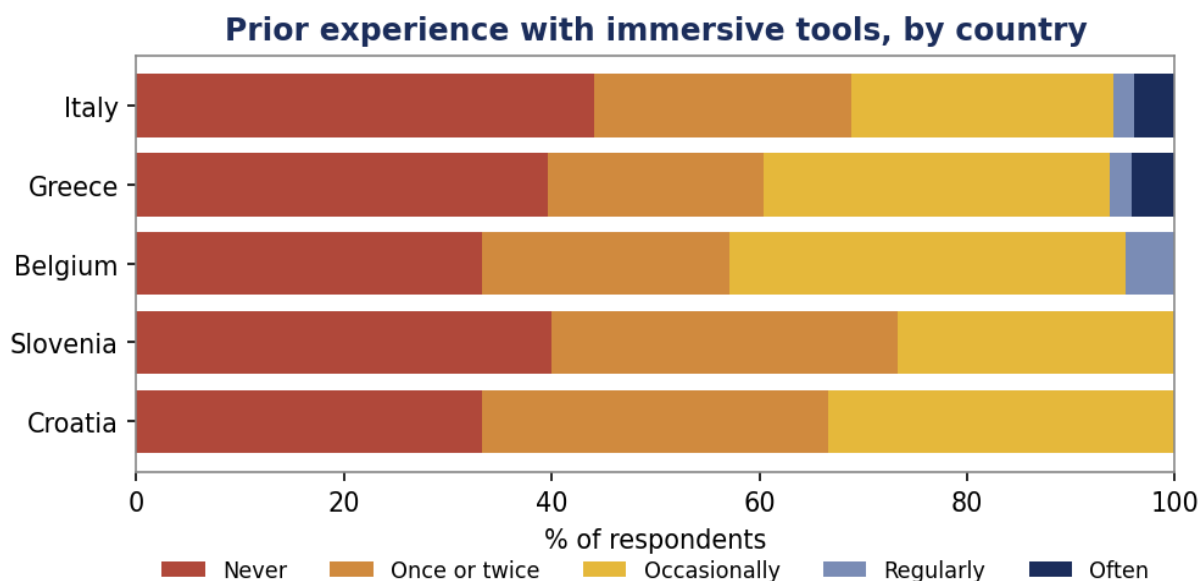


Figure 10 Prior experience with immersive tools, by national system.

In every system a majority of respondents have either never used immersive tools or tried them only once or twice — from 57% in Belgium, the most experienced sample, to 73% in Slovenia. Regular or frequent use is 6% in Italy, 6% in Greece, 5% in Belgium and absent in the Slovenian and Croatian returns. The training therefore addresses a population that is, with few exceptions, at the beginning of its engagement with the medium, which reinforces the

case for the exacting usability conditions teachers attach to use: activities that demand heavy preparation or specialist technical skill will not meet a workforce that reports little of either. Read together, the three implementation findings — consistent practical conditions of use, a consistent technical readiness gap, and a low base of prior experience — describe with some precision the design envelope within which the training and its scenarios must be built, and they are taken up in that form in the implementation guidance of Chapter 6.

5.5 Conclusions

This section draws the survey together as a set of numbered findings. It reports first the findings specific to each national system, then the findings that hold across the consortium, and states in each case what the finding establishes; it closes with the implications the survey carries for the project as a whole. The findings are numbered so that later work — the national-adaptation guidance of Chapter 6, the scenario design of the subsequent work packages, and the partner review of this deliverable — can refer to them directly. All findings rest on the evidence reported in Sections 5.2 to 5.4 and are subject to the reading rules of Section 5.1: the national findings for Italy and Greece rest on substantial samples, that for Belgium on a modest one, and those for Slovenia and Croatia on small samples read for direction rather than magnitude.

National findings

N1 · Italy: A large, experienced sample reporting moderate constraints, a recognised difficulty in teaching the target practices, and a primarily technical readiness gap.

The Italian sample of 202 responses is the largest and most experienced in the consortium, some three in five respondents reporting more than a decade of teaching. Its constraint profile is the most moderate of the five systems: no constraint reaches a mean severity of 3.4, the strongest being laboratory provision (3.33) and teaching time (3.30), while the unavailability of materials in Italian is the least of its concerns (2.11–2.13). Difficulty in engaging students as scientists is moderate and concentrated in hypothesis-testing (3.30) and model-building (3.29); anticipated benefit from immersive activities is high across all affordances (3.74–4.00). Readiness divides sharply along a single line: the pedagogical items sit above the scale mid-point while the technical items — knowledge of tools (2.81) and technical skill (2.82) — sit below it, and 139 of the 202 respondents have never used immersive tools or used them only once or twice. This establishes that in the consortium's largest system the obstacle to immersive teaching is technical capability and experience rather than pedagogical readiness or curricular resistance, and that the training must build technical fluency from a low base. The Italian open responses further establish, in the teachers' own

terms, both the content immersion should target — atomic and molecular structure, cell biology, and the unobservable phenomena of physics — and two firm design constraints: that immersive work complement rather than replace practical work, and that class sizes well below current norms are needed for it to be effective.

N2 · Greece: Acute systemic constraints, the highest reported teaching difficulty, and a distinctive dependence on curriculum alignment and national-language provision.

The Greek sample of 48 responses is defined by the intensity of its systemic constraints. Examination pressure (4.48), curriculum overload (4.48) and shortage of teaching time (4.38) all approach the top of the scale, each felt strongly by around nine in ten respondents, and Greece is the only system in which the unavailability of up-to-date materials and digital resources in the national language registers as a substantial constraint (3.62 and 3.29). Consistent with this pressure, Greek teachers report the greatest difficulty of any system in engaging students as scientists, hypothesis-testing reaching 4.06, and the highest anticipated benefit from immersive activities among the larger samples (every affordance above 4.1). Two conditions of use are raised more strongly here than anywhere else: alignment with the curriculum and examinations (4.21) and availability in Greek (4.06). This establishes the survey's principal localisation requirement — that for the Greek system, immersive provision must be curriculum- and examination-aligned and available in Greek to be usable at all — and it identifies examination-driven curriculum compression, rather than teacher disposition, as the binding constraint the national adaptation must work within. The open responses establish the same point from the teachers' side, connecting the appetite for immersion explicitly to the difficulty of teaching the abstract micro-world under examination pressure.

N3 · Belgium (Flemish Community): An organisational rather than curricular constraint profile, the only clear emphasis–depth shortfall, and the strongest prior experience.

The Flemish sample of 21 responses, drawn substantially from subsidised free schools and weighted towards lower-secondary teaching, presents a constraint profile unlike any other. Institutional and administrative support (3.95) is the single most strongly felt constraint on any item in any system, whereas examination pressure (2.43) and language availability (2.10–2.38) are among the lightest. It is also the only system in which the expected emphasis–depth shortfall appears, teachers reporting that they teach some competencies — prediction most clearly — to less depth than the attainment targets emphasise. Prior experience is the most favourable in the consortium, nine of the 21 using immersive tools at least occasionally, and readiness is correspondingly among the stronger, so that the inclusion of teacher training is required less strongly here (3.19) than anywhere else. This establishes that the barriers to immersive teaching are not uniform across the consortium: in the Flemish system they are organisational rather than curricular or resource-based, and the training and support model

appropriate here differs from that required in the examination-pressured systems, presuming a higher base of readiness and directing effort towards institutional enablement rather than initial technical training.

N4 · Slovenia: An indicative sounding of moderate constraints and lower reported difficulty, yet strong demand for immersion and an unprompted signal towards the AI strand.

The Slovenian sample of 15 responses is read as indicative. Its constraints resemble the moderate Italian pattern, led by class size and teaching time (each 3.93), and it reports the lowest frequency of any system for the classroom difficulties the instrument captures, no source exceeding 3.20. Even so, anticipated benefit from immersive activities is high, four of the five affordances rated 4.07, which establishes that demand for the medium does not depend on the underlying difficulties being acute — teachers value the affordances on their own terms. Distinctively, the inclusion of teacher training is required more strongly here (4.00) than in the more experienced Flemish sample. The most consequential Slovenian result is in the open material: respondents named the interpretation of machine-learning models, and specifically opaque "black-box" models, alongside fields and electromagnetic phenomena, as content where immersive or visual approaches would help. This establishes independent, teacher-originated support for extending the immersive approach into the artificial-intelligence strand that the deliverable currently treats as a supplement, and it is noted for the work that will elaborate that strand.

N5 · Croatia: A small sounding showing the widest gap in the survey between appetite for immersion and readiness to use it, with an unprompted signal towards the mathematics strand.

The Croatian return of six responses is reported as a national sounding, its value residing in direction rather than magnitude. Its constraints echo the Greek profile — examination pressure and curriculum overload (each 4.17) and laboratory provision (4.00) — and it records the highest anticipated benefit from immersive activities anywhere in the survey, every affordance rated in the top two categories by every respondent (means 4.33–4.67), against the lowest readiness of any system, no item reaching the mid-point and knowledge of tools rated 2.50. This establishes, even allowing for the small sample, the widest gap in the consortium between the value teachers place on the medium and their preparedness to use it — a gap that sharpens rather than weakens the case for training. The single developed open response establishes a further point: it identifies three-dimensional geometry, spatial sections and the graphs of functions as content students struggle to visualise, and specifies the enabling conditions immersive mathematics would demand, namely small classes and a workstation per student. This is teacher-originated support for the relevance of the immersive

approach in the mathematics strand, which the deliverable includes provisionally pending validation.

Cross-national findings

C1 The case for the intervention rests on difficulty, not on a perceived gap between curricular emphasis and depth of teaching.

In four of the five systems, teachers report teaching the eight target competencies to slightly greater depth, on average, than they judge their curriculum to emphasise them, the difference ranging from around a fifth of a scale point in Italy to around a third in Greece and Slovenia; only the Flemish sample shows the expected shortfall, and there weakly. The anticipated deficit between curricular ambition and delivered depth therefore does not appear. What does appear, consistently, is difficulty: in every system the scientific practices are reported as moderately to highly difficult to develop in students, concentrated in the same practices — forming and testing hypotheses and building and using models. This establishes that the framework addresses a need its intended users recognise, and that the need is properly framed as difficulty in developing demanding practices rather than as a curricular deficit, which is the more defensible basis for the intervention and the one the evidence supports.

C2 Teachers' reported classroom difficulties correspond, affordance by affordance, to the immersive affordances the project defines, and the anticipated benefit exceeds the reported difficulty in every case.

The instrument pairs each classroom source of difficulty with the immersive affordance intended to answer it. Pooled across all respondents, the anticipated benefit of each affordance exceeds the reported frequency of the corresponding difficulty for all five pairs, and the two affordances rated most beneficial — visiting inaccessible real places (4.07) and manipulating variables in real time (4.01) — answer the difficulties teachers report most strongly. This establishes the empirical warrant for the project's central design premise: that immersive technology is to be treated through specific pedagogical affordances matched to specific teaching difficulties rather than as a device category. It is the survey's principal contribution to the deliverable and the evidential foundation for the affordance-led scenario design of the subsequent work packages.

C3 The affordance-led premise holds not only in aggregate but within every national system, differing across systems in level rather than in pattern.

The perceived benefit of the five affordances has a broadly similar benefit profile across the five systems, with visiting real places and manipulating variables live the most valued in most, and no affordance falling out of favour in any; overall benefit ranges from 3.84 in Belgium to

4.53 in Croatia, and no country rates any single affordance below 3.57. The differences between systems are of level, and track the acuteness of the difficulties each reports, rather than of shape. This establishes that the affordance-led treatment is robust across materially different national contexts — constraints, curricula and readiness all varying — and is not an artefact of aggregation, which strengthens its use as a design principle for a five-country intervention.

C4 Teachers attach consistent, practical and demanding conditions to the use of immersive activities, and these conditions restate the constraints they report most strongly.

Across the consortium the most strongly required conditions of use are compatibility with class size and available devices (pooled 4.17), fit with lesson time (4.07) and readiness for use without heavy preparation (3.99) — operational rather than aspirational requirements, and the same constraints teachers report on their current teaching, restated as design requirements. This establishes the practical envelope within which immersive scenarios must be built: activities that demand heavy preparation, specialist equipment or reduced class sizes will not be usable by the workforce as it is. Two conditions vary by system and pass directly into national adaptation — curriculum and examination alignment, required most strongly in Greece and Croatia, and national-language availability, most pressing in Greece.

C5 Teachers judge themselves pedagogically prepared but technically unprepared, in the same pattern in every system.

Self-assessed readiness has an identical shape across the consortium: the pedagogical dimensions — classroom management and assessment — sit at or above the scale mid-point, while the technical dimensions — knowledge of immersive tools and technical skill — sit below it, knowledge of tools being the single weakest readiness item overall (pooled 2.86) and below the mid-point in four of the five systems. This establishes a precise instruction for the training: it must build technical fluency from a low base while assuming and drawing on the pedagogical competence teachers already hold, rather than treating immersive teaching as a wholly new pedagogy. It also indicates that train-the-trainer and course design should foreground the operation of the medium, not the pedagogy of teaching with it.

C6 The intervention addresses a workforce with little or no prior experience of the medium.

In every system a majority of respondents have never used immersive tools or have tried them only once or twice — from 57% in the most experienced sample, Belgium, to 73% in Slovenia — and regular or frequent use is rare everywhere, absent altogether in the Slovenian and Croatian returns. This establishes the baseline from which the training begins and reinforces the usability conditions of Finding C4: a first-encounter workforce cannot absorb

activities that presuppose fluency it does not yet have, so the earliest training and the earliest scenarios must be usable by novices.

C7 Unprompted, teachers extend the immersive rationale beyond the science core into the mathematics and artificial-intelligence strands.

Without being asked about specific strands, respondents in Croatia identified three-dimensional geometry and the graphs of functions, and respondents in Slovenia the interpretation of machine-learning models, as content where immersive approaches would help. This establishes independent, teacher-originated confirmation that the framework's logic holds in the mathematics and artificial-intelligence strands, which the deliverable treats provisionally and as a supplement respectively. It does not alter the operative framework, which remains as fixed by the derivation and the workshop, but it is evidence to be carried into the work that elaborates those strands.

What the survey establishes for the project

P1 The teacher-capability framework is validated for the pilot.

The survey confirms, among a wider body of teachers than took part in deriving or prioritising it, that the framework addresses a recognised difficulty and treats the medium in a way its intended users endorse. The framework may therefore be taken into the pilot as the tested basis for course and scenario design, its empirical standing extended from the consortium to the five national teaching populations without any change to its content.

P2 The affordance-led design principle is warranted for the scenario work of WP3 and beyond.

Finding C2, holding across all five systems by Finding C3, gives the scenario-design activities their empirical foundation: scenarios are to be built around the pedagogical affordance that answers a specific teaching difficulty, and the survey identifies which affordances teachers value most and against which difficulties. This is the principal handover from the survey to the downstream work packages.

P3 The design envelope for training and scenarios is specified.

Findings C4 and C6 together bound the design: scenarios and microcourses must be usable within current class sizes, lesson time and device provision, without heavy preparation, and by teachers with little prior experience. These are not preferences but conditions of adoption, and they constrain the scenario templates and the course designs that Chapter 6 and the subsequent work packages produce.

P4 The training must build technical fluency from a low base while leveraging existing pedagogy.

Finding C5 sets the pedagogical strategy for the training and the train-the-trainer model: technical operation of the medium is the priority for initial training, whereas the pedagogy of teaching with immersive activities can be developed from the competence teachers already report. This shapes the sequencing and emphasis of the course designs described in Chapter 6.

P5 National adaptation must be differentiated, and the mathematics and AI strands have independent support for their elaboration.

Findings N1 to N5 and C4 establish that national adaptation cannot be uniform: the Greek and Croatian systems require curriculum- and examination-aligned, language-available provision built for acute constraint, the Flemish system requires attention to institutional enablement and presumes higher readiness, and the Italian and Slovenian systems require technical capacity-building against a moderate constraint profile. Finding C7 further establishes that the mathematics and artificial-intelligence strands, though outside the operative framework, carry independent teacher support for their relevance and should be elaborated accordingly. Together these findings direct the national-adaptation guidance of Chapter 6 and the strand development of the subsequent work packages.



6. Implementation guidelines

6.1 Scope

The preceding chapters describe how the teacher-capability framework was derived and the components it contains; the framework is presented in full in Annex II. This chapter addresses its use. The framework is intended as a working instrument for the design of teacher education and for classroom practice, and its practical applications are not immediately apparent from the framework itself. These guidelines therefore set out how the framework is used, by whom, and for what purpose, so that it functions as a practical basis for course design and teaching, rather than merely as a reference document.

The guidance addresses two audiences. The first is course and programme developers, who require guidance on how a capability can be translated into a unit of training, how capabilities combine into structured pathways, and how the framework can be applied for purposes extending beyond the design of an individual course. The second is educators, who use the framework in their professional practice: as a statement of what a teacher must be able to do to develop a given ability in students, and as a reference against which to assess their own practice. The framework serves both audiences, with different applications for each, and the guidelines are organised accordingly.

These applications follow from the structure of the framework. Each capability is specified as a discrete unit, anchored to the competency from which it derives and resolved into a set of teacher skills that span the full teaching arc — what should be taught, where students are likely to struggle, how the content should be taught and framed, and how students' learning should be assessed — with each skill defined at a stated cognitive level. This modular structure allows the framework to be applied in more than one way and for more than one purpose. The sections that follow set out the principal ways in which its capabilities can be combined and recombined, the principal uses it supports, and a small number of worked examples showing how the framework can be applied to selected capabilities from Annex II.

6.2 Guidelines

The framework supports several distinct applications. This section identifies the principal scenarios in which it is used and, for each, sets out general guidance on how the framework is read and applied. The scenarios are not mutually exclusive: a single programme of work may draw on more than one of them. They are distinguished here because each uses the framework for a different purpose and at a different level of granularity and engages with its structure differently— some take the capability as the unit of work, others the individual teacher skill, and others the framework as a whole.



6.2.1 Designing courses and microcourses

The framework is constructed so that each capability provides the specification for a microcourse. Unlike a generic skill inventory, which must first be contextualised and then assigned an appropriate cognitive level before it can yield learning objectives, both steps have already been completed through the derivation process in its derivation: each teacher skill is already specific to the teaching of a defined ability, and each already carries a draft objective at a stated cognitive level. The **designer of a teacher training programme** therefore does not construct objectives — the framework supplies them — and course design can therefore begin from the supplied objectives and focus on sequencing, learning activities, resources, assessment and the allocation of learning time.

The transformation chain

As a default design rule, one capability can form the basis of one microcourse; each of its teacher skills yields one learning objective, taken directly from the objective column.

1 Capability → up to 5 Teacher Skills → up to 5 Learning Objectives

Because the skills span the five teaching dimensions, the resulting objectives cover the whole teaching arc — content, misconceptions, instruction, framing and assessment — and a microcourse built on one capability can provide systematic coverage of its principal pedagogical dimensions. Environmental capabilities, which carry three skills rather than five, yield three objectives by the same rule. The learning architecture this produces is set out below.

Learning architecture

Table 17 Learning architecture.

Component	Quantity	Contains	Obtained by
1 Capability	1 microcourse	up to 5 teacher skills	reading the framework row
1 Teacher skill	1 learning objective	objective at a fixed cognitive level	lifting the objective column
1 Microcourse	3–5 LOs	the full teaching arc	—

Credit and learning time

Allocating credit per objective keeps microcourses comparable and stackable. The figures below are indicative; the values applied in any given programme are fixed in course design.

Table 18 Credit and learning time.

Component	ECTS (indicative)	Learning time
Each learning objective (1 skill)	0.2 ECTS	5–6 hours
Each microcourse (1 capability, 3–5 LOs)	0.6–1 ECTS	15–30 hours

Component	ECTS (indicative)	Learning time
Each category (all its capabilities)	2–3 ECTS	50–75 hours

Worked example

Capability 2.1, Developing students' experimental-design reasoning, illustrates the chain. Its five skills become a five-objective microcourse with no further authoring, each objective carrying the cognitive level at which it is pitched.

Table 19 Worked example of teaching focus and learning objective.

Skill	Teaching focus	Learning objective (cognitive level)
2.1.1	What to teach	Identify the contexts that admit genuine experimental design (<i>Understand</i>)
2.1.2	Where students struggle	Anticipate typical design errors — confounded variables, missing controls, manipulations that cannot answer the question (<i>Analyse</i>)
2.1.3	How to teach it	Scaffold student-led design through partial templates, peer review and design-then-revise cycles (<i>Apply</i>)
2.1.4	How to frame it	Frame experimental design as a creative scientific act, not a fixed procedure (<i>Understand</i>)
2.1.5	How to assess it	Create rubrics that judge a design by its internal logic, independent of whether it produced results (<i>Create</i>)

The five objectives are the microcourse outcomes; their cognitive levels provide a progression from understanding to creation; and the final objective also provides the blueprint for assessment, so the course is assessed against criteria aligned with the outcomes it develops.

Learning pathways

Because each capability is a complete, a complete unit with its own explicit competency anchor, the same building blocks support several routes through the framework. A programme may offer any of the pathways set out below.

Table 20 Learning pathways.

Pathway	Components	Scope	Description
Single microcourse	1 capability = 3–5 LOs	0.6–1 ECTS	Complete development of one ability across the full teaching arc
Combination path (thematic track)	several capabilities across categories	sum of parts	A track on a broader competence — e.g. scientific argument, chaining source evaluation (3.1), argument construction (3.2) and decision-making (3.3)
Category pathway	all capabilities in one category	2–3 ECTS	Full coverage of one PISA competency — e.g. all of evidence, argument and decision

Pathway	Components	Scope	Description
Dimension slice (cross-cutting module)	one teaching dimension across many capabilities	varies	A focused module — e.g. assessment design, drawing the “how to assess it” skill (x.x.5) from every capability
Cognitive-progression pathway	skills ordered by objective level	varies	A route climbing from Understand-level to Create-level work rather than treating all skills as equivalent

Reuse across units

Objectives are not confined to one microcourse. A skill that is the primary outcome of its own capability may serve as a supporting outcome elsewhere — the data-interpretation objective (2.2) reinforcing an argument-construction microcourse, for instance — so the framework supports a many-to-many mapping between objectives and delivered units, with each unit carrying three to five objectives for a manageable scope.

6.2.2 Designing Train the trainer courses

A train-the-trainer programme prepares those who will, in turn, develop these capabilities in other teachers. It draws on the same framework but applies it at the next level of the training cascade: where a teacher microcourse develops a capability in students, a trainer module develops, in teachers, the ability to do that. The framework supports this directly, because the structure used to specify how a capability is enacted in teaching can also be adapted to specify how that capability should be developed in teachers.

The transformation

A trainer module is built on the same capability as the teacher microcourse that it is designed to cascade, with two things added. First, command of the capability is a prerequisite: a a trainer must first possess and be able to enact the capability before training it, so the teacher microcourse described in 6.2.1 is the foundation on which the trainer module rests. Second, the five teaching dimensions are applied again at the level of teacher education, taking the capability — rather than the underlying student ability — as the object of instruction. The trainer learns what to convey about the capability, where teachers struggle to acquire it, how to develop it in them, how to frame it, and how to judge that they have acquired it.

Teacher microcourse (command the capability) → Trainer module (develop it in teachers)

The added dimension

One dimension carries genuinely new content. In the teacher microcourse, the “where students struggle” skill names the misconceptions students bring to the ability. In the trainer module, the corresponding question is where teachers struggle to acquire and enact the capability — the points at which they revert to procedural teaching, under-estimate student difficulty, or avoid open-ended tasks. This is the trainer’s distinctive knowledge and cannot be

reduced to the trainer's own mastery of the capability. It is the part of the trainer module that the framework prompts but does not supply ready-made: it is developed through pilot delivery and refined as the train-the-trainer process is implemented and extended.

Worked example

Capability 2.1 again illustrates the move. Each teacher-level objective from 6.2.1 yields a trainer-level objective that takes the teacher, rather than the student, as the learner.

Table 21 Worked example of capability 2.1.

Skill	Teacher-level objective (from 6.2.1)	Trainer-level objective
2.1.1	Identify the contexts that admit genuine experimental design (<i>Understand</i>)	Enable teachers to identify such contexts, and model how to locate them within their own curriculum (<i>Apply</i>)
2.1.2	Anticipate typical student design errors (<i>Analyse</i>)	Prepare teachers to anticipate these student errors, and anticipate where teachers themselves overlook or reproduce them (<i>Analyse</i>)
2.1.3	Scaffold student-led experimental design (<i>Apply</i>)	Model the scaffolding moves and coach teachers in enacting them with their own classes (<i>Apply</i>)
2.1.4	Frame experimental design as a creative scientific act (<i>Understand</i>)	Instill this framing in teachers and address discomfort with open-ended, non-procedural design (<i>Understand</i>)
2.1.5	Create rubrics that judge a design by its internal logic (<i>Create</i>)	Train teachers to build and apply such rubrics, and assess the rubrics teachers produce (<i>Create</i>)

Selection and cascade

Not every capability is included in the training cascade. A programme selects the capabilities to be included in wider teacher-training provision and builds one trainer module for each. Trainers prepared in this way then deliver the corresponding teacher microcourses, so the trainer programme mirrors the teacher programme at the next level of the training cascade, and the same building blocks serve both. The resulting multiplier effect — each trained trainer reaching many teachers — is why selection at this level is deliberately narrower than the full framework.

6.2.3 Teacher self-assessment

The framework also serves educators directly, without an intermediary designer. A practising teacher interprets each capability as a statement of what they must be able to do to develop a given ability in their students, and uses it to judge which areas of their practice are well developed and which require further development. Used this way the framework is a self-diagnostic instrument as well as a design tool, and the two uses draw on the same structure.

The framework as a self-diagnostic

A separate instrument is not necessarily required, because each teacher skill already names a discrete professional action or area of knowledge that can be assessed. The five teaching dimensions of a capability therefore convert directly into five self-assessment prompts — one for content, one for student difficulty, one for instruction, one for framing, one for assessment — each answered on a simple scale such as not yet, partly, or confidently. A prompt that a teacher cannot answer confidently identifies an area in which professional development may be needed, identified precisely rather than in general terms.

1 Capability → 5 self-assessment prompts → a profile of strengths and gaps

Worked example

Capability 2.1 illustrates the instrument. Each of its five skills becomes a prompt the teacher answers about their own practice.

Table 22 Worked example connecting skill, teaching focus and self-assessment prompt.

Skill	Teaching focus	Self-assessment prompt
2.1.1	What to teach	Can I identify, in my own curriculum, the topics where students can design genuine experiments rather than follow set procedures?
2.1.2	Where students struggle	Can I anticipate the design errors my students typically make — confounded variables, missing controls, manipulations that cannot answer the question?
2.1.3	How to teach it	Can I scaffold student-led design — through partial templates, peer review and revision — without supplying the design myself?
2.1.4	How to frame it	Do I present experimental design as a creative scientific act rather than the execution of a fixed procedure?
2.1.5	How to assess it	Can I build a rubric that judges a design by its internal logic, independent of whether it produced results?

A teacher who answers the first two prompts confidently but not the last three knows that their gap is not in identifying where experimental design should be taught within the curriculum, but in scaffolding it, framing it and assessing it — a far more useful finding than a general sense that experimental design is difficult to teach.

From diagnosis to development

Applied across several capabilities, the prompts yield a capability profile: a map of which abilities a teacher can already develop in students across the full teaching arc, and which they cannot. That profile directs development. It identifies the microcourses of 6.2.1 most worth taking, and within a microcourse the dimensions that require the greatest attention, so that the time spent on professional development is concentrated where the identified development needs are greatest. The diagnostic and the design uses of the framework thus

reinforce one another: the same structure that specifies how a capability is built into a course specifies how a teacher judges their readiness to teach it.

6.2.4 National adaptation

The capabilities are expressed in terms general enough to hold across systems, but their realisation depends on where, and whether, the corresponding content is already present in a given curriculum. Adapting the framework to a national system therefore means mapping each capability against the national curriculum and determining what form of adaptation it requires in that context. The desktop survey reported earlier in this document provides the evidence for that judgement, and it is what makes the guidance in this deliverable country-specific rather than generic.

The three-status decision

For any capability, a curriculum developer establishes its status in the national curriculum. It falls into one of three cases, and each implies a different adaptation move.

Unlabelled → make explicit · Latent → surface and strengthen · Absent → introduce

A capability that is already substantially present in practice but not named or deliberately designed; the adaptation makes it explicit and attaches it to the provision that already carries it. A capability present only in latent form is partially developed and implicit; the adaptation surfaces it and supplies the scaffolding it lacks. A capability that is absent has not been identified anywhere in the curriculum; the adaptation introduces it as new content. The three moves differ in effort and in risk, which is why classifying correctly precedes designing.

Reading the status from the desktop survey

The classification is derived from the national analysis, not guessed. Three of that analysis's cross-system findings bear directly on it. The arts as a mode of scientific inquiry are not systematically embedded in any of the systems examined, so any capability resting on creative or design-based reasoning is, for now, in the absent case everywhere and must be introduced. STEAM-specific assessment is the weakest dimension throughout and most severe in one system, so the "how to assess it" skill of each capability is typically the one to be introduced rather than merely surfaced. And alignment to the PISA descriptors is implicit across systems and most acute for the epistemic and environmental dimensions, so capabilities of that kind are usually latent — present but unmapped — and the adaptation is chiefly one of making them explicit.

Worked example

Capability 4.1, Building students' environmental awareness, shows the same capability requiring different moves in different systems. Its status is determined from the cross-system synthesis.

Table 23 Comparison between countries - status in the curriculum and adaptation move.

System	Status in the curriculum	Adaptation move
Slovenia	Latent — carried as cross-curricular sustainability education, not labelled as a science capability	Make explicit; attach to existing science provision
Greece	Latent — present in recent curricula and cross-curricular skills provision	Surface and align to the capability
Croatia	Asserted at policy level, with little subject-level scaffolding	Introduce subject-level scaffolding
Italy	Named in national guidelines, weak in mainstream provision	Introduce, and build the assessment that is missing
Belgium (Fl.)	Embedded in the competence architecture, devolved to schools	Attach to that architecture; specify enactment

The capability does not change; what changes is the role it must play in each national context. In two systems it is already present and the task is to name and align it; in two others it is articulated in policy but lacks pedagogical scaffolding and must be developed more fully; in the fifth it is devolved and must be tied to an existing competence structure. A single generic instruction would not adequately address any of these contexts, which is precisely why adaptation is handled capability by capability and system by system.

From classification to a national map

Applied across the full framework, the three-status decision yields a national adaptation map: for each capability, its status in the country and the corresponding adaptation strategy. That map is the country-specific output the deliverable calls for. It also locates the points at which national systems genuinely diverge — where the same capability is latent in one country and absent in another — and so marks where shared course material can be used directly and where it must be adapted to existing national curriculum provision.

6.2.5 Scenario-based learning

The framework also supports learning organised around authentic scenarios rather than discrete exercises. Here a capability is not incorporated into a course for direct instruction, but translated into a situation in which students exercise the underlying ability themselves. The framework provides the design brief, so the scenario is assembled from the capability rather than devised around a theme alone.

The capability as a scenario brief

A scenario needs three things: an authentic situation, a task within it that requires the target ability, and a means of telling whether students succeeded. A capability provides all three through its constituent skills. The “how to teach it” skill, with its objective, gives the core task — the target ability enacted at the required cognitive level, at the level at which it is

performed. The “where students struggle” skill supplies the challenge: the difficulties it names become the decision points designed into the situation, so that the scenario tests whether the ability holds under the conditions that typically undermine successful performance. The “how to frame it” skill establishes the epistemic stance the scenario should convey, and the “how to assess it” skill provides the success criterion. The scenario is therefore not an exercise in unrestricted invention; its principal elements are derived systematically from the capability.

“how to teach it” = the task · “where students struggle” = the challenge · “how to assess it” = the success check

Worked example

Capability 3.3, Developing students’ science-based decision-making, assembles into a scenario brief as follows.

Table 24 Capability 3.3 connecting the scenario element to its contribution.

Scenario element	Drawn from	What it contributes
Authentic situation	3.3.1	A contemporary issue on which scientific data bears on a real decision, carrying genuine trade-offs
Core task (the ability in action)	3.3.3	Students move from the evidence to a justified decision, weighing data against values and constraints — the objective, at Apply level
Built-in challenge	3.3.2	The situation forces students to prioritise among competing considerations, exposing the habit of weighting everything equally
Framing	3.3.4	The scenario treats how a decision is justified as itself an object of analysis, surfacing competing scientific perspectives
Success criterion	3.3.5	Students must justify the decision under explicit constraints, not assert a preference — the assessment trigger

The result is a decision scenario in which students confront a real issue, are required by the scenario design to prioritise rather than assign equal weight to every consideration and are assessed on their ability to justify their decision under constraints rather than on the particular decision they reach.

Anchoring and authenticity

Because every element of the scenario is drawn from one capability, the scenario stays anchored to a defined ability and the cognitive level at which it is intended to be performed. This is what keeps scenario-based learning rigorous: the authenticity of the situation does not obscure the competence being developed, since the task, the challenge and the success criterion are all explicitly derived from the same capability. A scenario built in this way can then be implemented through the medium most appropriate to its objectives — discussion,

role-play, case study or an immersive simulated environment— without weakening its alignment with the target competence.

6.3 Worked Examples

6.3.1 STEAM Immersive Investigation

This example traces one teacher skill from the framework to a concrete, immersive STEAM investigation, to show how the framework, a pathway combining several capabilities and an immersive medium are integrated in a practical application. The setting is a polluted urban river. Rather than running a set protocol, students design and carry out their own investigation of the pollution, working towards specified learning goals. The teacher's task is to make that student-led investigation possible without determining the investigation design on their behalf — which is the skill on which the example is built.

The skill and the combination path

The core teacher skill is 2.1.3, scaffolding student-led experimental design (How to teach it, Apply): supporting students through the design of an investigation without supplying the method. It forms part of Capability 2.1, developing students' experimental-design reasoning, and the investigation integrates several additional capabilities within a single learning pathway — which is what makes it STEAM rather than science alone.

Table 25 Combination between capabilities and their contribution to the investigation.

Capability	What it contributes to the investigation
2.1 Experimental-design reasoning	The spine: students design the investigation themselves — the Science of the task
2.2 Data interpretation	Reading the measurements to locate the pollution source and warrant the claim — Science and Mathematics
4.1–4.3 Environmental science identity	Treating the river as genuine environmental-science content and moving from awareness to action — the Environmental dimension
3.3 Science-based decision-making	Justifying a remediation under real constraints, and communicating it — carrying the Arts-as-reasoning dimension
— Immersive medium	The river is instrumented with virtual sensors the students deploy and read — the Technology dimension

The investigation and its goals

Students enter a virtual stretch of the river they can walk, sample and instrument. They work towards five goals, each developing part of the combination path.

Table 26 Student goal, STEAM and framework linking.

#	Student goal	STEAM	Framework
G1	Frame a researchable question — what is polluting the river, and how far does it spread?	S	2.1
G2	Design a valid sampling plan — sites above and below the suspected source, the variables to measure (pH, dissolved oxygen, turbidity, nitrates, temperature), controls and repetition, without confounding	S · T · M	2.1.3
G3	Carry out the investigation in the immersive river and record data from the virtual sensors	S · T	2.1 / 2.2
G4	Interpret the data to locate the source and warrant the claim, distinguishing correlation from causation	S · M	2.2
G5	Propose and communicate a remediation or monitoring action, justified under real constraints	E · A	4.3 / 3.3

The teacher's work: scaffolding without supplying

Skill 2.1.3 is enacted through the manner in which the teacher supports each goal, providing structure but not the design itself.

Table 27 Teacher's scaffolding move.

Goal	The teacher's scaffolding move
G1	Offers a partial question template and a starting list of variables, then has the students narrow the question themselves
G2	Supplies a design checklist — variables, controls, sampling sites, repetition — but not the plan; runs structured peer review of draft plans and permits a design-then-revise cycle
G3	Lets flawed designs run in the simulation rather than correcting them in advance, so students meet the consequences of their own choices in a fail-friendly space
G4	Prompts the correlation-versus-causation check without resolving it for the students
G5	Sets out the real constraints — cost, feasibility, competing uses of the river — and leaves the weighing to the students

Throughout the investigation the teacher maintains the distinction that defines 2.1.3: responsibility for the design remains with the students, while the teacher provides the scaffolding needed to make a sound design attainable. Enacting this skill is most demanding at G2, where the temptation to provide students with a correct plan is strongest and where withholding it is precisely what the skill requires.

The contribution of the immersive medium

The immersive medium is not incidental in this example; it creates the conditions needed to teach and practise the skill effectively at goals G2 and G3. A real polluted river cannot be sampled repeatedly and safely whenever required by a class, and a conventional paper-based exercise cannot reproduce the consequences of a flawed design with the same immediacy. The simulation permits both: a plan that samples only downstream, or that omits an upstream control, produces data that are demonstrably insufficient to answer the research

question, so that the weakness of the design is revealed by the investigation itself rather than explained to students in advance by the teacher. Because the virtual sensors return readings immediately, students can revise a design and re-test it within a single lesson, and the dispersion of the pollutant can be represented on a compressed timescale. This is the affordance that the project's immersive approach is intended to exploit, applied to the requirements of a clearly defined teacher skill rather than introduced for its own sake.

6.3.2 Green Capability

This second example concerns the environmental capabilities of the framework. It follows the same method as the first: a single teacher skill is translated into an immersive STEAM investigation, while integrating additional capabilities into a combined learning pathway. The investigation takes place in a simulated city and focuses on its carbon emissions. Students investigate the sources of the emissions and which interventions could reduce them, working towards specified learning goals, and the teacher's task is to develop their ability to analyse environmental action at different scales without prescribing the conclusions students should reach.

The skill and the combination path

The core teacher skill is 4.3.2, supporting students to analyse environmental action at different scales (How to teach it, Apply): equipping students to analyse the levels at which environmental action can be taken— individual, community, institutional and structural— and to distinguish between points of leverage that are more or less amenable to intervention it forms part of Capability 4.3 confronts the misconception recorded in 4.3.1, developing students' environmental agency, and addresses the misconception identified in Skill 4.3.1, that environmental problems can be addressed primarily through individual consumer choices. The investigation integrates the other relevant environmental capabilities, together with data interpretation, into a single combination path.

Table 28 Connection between capabilities and their contribution to investigation.

Capability	What it contributes to the investigation
4.3 Environmental agency	The spine: analysing where action operates — individual, community, institutional and structural — and distinguishing tractable from intractable leverage. The Environmental dimension
4.1 Environmental awareness	Building the systems-level map of the problem across sectors rather than as a single cause
4.2 Environmental concern	Sustaining constructive, evidence-based engagement in place of fatalism as the scale of the problem becomes apparent
2.2 Data interpretation	Quantifying and comparing the impact of interventions — Science and Mathematics
— Immersive medium	A walkable, instrumented city in which interventions are implemented and their effects read — the Technology dimension

The investigation and its goals

Students enter a simulated city they can traverse and instrument, in which sources of emissions and candidate interventions are both represented. They work towards five goals, each developing part of the combination path.

Table 29 Linking between student goal, STEAM and framework.

#	Student goal	STEAM	Framework
G1	Map the system — identify the principal sources of the city's emissions across sectors (transport, buildings, energy, industry)	S · E	4.1
G2	Classify candidate interventions by scale of action — personal, civic, collective and structural	E	4.3.2
G3	Model and compare the impact of interventions at each scale using the simulation's data	S · M	4.3.2 / 2.2
G4	Identify the points of genuine leverage, recognising the limited structural reach of individual consumer choices	E	4.3.2 / 4.3.1
G5	Propose and justify a structurally framed action plan that integrates scientific evidence with ethical and equity considerations	E · A	4.3.3 / 3.3

The teacher's work: developing the analysis without prescribing it

Skill 4.3.2 is enacted through the manner in which the teacher develops the scale analysis, supplying the analytic frame but not the judgement that follows from it.

Table 30 Skill 4.3.2 and teacher's scaffolding move.

Goal	The teacher's scaffolding move
G1	Provides a system-map template and sector prompts; the students populate the emission sources themselves
G2	Supplies the four scales of action as an analytic frame, and has the students classify interventions rather than ranking them in advance
G3	Allows the students to implement interventions in the simulation and read the impact data directly, rather than indicating which will succeed
G4	Has the students compare the aggregate effect of personal-only and structural interventions in the simulation, so that the individualisation assumption is tested against the model's output rather than corrected by assertion
G5	Sets out the ethical and equity constraints — who bears the cost and who benefits — and requires the action plan to be framed at structural as well as personal scales

Across these moves the teacher develops the capacity the skill defines: students learn to identify the scale at which action is most appropriate and to assess its likely impact using

evidence. The key stage is G4, where the assumption that responsibility lies primarily with individuals is most directly tested, and where the teacher's role is to allow the evidence from the comparison to speak for itself rather than anticipating or prescribing the conclusion.

The contribution of the immersive medium

The immersive medium is central to goals G3 and G4. The comparison central to this skill — the differential reach of action at the individual and structural levels — is difficult to demonstrate directly in a conventional lesson, where the relative impact of an individual choice and a structural policy is often discussed only in hypothetical or qualitative terms. In a simulated city the comparison becomes observable and quantitative: students can implement a change in personal behaviour across the population, and then model a structural intervention, and read the resulting difference in system-level emissions. The misconception identified in Skill 4.3.1" is thereby challenged by evidence generated by the students themselves. A real city cannot be subjected to such experimentation; the simulation permits interventions to be tested and their system-level consequences represented over compressed time, which is the affordance the project's immersive approach is intended to exploit, applied to a specific requirement of a specific skill.



7. Conclusions

This deliverable set out to translate the competency architecture of the PISA 2025 Science Framework into a structured and usable account of the capabilities teachers require in order to develop scientific competence in their students, and to establish how that account is applied in the design of immersive STEAM education. It has done so through a single sustained argument, carried across the preceding chapters, that reasons from an authoritative account of what students should be able to do to the capabilities a teacher must hold for them to do it. This concluding chapter draws the elements of that argument together, states what the deliverable establishes, records the boundaries within which it holds, and sets out its relationship to the curriculum and scenario-design work that follows.

7.1 What the deliverable establishes

The method of the deliverable is a reversal. Rather than beginning from an inventory of what teachers lack and reasoning towards remediation, it begins from the student competencies defined by the PISA 2025 Science Framework and derives, competency by competency, the teacher capability required to develop each. This derivation-first approach is the deliverable's organising principle, and each stage of the work contributes a distinct element to it. The desk-based review established the baseline: across five independently analysed national systems, the competences that immersive STEAM education seeks to develop are broadly present but latent — seldom explicitly labelled, rarely deliberately designed or assessed, and not mapped to the PISA architecture — and the convergence of this finding across systems that otherwise differ markedly indicated that the deficits are structural rather than idiosyncratic, and so amenable to a common response adapted to national contexts. The systematic derivation then translated the student-side competencies of PISA 2025 into a pool of teacher capabilities through a four-layer chain, anchoring each capability in the student outcome it serves, the pedagogical content knowledge it requires, the immersive affordance through which it can be developed, and the European policy context in which it sits. The consortium workshop prioritised this pool into the subset taken forward for the pilot and elaborated each selected capability into the concrete teaching actions that enact it. The online survey then carried the result to a wider body of teachers, validating the framework and gathering the evidence on localisation and on the immersive medium that the later guidance requires.

The principal output of this work is the ImSTEAM Teacher Capability Framework, presented in Annex II. Derived entirely from the PISA 2025 Science Framework, it comprises four categories, eleven capabilities and forty-nine teacher skills, covering the three science



competencies — explaining phenomena scientifically, constructing and evaluating scientific enquiry and interpreting data, and using evidence, argument and information to reach decisions — together with the environmental science-identity dimension the framework defines alongside them. Each capability is resolved into teacher skills that span the full teaching arc, from what should be taught, through where students are likely to struggle and how the content should be taught and framed, to how learning should be assessed, and each skill carries a draft objective at a defined cognitive level. The consequence is that a capability functions as a specified unit of training rather than a general statement of desirable competence, which is what allows the framework to be translated directly into the design of courses, microcourses and professional-learning pathways.

The survey established three things about this framework, and they are the empirical foundation on which the subsequent work rests. First, the framework addresses a need its intended users recognise: teachers across the consortium report the scientific practices the framework is built around as genuinely difficult to develop in students, consistently and in the same practices in most systems, with some national variation, so that the case for the intervention rests on difficulty rather than on a perceived curricular deficit. Second, the project's affordance-led treatment of immersive technology is sound: the classroom difficulties teachers report correspond, affordance by affordance, to the benefits they anticipate from immersive activities, and for every paired affordance the anticipated benefit exceeds the reported difficulty, in each national context as well as in aggregate. Third, the conditions of use and the readiness against which the training must be pitched are now specified: teachers require immersive activities to be compatible with their class sizes, devices and lesson time, and judge themselves pedagogically prepared but technically unready, the great majority having little or no prior experience of the medium. Taken together, these results extend the empirical basis of the framework from the consortium to a wider body of teachers without altering its content.

7.2 Boundaries of the deliverable

The deliverable establishes the framework and the principles governing its use; it does not itself constitute the courses, curricula or learning resources built from it, which are the object of the subsequent work. Three further boundaries, each a matter of deliberate design rather than omission, qualify what is presented here. The framework covers the scientific and environmental capabilities in full; the mathematics strand remains provisional and requires dedicated validation; the survey provides only an exploratory signal of its relevance; and the treatment of artificial intelligence is retained as a supplement rather than incorporated into



the operative framework — in both cases reflecting the principle that a capability is specified only where both the source framework and the relevant pedagogy support one. Notably, the survey's open responses lend independent, teacher-originated support to the relevance of both strands, pointing unprompted to three-dimensional geometry and the interpretation of artificial-intelligence models as content where immersive approaches would help, which is evidence to be carried into the work that elaborates them. The arts dimension, identified in the desk review as the most consistent gap across the five systems and therefore the clearest opportunity for intervention, is addressed at the level of enactment in the scenario-design work rather than as a parallel competency set within this framework. Finally, immersive technology is treated throughout as one medium among others, considered only where it develops a capability more effectively than conventional means; the deliverable is not a technical specification for immersive scenarios.

7.3 Relationship to the work that follows

This deliverable is the analytical foundation for the intervention that follows, and its outputs pass into that work in specific ways. The teacher-capability framework of Annex II is the PISA-anchored reference from which the courses and microcourses of the training are designed, and the implementation guidance of Chapter 6 sets out how a capability is translated into a unit of training, how capabilities combine into pathways, and how the framework is adapted to a national curriculum. The survey findings shape that design directly: the affordance-led correspondence between teaching difficulty and immersive benefit is the empirical warrant for the affordance-led scenario design that follows; the conditions of use and the low base of prior experience define the envelope within which scenarios and microcourses must be built; and the readiness profile — pedagogically prepared, technically unready — sets the emphasis and sequencing of the training and the train-the-trainer model. The national differences the survey and the desk review record, most sharply the acute examination and language-resource constraints reported in Greece and the organisational constraints reported in the Flemish Community, pass into the national-adaptation guidance, which the subsequent work packages apply as they localise the provision. On this foundation the curriculum and scenario-design activities that build on this deliverable proceed, carrying the framework from a validated specification into the immersive STEAM teacher education the project is established to deliver.



Annex I The PISA skills Pool

Purpose of the Skills Pool

The Skills Pool presented in this Annex is not intended as a reproduction of the PISA 2025 Science Framework. Rather, it constitutes the complete derivation space from which the ImSTEAM Teacher Capability Framework is constructed.

PISA defines what students should know and be able to do. The purpose of this Annex is to make explicit the pedagogical implications of those student competencies by identifying the corresponding capabilities teachers require in order to develop them. The resulting capability statements therefore represent an original teacher-side interpretation of the PISA competency architecture and should be understood as project-generated constructs rather than OECD descriptors.

The derivation follows the backward-design principle described in Chapter 3. Each student sub-ability is examined through the lens of Pedagogical Content Knowledge (PCK) and transformed into one or more teacher capabilities. The process is traceable and auditable because every capability can be linked directly to:

1. the originating PISA competency;
2. the specific student sub-ability;
3. the PCK dimension used in the derivation;
4. the resulting teacher capability.

This Annex therefore serves two purposes:

- to document the complete derivation process;
- to preserve traceability between the source framework and the operational framework adopted by ImSTEAM.

The full pool contains capabilities that exceed the scope of the pilot implementation. Their inclusion is deliberate, ensuring transparency in coverage and providing a resource for future exploitation beyond the project lifetime.

Relationship with Annex II

Not all capabilities included in the Skills Pool are carried forward into the operational ImSTEAM Teacher Capability Framework. Selection was conducted using the prioritisation process described in Chapter 4.

For transparency, each capability is assigned one of the following statuses:

- **Selected for Framework** – capability included in the operational ImSTEAM Teacher Capability Framework (Annex II);



- **Retained for Reference** – capability preserved within the Skills Pool but not prioritised for pilot implementation;
- **Research and Future Exploitation** – capability retained for future curriculum development and further research activities.

Student Competency A1: Explain phenomena scientifically

PISA Competency: A1 — Explain phenomena scientifically (Box 3, p. 21).

Sub-ability A1.1 — Recalling and applying scientific knowledge

Verbatim PISA: *"Recall and apply appropriate scientific knowledge." (Box 3, p. 21)*

Table 31 Sub-ability A1.1 - Recalling and applying scientific knowledge.

ID	PCK anchor (RCM)	Teacher capability	Bloom	One-line gloss
T-A1.1.a	Curricular Saliency	Identifying core scientific knowledge for application	Understand	Distinguishes essential ideas students must internalise from supporting detail, in light of what application tasks demand.
T-A1.1.b	Students' Understanding	Anticipating student misapplications	Analyze	Anticipates where students typically misfire when applying scientific knowledge — wrong-context formula use, surface pattern-matching, over-generalisation.
T-A1.1.c	Instructional Strategies	Selecting representations that support transfer	Apply	Chooses and sequences representations, analogies, and worked-example/problem pairs that help

ID	PCK anchor (RCM)	Teacher capability	Bloom	One-line gloss
				students move from recall to application.
T-A1.1.d	Orientations to Teaching Science	Framing science as applied, not memorised	Understand	Conveys an epistemic stance that scientific knowledge is provisional and situated, and is for use rather than for retention alone.
T-A1.1.e	Assessment Knowledge	Creating tasks that need real understanding	Create	Designs tasks that require genuine application of scientific knowledge rather than recall of a familiar problem type. (Worked example, Methodology v1.1 Section 9.)

Sub-ability A1.2 — Using and translating between forms of representation

Verbatim PISA: "Use different forms of representations and translate between these forms." (Box 3, p. 21)

Table 32 Using and translating between forms of representation.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A1.2.a	Curricular Saliency	Identifying the representations that carry the topic	Understand	Identifies which representations are constitutive of a given topic (e.g. Punnett squares for genetics; free-body diagrams for mechanics) and which are illustrative.
T-A1.2.b	Students' Understanding	Diagnosing representational fluency gaps	Analyze	Diagnoses where students stall when moving between representations — e.g. graph to equation, particulate to macro, prose to diagram.
T-A1.2.c	Instructional Strategies	Constructing tasks that require translation between representations	Apply	Constructs tasks that explicitly require students to produce one representation from another, not merely interpret one in isolation.
T-A1.2.d	Orientations to Teaching Science	Framing representations as thinking tools, not illustrations	Understand	Conveys to students that representations are how scientists reason, rather than decorative or mnemonic aids.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A1.2.e	Assessment Knowledge	Designing assessments of translation fluency, not recognition	Create	Designs assessments where the student must generate a representation from another, distinguishing surface familiarity from genuine fluency.

Sub-ability A1.3 — Making and justifying scientific predictions

Verbatim PISA: *"Make and justify appropriate scientific predictions and solutions."* (Box 3, p. 21)

Table 33 Making and justifying scientific predictions.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A1.3.a	Curricular Saliency	Identifying topics where prediction is the central scientific move	Understand	Distinguishes curricular topics in which prediction is constitutive of the science (mechanics, ecology, chemical equilibrium) from those in which it is incidental.
T-A1.3.b	Students' Understanding	Anticipating where students mistake plausibility for justification	Analyze	Anticipate the typical patterns where students predict on intuition or plausibility rather than on principled reasoning from theory or prior data.
T-A1.3.c	Instructional Strategies	Scaffolding the prediction-justification-revision cycle	Apply	Designs instruction in which prediction precedes data, justification is made explicit, and predictions are revised in light of evidence.
T-A1.3.d	Orientations to Teaching Science	Framing prediction as constitutive of scientific work	Understand	Conveys to students that prediction-and-test is what science does, not an optional pedagogical add-on or a guessing game.
T-A1.3.e	Assessment Knowledge	Designing rubrics that reward justification, not prediction correctness	Create	Designs rubrics that reward the reasoning chain — premises, inference,

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
				scope — rather than only whether the prediction matched the outcome.

Sub-ability A1.4 — Identifying, constructing, and evaluating models

Verbatim PISA: *"Identify, construct, and evaluate models."* (Box 3, p. 21)

Table 34 Identifying, constructing, and evaluating models.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A1.4.a	Curricular Saliency	Identifying foundational models for a topic	Understand	Distinguishes essential scientific models (e.g. particle model, plate tectonics) from didactic simplifications and recognises what each model is for.
T-A1.4.b	Students' Understanding	Anticipating student confusion that may arise from misunderstanding the model in relation to reality	Analyse	Anticipates the two failure modes — taking the model literally (e.g. orbiting electrons) or dismissing it as 'just a model' — and the contexts in which each appears.
T-A1.4.c	Instructional Strategies	Sequencing modelling — building, testing, revising	Apply	Designs instruction in which students construct, evaluate, and revise models against evidence, rather than receiving them as finished products.
T-A1.4.d	Orientations to Teaching Science	Framing models as provisional and purpose-built	Understand	Conveys that models are constructed to answer specific questions and are evaluated by fitness-for-purpose, not by truth or correctness.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A1.4.e	Assessment Knowledge	Designing assessments of model evaluation against fitness-for-purpose criteria	Create	Designs assessments where the student judges a model against criteria (predictive scope, parsimony, agreement with data), not against 'the right answer'.

Sub-ability A1.5 — Recognising and developing explanatory hypotheses

Verbatim PISA: "Recognise and develop explanatory hypotheses of phenomena in the material world." (Box 3, p. 21)

Table 35 Recognising and developing explanatory hypotheses.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A1.5.a	Curricular Saliency	Identifying phenomena that admit student-level hypothesising	Understand	Identifies which curricular phenomena are productive sites for student-generated hypotheses, and which require prior content beyond the students' current reach.
T-A1.5.b	Students' Understanding	Distinguishing hypothesis from guess in student work	Analyze	Recognises the difference between a student offering a testable explanation and one naming a possibility, and surfaces the distinction productively.
T-A1.5.c	Instructional Strategies	Eliciting and refining student hypotheses toward testability	Apply	Facilitates discussions in which student hypotheses are surfaced, compared, and refined — through questioning and peer critique — toward testable form.
T-A1.5.d	Orientations to Teaching Science	Framing hypothesising as a student-accessible scientific practice	Understand	Conveys that students can do real hypothesising with appropriate scaffolding, rather than only receiving

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
				teacher-given hypotheses to test.
T-A1.5.e	Assessment Knowledge	Designing rubrics for hypothesis quality on testability, scope, and theoretical coherence	Create	Designs rubrics that recognise the structural criteria of a good explanatory hypothesis, distinct from whether it turns out to be confirmed.

Sub-ability A1.6 — Explaining the societal implications of scientific knowledge

Verbatim PISA: "Explain the potential implications of scientific knowledge for society." (Box 3, p. 21)

Table 36 Explaining the societal implications of scientific knowledge.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A1.6.a	Curricular Saliency	Identifying socially-consequential content within the curriculum	Understand	Identifies where the curriculum's scientific content has live societal stakes — vaccines, climate, energy, GM food, AI — and which are most live in the local context.
T-A1.6.b	Students' Understanding	Anticipating value-laden student reasoning on science-society questions	Analyze	Recognises when student responses are driven by identity, family position, or values rather than by understanding, without dismissing those drivers.
T-A1.6.c	Instructional Strategies	Facilitating evidence-based deliberation on science-society questions	Apply	Leads classroom deliberation in which evidence and values are kept distinct but both visible, and disagreement is conducted on principled grounds.
T-A1.6.d	Orientations to Teaching Science	Framing science as having legitimate societal stakes	Understand	Conveys that science is in society, not separate from it, while preserving the fact-value distinction and avoiding advocacy.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A1.6.e	Assessment Knowledge	Designing tasks that evaluate reasoning about implications, not student opinion	Create	Designs tasks that evaluate the structure of students' reasoning about implications — evidence used, scope, trade-offs — independently of which side they take.

Student Competency A2: Construct and evaluate designs for scientific enquiry; interpret data critically

PISA Competency: A2 — Construct and evaluate designs for scientific enquiry and interpret scientific data and evidence critically (Box 4, p. 21).

Sub-ability A2.1 — Identifying the research question of a scientific study

Verbatim PISA: "Identify the question in a given scientific study." (Box 4, p. 21)

Table 37 Identifying the research question of a scientific study.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A2.1.a	Curricular Saliency	Identifying curricular sites for research-question analysis	Understand	Knows which curricular topics carry primary studies suitable for student-level analysis, and where in the sequence such analysis is most productive.
T-A2.1.b	Students' Understanding	Diagnosing topic-versus-question confusion in student readings	Analyze	Anticipates that students typically name what a study is about (its topic) when asked for its question, and detects this conflation in student work.
T-A2.1.c	Instructional Strategies	Modelling question-extraction from primary sources	Apply	Demonstrates the moves for finding the actual research question in an abstract or methods section — including the test 'what would count as an answer?'.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A2.1.d	Orientations to Teaching Science	Framing science as fundamentally question-driven	Understand	Conveys that every scientific activity has a question behind it — explicit or implicit — and that finding the question is part of reading scientifically.
T-A2.1.e	Assessment Knowledge	Designing tasks for question-identification on unfamiliar studies	Create	Designs tasks where students must articulate the research question of a study they have not been taught about, distinguishing this from topic recall.

Sub-ability A2.2 — Proposing an appropriate experimental design

Verbatim PISA: "Propose an appropriate experimental design." (Box 4, p. 21)

Table 38 Proposing an appropriate experimental design.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A2.2.a	Curricular Saliency	Identifying curricular contexts that admit genuine experimental design	Understand	Distinguishes topics in which students can design real experiments from those in which design is constrained to recipe-following or thought-experiments.
T-A2.2.b	Students' Understanding	Anticipating typical student design errors	Analyze	Anticipate recurring patterns — confounded variables, missing controls, inconclusive manipulations, and sample-size confusion — and determine where each appears
T-A2.2.c	Instructional Strategies	Scaffolding student-led experimental design	Apply	Uses partial design templates, structured peer review, and 'design-then-revise' cycles to support students through design without supplying the design.
T-A2.2.d	Orientations to Teaching Science	Framing experimental design as a creative scientific act	Understand	Conveys that designing an experiment is part of doing science — a creative and consequential decision space —

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
				not the execution of a fixed procedure.
T-A2.2.e	Assessment Knowledge	Designing rubrics for design quality on internal logic, not whether it 'worked'	Create	Designs rubrics that reward principled design choices — variable control, alignment to question, justified sample — independently of whether the experiment yielded results.

Sub-ability A2.3 — Evaluating the fit of an experimental design to its question

Verbatim PISA: "Evaluate whether an experimental design is best suited to answer the question."
(Box 4, p. 21)

Table 39 Evaluating the fit of an experimental design to its question.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A2.3.a	Curricular Saliency	Identifying curricular contexts where design evaluation is productive	Understand	Identifies where in the curriculum competing or contested studies, replications, or methodological debates can be brought in for student evaluation.
T-A2.3.b	Students' Understanding	Anticipating familiarity-based design evaluation	Analyze	Recognises that students tend to favour designs that resemble ones they have seen, regardless of whether those designs fit the question being asked.
T-A2.3.c	Instructional Strategies	Modelling design-to-question fit checking	Apply	Demonstrates explicit move-by-move evaluation of a published design against the study's stated question, including identifying mismatches and trade-offs.
T-A2.3.d	Orientations to Teaching Science	Framing design critique as constructive scientific practice	Understand	Conveys that critique of designs — including published ones — is constitutive of how science self-corrects, and is not personal

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
				criticism of researchers.
T-A2.3.e	Assessment Knowledge	Designing assessments of principled critique of published designs	Create	Designs assessments where students critique a published design against its question, with rubrics rewarding identified trade-offs and grounded alternative proposals.

Sub-ability A2.4 — Interpreting data, drawing conclusions, and evaluating their relative merits

Verbatim PISA: "Interpret data presented in different representations, draw appropriate conclusions from data and evaluate their relative merits." (Box 4, p. 21)

Table 40 Interpreting data, drawing conclusions, and evaluating their relative merits.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A2.4.a	Curricular Saliency	Identifying curricular sites for genuine data interpretation	Understand	Distinguishes datasets and claims in the curriculum that admit real interpretation work from those that arrive pre-interpreted as a 'take-home message'.
T-A2.4.b	Students' Understanding	Diagnosing data-misreading patterns	Analyze	Anticipates correlation–causation confusion, scale-misreading, single-data-point inference, and disregard of uncertainty, and detects these in student work.
T-A2.4.c	Instructional Strategies	Teaching the data-to-claim-to-warrant chain	Apply	Makes the moves from data to claim, and from claim to its scope of warrant, explicit and practised — including what conclusions the data does and does not support.
T-A2.4.d	Orientations to Teaching Science	Framing conclusions as warranted, not stated	Understand	Conveys that scientific conclusions are claims earned by

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
				evidence with declared scope, not opinions or beliefs to be stated and defended afterwards.
T-A2.4.e	Assessment Knowledge	Designing assessments of defence of conclusions against alternatives	Create	Designs assessments where students must defend a conclusion against named alternatives, with rubrics rewarding the strength of the defence rather than the choice of conclusion.

Student Competency A3: Researching, evaluating and using scientific information for decision making and action

PISA Competency: A3 — Research, evaluate and use scientific information for decision making and action (Box 5, p. 23).

Sub-ability A3.1 — Searching, evaluating and communicating the merits of information sources

Verbatim PISA: "Search, evaluate and communicate the relative merits of different sources of information (scientific, social, economic and ethical) that may have significance or merit in arriving at decisions on science-related issues, and whether they support an argument or a solution." (Box 5, p. 23)

Table 41 Searching, evaluating and communicating the merits of information sources.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A3.1.a	Curricular Saliency	Identifying curricular sites for multi-source engagement	Understand	Knows which curricular topics genuinely require students to engage with multiple sources rather than a single textbook account — climate, health, AI, energy.
T-A3.1.b	Students' Understanding	Anticipating binary trust judgments in source evaluation	Analyze	Recognises that students typically treat sources as either trustworthy or untrustworthy, missing gradations, and conflate institutional authority with epistemic warrant.
T-A3.1.c	Instructional Strategies	Teaching source-evaluation moves explicitly	Apply	Models the named moves — provenance, primary versus secondary, peer review, authorship,

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
				recency, declared interests — rather than relying on tacit 'good judgment'.
T-A3.1.d	Orientations to Teaching Science	Framing source criticism as scientific practice	Understand	Conveys that source criticism is part of doing science, not a separate 'media literacy' skill bolted onto the science curriculum.
T-A3.1.e	Assessment Knowledge	Designing tasks for source comparison against principled criteria	Create	Designs tasks where students compare sources of differing quality on the same issue, with rubrics rewarding the criteria applied rather than the conclusion reached.

Sub-ability A3.2 — Distinguishing claims by type and warrant

Verbatim PISA: "Distinguish among claims based on strong scientific evidence, expert vs. non-expert, and opinion, and provide reasons for the distinction." (Box 5, p. 23)

Table 42 Distinguishing claims by type and warrant.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A3.2.a	Curricular Saliency	Identifying topics where claim-types are heterogeneous	Understand	Recognises which curricular topics carry a mix of evidence-based claims, expert judgment, and informed opinion, and which are largely homogeneous on this dimension.
T-A3.2.b	Students' Understanding	Diagnosing confidence-evidence conflation	Analyze	Anticipates that students judge claims by the speaker's confidence rather than by evidential warrant, and that expert disagreement is read as a license for opinion.
T-A3.2.c	Instructional Strategies	Modelling claim-type distinction explicitly	Apply	Teaches the typology — evidence-based claim, expert judgment, informed opinion, opinion — and demonstrates the tests that distinguish them in concrete examples.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A3.2.d	Orientations to Teaching Science	Framing expertise as accountable, not authoritative	Understand	Conveys that experts are correct because of evidence and method, not the other way round — credentials are necessary but not sufficient warrant.
T-A3.2.e	Assessment Knowledge	Designing assessments of claim classification with reasoning	Create	Designs assessments where students classify claims by type and justify the classification, with rubrics rewarding classification quality independent of agreement with the teacher.

Sub-ability A3.3 — Constructing scientific arguments from data

Verbatim PISA: "Construct an argument to support an appropriate scientific conclusion from a set of data." (Box 5, p. 23)

Table 43 Constructing scientific arguments from data.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A3.3.a	Curricular Saliency	Identifying curricular sites for data-driven argument construction	Understand	Knows which curricular topics naturally afford the construction of arguments from data — contested findings, environmental data, public-health figures — rather than data-as-illustration.
T-A3.3.b	Students' Understanding	Diagnosing missing-warrant patterns in student arguments	Analyze	Anticipates that students typically state a claim and present data but skip the warrant — the explicit reasoning that connects the data to the claim — and detects this in student work.
T-A3.3.c	Instructional Strategies	Scaffolding argument structure with claim, evidence, and reasoning	Apply	Uses an explicit argument schema (claim–evidence–reasoning, with qualifier) to scaffold student argument construction, including in oral and written form.
T-A3.3.d	Orientations to Teaching Science	Framing science as constitutively argumentative	Understand	Conveys that science advances through argument among scientists,

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
				and that data does not 'reveal' conclusions on its own — it is mobilised within arguments.
T-A3.3.e	Assessment Knowledge	Designing assessments of argument construction against structural criteria	Create	Designs assessments in which students construct, not just state, an argument from given data, with rubrics that assess argument structure rather than the choice of conclusion.

Sub-ability A3.4 — Critiquing standard flaws in science-related arguments

Verbatim PISA: "Critique standard flaws in science-related arguments using epistemic and procedural knowledge e.g., poor assumptions, cause vs. correlation, faulty explanations, generalisations from limited data." (Box 5, p. 23)

Table 44 Critique standard flaws in sciences-related arguments.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A3.4.a	Curricular Saliency	Identifying live debates as contexts for argument critique	Understand	Selects contemporary debates where flawed arguments are abundant and pedagogically productive — vaccines, climate, GM food, AI risk — and locates them within the curriculum.
T-A3.4.b	Students' Understanding	Diagnosing stance-contaminated flaw recognition	Analyze	Anticipates that students readily identify flaws in arguments they disagree with and overlook identical flaws in arguments they agree with, and surfaces the asymmetry.
T-A3.4.c	Instructional Strategies	Teaching a working repertoire of standard argumentative flaws	Apply	Teaches a named, practised repertoire — cherry-picking, false dichotomy, base-rate neglect, ad hominem, motivated scepticism — and the moves to detect each.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A3.4.d	Orientations to Teaching Science	Framing critique as required of all positions	Understand	Conveys that critique applies to all positions, including ones the student or the teacher agrees with — critique is not a tool for defeating opponents.
T-A3.4.e	Assessment Knowledge	Designing assessments of even-handed argument critique	Create	Designs assessments in which students critique arguments on both sides of an issue, with rubrics rewarding the consistent application of critique criteria across stances.

Sub-ability A3.5 — Justifying decisions using scientific arguments on contemporary issues

Verbatim PISA: *"Justify decisions using scientific arguments, either individual or communal, that contribute to solving contemporary issues or sustainable development."* (Box 5, p. 23)

Table 45 Justifying decisions using scientific arguments on contemporary issues.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-A3.5.a	Curricular Saliency	Identifying decision-relevant content in the curriculum	Understand	Identifies curricular content that genuinely informs personal, civic, or policy decisions — distinguishing it from content that is merely topical or contextual.
T-A3.5.b	Students' Understanding	Anticipating undifferentiated weighting in student decision reasoning	Analyze	Recognises that students treat science as one consideration among many in decisions but cannot articulate how it weighs against values, costs, and trade-offs.
T-A3.5.c	Instructional Strategies	Modelling the move from evidence to justified decision	Apply	Demonstrates how scientific evidence does and does not determine decisions in the presence of values and trade-offs, including the structure of a justified decision under uncertainty.
T-A3.5.d	Orientations to Teaching Science	Framing decision-making as a legitimate scientific object of analysis	Understand	Conveys that decisions on science-related

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
				issues are themselves analysable — students can be asked to justify decisions, not only to know facts.
T-A3.5.e	Assessment Knowledge	Designing assessments of decision-justification under explicit constraints	Create	Designs assessments where students justify a specific decision using scientific reasoning under explicit constraints (resources, time, values), with rubrics rewarding the structure of justification.

E-strand: Environmental Science Identity

PISA reference: Box 11 (p. 35), 'Environmental awareness, concern and agency'.

Component E-1 — Environmental Awareness

PISA reference: Box 11 (p. 35), awareness component.

TEACHER CAPABILITIES	PISA SCIENCE IDENTITY COMPONENT	PISA CONSTRUCT
E1 — Environmental Awareness	Environmental awareness	Science Identity (Box 11)

Table 46 Environmental awareness.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-E1.a	Curricular Saliency	Identifying curricular sites that build cumulative socio-ecological systems awareness	Understand	Knows where the curriculum can carry Earth-system content cumulatively across years and topics — atmosphere, hydrosphere, biosphere, geosphere and their feedbacks — rather than as a single chapter or a one-off awareness day.
T-E1.b	Students' Understanding	Anticipating linear-causal default reasoning about Earth systems	Analyze	Anticipates that students reach for single-cause linear chains when asked about Earth-system phenomena (climate, biodiversity loss), missing feedback, tipping points and leverage points, and detects these patterns in student work.
T-E1.c	Instructional Strategies	Teaching systems thinking with	Apply	Uses systems-thinking moves — diagrams of

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
		feedbacks, tipping points and leverage points		stocks and flows, identification of positive and negative feedback, scenario tracing — to help students reason about socio-ecological systems beyond linear cause-and-effect.
T-E1.d	Orientations to Teaching Science	Framing environmental phenomena as legitimate scientific content	Understand	Conveys that environmental science is part of the curriculum's scientific work — subject to the same scrutiny and evidential standards — and not a 'soft' or extracurricular topic.

Component E-2 — Environmental Concern

PISA reference: Box 11 (p. 35).

TEACHER CAPABILITIES	PISA SCIENCE IDENTITY COMPONENT	PISA CONSTRUCT
E2 — Environmental Concern	Environmental concern	Science Identity (Box 11)

Table 47 Environmental concern.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-E2.a	Curricular Saliency	Connecting environmental phenomena to students' lived experience	Understand	Identifies points in the curriculum where environmental issues can be grounded in students' immediate experience — local data, fieldwork, comparison with familiar contexts — rather than treated as distant abstractions.
T-E2.b	Students' Understanding	Diagnosing eco-anxiety and disengagement responses to environmental content	Analyze	Recognises that students encountering environmental content commonly respond with paralysing anxiety or disengaged fatalism, distinguishes these from constructive concern, and surfaces the distinction productively rather than dismissing the affective response.
T-E2.c	Instructional Strategies	Cultivating evidence-based hope through documented success cases	Apply	Integrates concrete documented cases of environmental success — ecosystem recovery, energy transition, effective policy — to build students' sense

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
				that change is possible, distinguishing evidence-based hope from naive optimism.
T-E2.d	Orientations to Teaching Science	Framing concern as an evidence-supported response, not a curriculum requirement	Understand	Conveys that concern about environmental outcomes is a justified response to evidence about those outcomes, distinct from a stance the teacher imposes on the class.

Component E-3 — Environmental Agency

PISA reference: Box 11 (p. 35).

TEACHER CAPABILITIES	PISA SCIENCE IDENTITY COMPONENT	PISA CONSTRUCT
E3 — Environmental Agency	Environmental agency	Science Identity (Box 11)

Table 48 Environmental agency.

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
T-E3.a	Curricular Saliency	Identifying decision-relevant environmental content with equity dimensions	Understand	Identifies curricular content that genuinely informs environmental decisions and surfaces the distributional dimension — who is most affected, who has the least voice — rather than treating environmental issues as evenly distributed.
T-E3.b	Students' Understanding	Anticipating individualisation of structural environmental problems	Analyze	Recognises that students typically reduce environmental crises to individual consumer choices (carbon footprint as a personal moral question), missing the structural and collective dimensions, and detects this pattern in student reasoning.
T-E3.c	Instructional Strategies	Teaching the analysis of action at different scales	Apply	Equips students to analyse where action operates — personal, civic, collective, structural — and to distinguish tractable from intractable points

ID	PCK anchor (RCM)	Teacher capability (gerund title)	Bloom	One-line gloss
				of leverage, rather than treating 'doing something' as a single category.
T-E3.d	Orientations to Teaching Science	Framing scientific knowledge as informing but not determining environmental decisions	Understand	Conveys that science informs environmental decisions but does not resolve them: decisions also involve values, distributional judgements, and plural knowledge systems (scientific, indigenous, local), all of which are legitimate objects of analysis.
T-E3.e	Assessment Knowledge	Designing authentic impact-analysis tasks with structural framing	Create	Designs tasks (carbon footprint, life-cycle analysis, dietary assessment) that require students to integrate scientific evidence with structural and ethical analysis, with rubrics rewarding the quality of the action plan at multiple scales rather than the ambition of the goal.

M1 — Mathematical Reasoning

PISA 2022 Maths: deductive and inductive reasoning to evaluate situations, select strategies, draw conclusions, and develop arguments.

Table 49 Mathematical reasoning.

ID	PCK anchor	Capability (gerund title)	Bloom	One-line gloss
T-M1.b	Students' Understanding	Diagnosing authority-based reasoning in mathematical claims	Analyze	Recognises when students accept mathematical claims by appeal to authority (textbook, teacher, internet) rather than by deductive or inductive reasoning, and surfaces this productively.
T-M1.c	Instructional Strategies	Teaching deductive and inductive reasoning as distinct moves	Apply	Models the difference between deductive certainty and inductive support; designs tasks where students use each appropriately rather than conflating them.
T-M1.d	Orientations to Teaching Science	Framing mathematical reasoning as protection from manipulation	Understand	Conveys that reasoning rigour is what allows students to defend themselves against quantitative misinformation, not an academic exercise.

M2 — Formulating Situations Mathematically

PISA 2022 Maths: recognising opportunities to use mathematics, providing structure to contextualised problems, identifying significant variables and constraints.

Table 50 Formulating situations mathematically.

ID	PCK anchor	Capability (gerund title)	Bloom	One-line gloss
T-M2.a	Curricular Saliency	Identifying real-world contexts that admit genuine mathematical formulation	Understand	Distinguishes contexts where mathematical formulation is the substantive cognitive work from contexts where it is decorative wrapping on a procedural problem.
T-M2.c	Instructional Strategies	Scaffolding the move from context to mathematical structure	Apply	Demonstrates the moves of selecting variables, identifying assumptions, recognising patterns, and translating context-specific language into mathematical form.
T-M2.d	Orientations to Teaching Science	Framing modelling as a creative act under simplifying assumptions	Understand	Conveys that mathematical modelling is choice-laden — what to include and what to ignore — rather than a single right structure waiting to be found.

M3 — Employing Mathematical Concepts, Facts, Procedures and Reasoning

PISA 2022 Maths: performing the procedures and manipulations needed to derive mathematical solutions. Lightly populated — procedural mathematics is conventional teacher territory.

Table 51 Employing mathematical concepts, facts, procedures and reasoning.

ID	PCK anchor	Capability (gerund title)	Bloom	One-line gloss
T-M3.c	Instructional Strategies	Teaching switching between representations during procedural work	Apply	Develops fluency in moving between symbolic, graphical, tabular, and verbal representations as part of solution work, not merely at the end of it.
T-M3.d	Orientations to Teaching Science	Framing procedure as means to mathematical insight, not end in itself	Understand	Conveys that procedural fluency exists in service of reasoning and modelling, and that procedure without reasoning is not yet mathematics.

M4 — Interpreting, Applying and Evaluating Mathematical Outcomes

PISA 2022 Maths: reflecting on solutions, evaluating reasonableness in context, critiquing the limits of the model. Heavily dual-traced to A2.4 (Interpreting data); items here are net-new where the mathematical-context dimension is genuinely additive.

Table 52 Interpreting, applying and evaluating mathematical outcomes.

ID	PCK anchor	Capability (gerund title)	Bloom	One-line gloss
T-M4.b	Students' Understanding	Anticipating context-detachment in student conclusions	Analyze	Recognises when students arrive at correct mathematical answers but fail to translate them back to the real-world context, and surfaces the missing translation.
T-M4.c	Instructional Strategies	Teaching the critique of model fitness against context	Apply	Designs instruction in which students evaluate whether a mathematical model captures what matters in the real situation — not only whether the calculation is correct.
T-M4.e	Assessment Knowledge	Designing assessments that reward reasonableness checking	Create	Designs assessments where the rubric rewards students for testing answer reasonableness against context, not only for procedural correctness.

AI-strand — AILit Framework

Covers only the intersection of AI literacy and scientific reasoning. Broader AI literacy is addressed by dedicated AI Teacher Academies projects, to which ImSTEAM defers.

AI1 — Engaging with AI (in scientific reasoning contexts)

AILit: using AI as a tool to access content, information, or recommendations. ImSTEAM scope: where AI mediates scientific information environments and student-facing scientific claims.

Table 53 Engaging with AI.

ID	PCK anchor	Capability (gerund title)	Bloom	One-line gloss
T-AI1.b	Students' Understanding	Anticipating uncritical acceptance of AI-generated scientific claims	Analyze	Recognises that students typically treat AI-generated scientific explanations as authoritative without distinguishing them from peer-reviewed sources, and surfaces the distinction productively.
T-AI1.c	Instructional Strategies	Teaching evaluation of AI-generated scientific outputs	Apply	Develops in students the habit of cross-referencing AI-generated scientific claims against primary sources, identifying confabulation, and recognising AI's training-data limitations.
T-AI1.d	Orientations to Teaching Science	Framing AI as a generator of plausible-sounding claims, not a source of warranted knowledge	Understand	Conveys that AI outputs are predictive completions rather than epistemic claims — a distinction students must hold even when outputs appear scientifically credible.

ID	PCK anchor	Capability (gerund title)	Bloom	One-line gloss
T-AI1.e	Assessment Knowledge	Designing assessments of AI-output critique against scientific criteria	Create	Designs assessments where students evaluate AI-generated scientific content against named criteria (source warranting, scope of training, identifiable failure modes), not against general plausibility.



AI3 — Managing AI (selectively, in scientific work)

AILit: intentionally choosing how AI supports human work, including task decomposition and delegation. ImSTEAM scope: helping students decide when AI is and is not an appropriate partner in scientific reasoning tasks.

Table 54 Managing AI.

ID	PCK anchor	Capability (gerund title)	Bloom	One-line gloss
T-AI3.c	Instructional Strategies	Teaching the decomposition of scientific tasks into AI-suitable and human-required parts	Apply	Develops in students the habit of distinguishing parts of a scientific task that AI can usefully assist (literature scanning, calculation, draft generation) from parts that require human judgement (warranting, ethical evaluation, design choices).
T-AI3.d	Orientations to Teaching Science	Framing AI use in scientific work as a deliberate choice with consequences	Understand	Conveys that whether to use AI in a given scientific task is itself a methodological decision affecting the warrant of the result, not a frictionless productivity move.

21st-century skills — A-strand absorption and integration map

PISA 2022 Mathematics names eight 21st-century skills woven into mathematical literacy. Six of the eight are operationally covered by existing A-strand capabilities and are dual-traced to PISA Maths' 21C inventory at Layer 1 rather than authored as separate items. Two — Creativity and Self-direction, initiative, and persistence — are genuinely uncovered and absorbed into the A-strand as new capabilities below.

New A-strand capabilities

Table 55 New A-strand capabilities.

ID	PCK anchor	Capability (gerund title)	Bloom	One-line gloss
T-A1.5.f*	Instructional Strategies	Eliciting divergent generation before convergent selection in scientific work	Apply	Designs hypothesising activities where students first generate multiple candidate explanations without filtering, then evaluate against criteria — protecting the creative phase from premature convergence. New capability absorbing PISA Maths 21C: Creativity.
T-A2.2.f*	Orientations to Teaching Science	Framing scientific learning as requiring persistence through productive failure	Understand	Conveys that scientific work involves sustained effort across confusion, dead ends, and revision; cultivates student self-direction by treating productive failure as part of the work rather than as a personal deficit. New capability absorbing PISA Maths 21C: Self-direction, initiative, and persistence.

Integration map (six 21C skills covered through dual-tracing)

Each row names a 21C skill from the PISA Maths inventory, identifies where it is operationally covered in the A-strand, and indicates the existing capabilities to which a Layer-1 secondary citation to PISA Maths 21C should be added.

Table 56 Integration map.

21C skill	Where covered in A-strand	Capabilities to dual-trace	Predominant PCK anchor
Critical thinking	Pervasive across A3 (source evaluation, claim distinction, argument critique) and Orientations across all sub-abilities	T-A3.1.c, T-A3.2.c, T-A3.4.c; A1.x.d Orientations broadly	Instructional Strategies + Orientations
Research and inquiry	Constitutive of A2 (research-question identification, experimental design) and A3.1 (multi-source engagement)	T-A2.1.c, T-A2.2.c, T-A3.1.c	Instructional Strategies
Information use	Constitutive of A3.1 (searching, evaluating, communicating sources)	T-A3.1.c, T-A3.1.e	Instructional Strategies + Assessment
Systems thinking	Primarily in E-strand (E3 environmental agency, systems mapping); selectively in A1.6 societal implications	T-E3.b (after augmentation); T-A1.6.c	Instructional Strategies
Communication	A1.2 (translating between representations) and A3.3 (constructing arguments)	T-A1.2.c, T-A3.3.c	Instructional Strategies
Reflection	A1.3 (prediction-justification-revision cycle); A1.5 (hypothesis revision); Assessment Knowledge anchors broadly	T-A1.3.c, T-A1.5.c, A1.x.e Assessment broadly	Instructional Strategies + Assessment



Annex II The ImSTEAM Teachers Capability framework

About this framework

This framework sets out the teacher capabilities that ImSTEAM develops, expressed entirely from the teacher's side: what a teacher must be able to do to develop a given scientific capability in students. It is a standalone reference, derived from the **PISA 2025 Science Framework**. It covers the science competencies — explaining phenomena, scientific enquiry and data, and evidence, argument and decision — together with the environmental science-identity dimension.

How to read an entry

The framework is organised in three levels: category, capability and teacher skill. Each skill is presented in a four-column row:

- **N_e** — the position of the category, capability and skill (for example 2.1.3: category 2, capability 1, skill 3). Numbering is sequential throughout.
- **Teacher skill** — the named skill and a short description of what enacting it involves.
- **Teaching focus** — which part of teaching the skill addresses: what to teach, where students struggle, how to teach it, how to frame it, or how to assess it.
- **Objective** — a draft learning objective written at the skill's cognitive level, with that level shown in brackets, ready to seed course objectives.

At category and capability level, an **anchoring reference** states the PISA 2025 competency or construct the item derives from.

Teacher skills and the five teaching dimensions

Each capability comprises up to five teacher skills, one for each dimension of the teaching it requires. Read together, the five form a complete teaching arc:

- **What to teach** — the core content the capability rests on.
- **Where students struggle** — the misconceptions and difficulties students typically bring.
- **How to teach it** — the instructional moves that develop the capability.
- **How to frame it** — the stance toward science the teacher conveys.
- **How to assess it** — how mastery is evidenced and judged.

A capability covered across all five dimensions equips a teacher with complete command of it. Where a capability shows fewer than five skills, the absent dimension is a deliberate scope decision



rather than an omission — most visibly in the environmental capabilities of awareness and concern, which carry no dedicated assessment skill.

Scope and notation

- The framework comprises **four categories, eleven capabilities and forty-nine teacher skills**.
- Anchoring references use the form “Competency n” for the three PISA 2025 science competencies (Boxes 3–5) and “Box 11” for the environmental awareness–concern–agency construct.
- A cross-reference table at the end maps every framework number to its identifier in the underlying capability pool.

Categories

Table 57 Categories, anchoring reference, capabilities and skills.

No	Category	Anchoring reference	Capabilities	Skills
1	Explaining phenomena	PISA 2025 · Competency 1 (Box 3)	3	15
2	Scientific enquiry and data	PISA 2025 · Competency 2 (Box 4)	2	10
3	Evidence, argument and decision	PISA 2025 · Competency 3 (Box 5)	3	15
4	Environmental science identity	PISA 2025 · Box 11	3	9

Capabilities

Table 58 Capabilities, anchoring reference and skills.

No	Capability	Anchoring reference	Skills	Evidence of Mastery
1.1	Developing students' fluency across scientific representations	C1 · representations (Box 3)	5	Explains scientific concepts accurately using age-appropriate language. • Selects authentic examples that connect scientific knowledge to real-world contexts. • Identifies and addresses common student misconceptions. • Designs learning activities that require students to apply scientific knowledge in unfamiliar situations.

Nº	Capability	Anchoring reference	Skills	Evidence of Mastery
1.2	Developing students' scientific prediction and justification	C1 · predictions (Box 3)	5	Uses multiple forms of scientific representation (graphs, diagrams, models, simulations). • Guides students in translating between different representations. • Selects representations appropriate to learners' needs. • Assesses students' ability to interpret and construct scientific models.
1.3	Developing students' work with scientific models	C1 · models (Box 3)	5	Designs learning tasks requiring prediction, explanation and justification. • Encourages evidence-based reasoning during classroom discussion. • Supports students in constructing coherent scientific explanations. • Evaluates the quality of students' scientific reasoning using explicit criteria.
2.1	Developing students' experimental-design reasoning	C2 · experimental design (Box 4)	5	Designs inquiry-based learning activities aligned with learning objectives. • Formulates investigable scientific questions. • Identifies appropriate variables, controls and methods. • Justifies methodological choices according to the inquiry goals.
2.2	Developing students' data interpretation and conclusion-drawing	C2 · data & conclusions (Box 4)	5	Selects appropriate methods for data collection and analysis. • Supports students in interpreting experimental evidence. • Identifies limitations and possible sources of error. • Guides students in drawing evidence-based conclusions.
3.1	Developing students' evaluation of information sources	C3 · source evaluation (Box 5)	5	Selects reliable scientific information from multiple sources. • Supports students in distinguishing evidence from opinion.

Nº	Capability	Anchoring reference	Skills	Evidence of Mastery
				• Critically evaluates the credibility of scientific claims. • Promotes informed decision-making based on evidence.
3.2	Developing students' construction of scientific arguments	C3 · argument from data (Box 5)	5	Facilitates structured scientific argumentation in classroom activities. • Guides students in connecting claims, evidence and reasoning. • Encourages critical evaluation of alternative explanations. • Assesses the quality of students' scientific arguments using transparent criteria.
3.3	Developing students' science-based decision-making	C3 · decision-making (Box 5)	5	Designs authentic learning scenarios involving socio-scientific issues. • Supports balanced evaluation of scientific, ethical and societal dimensions. • Facilitates evidence-informed classroom discussions. • Evaluates students' ability to justify decisions using scientific evidence.
4.1	Building students' environmental awareness	Box 11 · awareness	3	Integrates environmental concepts into science learning activities. • Uses authentic environmental case studies and local examples. • Promotes systems thinking in analysing environmental issues. • Assesses students' understanding of environmental interrelationships.
4.2	Developing students' evidence-based environmental concern	Box 11 · concern	3	Encourages reflection on environmental responsibility. • Designs activities that promote informed attitudes towards sustainability. • Facilitates discussion on environmental challenges from multiple perspectives.

Nº	Capability	Anchoring reference	Skills	Evidence of Mastery
				• Evaluates students' ability to justify environmentally responsible positions.
4.3	Developing students' environmental agency	Box 11 · agency	3	Designs action-oriented learning activities addressing sustainability challenges. • Supports collaborative problem-solving for environmental issues. • Encourages students to plan and implement evidence-based environmental actions. • Evaluates students' capacity to reflect on the impact of their decisions and actions.



Explaining phenomena

Anchored in PISA 2025 Science · Competency 1 — explain phenomena scientifically (Box 3).

Developing students' fluency across scientific representations

PISA 2025 · Competency 1: using and translating between forms of representation (Box 3)

Capability. Develop students' ability to use and move between the representations that carry scientific meaning — text, diagrams, graphs and models.

Table 59 PISA 2025 Competency 1: using and translating between forms of representation.

No	Teacher skill	Teaching focus	Objective
1.1.1	Identifying the representations that carry the topic Identifies the written, diagrammatic, graphical and symbolic forms a topic depends on, drawing on real-life contexts students already recognise.	What to teach	[Understand] identify the representations that carry a topic
1.1.2	Diagnosing representational fluency gaps Detects where students can read one representation but cannot move between forms, and adapts the level of teaching to the gap observed.	Where students struggle	[Analyse] diagnose gaps in students' representational fluency
1.1.3	Constructing tasks that require translation between representations Constructs tasks that require students to produce one representation from another rather than interpret one in isolation, including designing their own forms and exchanging them between groups.	How to teach it	[Apply] set tasks requiring translation between representations
1.1.4	Framing representations as thinking tools in their own right Conveys that a representation is a tool for reasoning about a phenomenon, not a decorative illustration of it.	How to frame it	[Understand] frame representations as thinking tools in their own right
1.1.5	Designing assessments of genuine translation fluency Designs assessments that test genuine translation between forms rather than recognition, differentiating the assessment level and guiding student self-assessment and reflection.	How to assess it	[Create] design assessments of genuine translation fluency

Developing students' scientific prediction and justification

PISA 2025 · Competency 1: making and justifying scientific predictions (Box 3)

Capability. Develop students' ability to make scientific predictions and justify them, treating prediction as a core scientific move.

Table 60 PISA 2025: Competency 1: making and justifying scientific predictions.

No	Teacher skill	Teaching focus	Objective
1.2.1	Identifying topics where prediction is the central scientific move Identifies the topics in which predicting and justifying an outcome is the substantive scientific work rather than an add-on.	What to teach	[Understand] identify topics where prediction is the central scientific move
1.2.2	Anticipating where students mistake plausibility for justification Anticipates that students treat a plausible-sounding guess as a justified prediction, and helps them weigh evidence and data against hearsay.	Where students struggle	[Analyse] anticipate where students mistake plausibility for justification
1.2.3	Scaffolding the prediction–justification–revision cycle Scaffolds the cycle of predicting, justifying and revising using a framework or template and formats such as brainstorming, case study and structured debate.	How to teach it	[Apply] scaffold the prediction–justification–revision cycle
1.2.4	Framing prediction as constitutive of scientific work Conveys that making and testing predictions is part of how science proceeds, not a classroom exercise detached from it.	How to frame it	[Understand] frame prediction as constitutive of scientific work
1.2.5	Designing rubrics that reward the quality of justification Designs rubrics that reward the quality of a student's justification and reasoning, independently of whether the prediction turned out correct.	How to assess it	[Create] design rubrics rewarding the quality of justification

Developing students' work with scientific models

PISA 2025 · Competency 1: identifying, constructing and evaluating models (Box 3)

Capability. Develop students' ability to build, test and evaluate scientific models as provisional, purpose-built tools.

Table 61 PISA 2025, Competency 1: identifying, constructing and evaluating models.

No	Teacher skill	Teaching focus	Objective
1.3.1	Identifying foundational models for a topic Identifies the foundational models a topic rests on and recognises when each should be introduced and used.	What to teach	[Understand] identify the foundational models for a topic
1.3.2	Anticipating student confusion between model and reality Anticipates that students treat a model as a literal description of reality, and detects this confusion in their work.	Where students struggle	[Analyse] anticipate confusion between model and reality
1.3.3	Sequencing modelling — building, testing, revising Sequences modelling as building, testing and revising, constructing scenarios in which students make predictions with a model and settings in which they debate and explain competing solutions.	How to teach it	[Apply] sequence modelling as building, testing and revising
1.3.4	Framing models as provisional and purpose-built Conveys that models are provisional, purpose-built tools to be judged by their usefulness rather than their literal truth.	How to frame it	[Understand] frame models as provisional and purpose-built
1.3.5	Designing assessments of model evaluation against fitness-for-purpose criteria Designs assessments in which students judge a model against fitness-for-purpose criteria rather than against a single right answer.	How to assess it	[Create] design assessments of model evaluation against fitness-for-purpose

Scientific enquiry and data

Anchored in PISA 2025 Science · Competency 2 — construct and evaluate designs for scientific enquiry and interpret data critically (Box 4).

Developing students' experimental-design reasoning

PISA 2025 · Competency 2: proposing an appropriate experimental design (Box 4)

Capability. Develop students' ability to design sound scientific investigations and reason about what a design can and cannot show.

Table 62 PISA 2020, Competency 2: proposing an appropriate experimental design.

No	Teacher skill	Teaching focus	Objective
2.1.1	Identifying curricular contexts that admit genuine experimental design Distinguishes topics where students can design real experiments from those limited to recipe-following or thought-experiment, weighing factors such as student age and identifying the contextual components that make genuine design possible.	What to teach	[Understand] identify contexts that admit genuine experimental design
2.1.2	Anticipating typical student design errors Anticipates the recurring patterns — confounded variables, missing controls, manipulations that cannot answer the question, sample-size confusion — and recognises where each appears in student work.	Where students struggle	[Analyse] anticipate typical student design errors
2.1.3	Scaffolding student-led experimental design Supports students through design without supplying it — partial templates, structured peer review and design-then-revise cycles — coaching students to pose their own questions in a fail-friendly space and deconstructing the phases of design.	How to teach it	[Apply] scaffold student-led experimental design
2.1.4	Framing experimental design as a creative scientific act Conveys that designing an experiment is a creative, consequential scientific act — inviting students to design a problem of their own — rather than the execution of a fixed procedure.	How to frame it	[Understand] frame experimental design as a creative scientific act
2.1.5	Designing rubrics that judge design quality by its internal logic	How to assess it	[Create] create rubrics judging a design by its internal logic

No	Teacher skill	Teaching focus	Objective
	Designs rubrics that reward principled choices — control of variables, alignment of the design to the question of what and how to measure, credibility, and the capacity to falsify a prediction — independently of whether the experiment yielded results, using counter-examples to test design logic.		

Developing students' data interpretation and conclusion-drawing

PISA 2025 · Competency 2: interpreting data, drawing and evaluating conclusions (Box 4)

Capability. Develop students' ability to interpret data and draw conclusions that are warranted by the evidence.

Table 63 PISA 2025, Competency 2: interpreting data, drawing and evaluating conclusions.

No	Teacher skill	Teaching focus	Objective
2.2.1	Identifying curricular sites for genuine data interpretation Identifies the points in the curriculum where students interpret real data rather than confirm a known result.	What to teach	[Understand] identify sites for genuine data interpretation
2.2.2	Diagnosing data-misreading patterns Diagnoses the recurring ways students misread data, including confusion of correlation with causation, and helps them recognise genuine patterns.	Where students struggle	[Analyse] diagnose data-misreading patterns
2.2.3	Teaching the data-to-claim-to-warrant chain Teaches the chain from data to claim to warrant, having students categorise and validate data, relate conclusions to the evidence, and judge whether a conclusion is warranted.	How to teach it	[Apply] teach the data-to-claim-to-warrant chain
2.2.4	Framing conclusions as claims warranted by evidence Conveys that a conclusion is a claim that must be warranted by evidence and defended, not simply stated.	How to frame it	[Understand] frame conclusions as claims warranted by evidence
2.2.5	Designing assessments of how students defend conclusions against alternatives Designs assessments in which students defend their conclusions against alternatives, including counter-factual arguments.	How to assess it	[Create] design assessments of conclusions defended against alternatives

Evidence, argument and decision

Anchored in PISA 2025 Science · Competency 3 — research, evaluate and use scientific information for decision making and action (Box 5).

Developing students' evaluation of information sources

PISA 2025 · Competency 3: searching, evaluating and communicating the merits of sources (Box 5)

Capability. Develop students' ability to find, compare and judge the merits of competing information sources.

Table 64 PISA 2025, Competency 3: searching, evaluating and communicating the merits of sources.

No	Teacher skill	Teaching focus	Objective
3.1.1	Identifying curricular sites for multi-source engagement Identifies the curricular sites where students compare reliable and relevant sources rather than work from a single given text.	What to teach	[Understand] identify sites for multi-source engagement
3.1.2	Anticipating binary trust judgements in source evaluation Anticipates that students treat sources as simply trustworthy or not, missing the graded judgement that evaluation requires.	Where students struggle	[Analyse] anticipate binary trust judgements about sources
3.1.3	Teaching source-evaluation moves explicitly Teaches source-evaluation moves explicitly — weighing origin, audience, purpose, currency and possible conflicts of interest — and explaining what makes a source credible.	How to teach it	[Apply] teach explicit source-evaluation moves
3.1.4	Framing source criticism as scientific practice Conveys that critically evaluating sources is part of scientific practice, not an obstacle to it.	How to frame it	[Understand] frame source criticism as scientific practice
3.1.5	Designing tasks for source comparison against principled criteria Designs tasks in which students compare sources against principled criteria, including rubrics co-created with students for classroom use.	How to assess it	[Create] design source-comparison tasks against principled criteria

Developing students' construction of scientific arguments

PISA 2025 · Competency 3: constructing scientific arguments from data (Box 5)

Capability. Develop students' ability to build scientific arguments that connect claim, evidence and reasoning.

Table 65 PISA 2025, Competency 3: constructing scientific arguments from data.

No	Teacher skill	Teaching focus	Objective
3.2.1	Identifying curricular sites for data-driven argument construction Identifies the curricular sites where students build arguments from data rather than restate established conclusions.	What to teach	[Understand] identify sources for data-driven arguments
3.2.2	Diagnosing missing-warrant patterns in student arguments Diagnoses where student arguments jump from evidence to claim without a warrant, checking them against common faulty logic such as circular reasoning and cherry-picking.	Where students struggle	[Analyse] diagnose missing warrants in student arguments
3.2.3	Scaffolding argument structure with claim, evidence and reasoning Scaffolds argument structure — claim, evidence and reasoning — using schemes pitched to different levels and matched to the subject.	How to teach it	[Apply] scaffold argument structure (claim, evidence, reasoning)
3.2.4	Framing science as constitutively argumentative Conveys that science is constitutively argumentative, so that the same dataset can support more than one valid argument.	How to frame it	[Understand] frame science as constitutively argumentative
3.2.5	Designing assessments of argument construction against structural criteria Designs assessments that judge a student's argument against structural criteria — the presence and quality of claim, evidence and reasoning.	How to assess it	[Create] design assessments of argument construction

Developing students' science-based decision-making

PISA 2025 · Competency 3: justifying decisions on contemporary issues (Box 5)

Capability. Develop students' ability to justify decisions on contemporary issues using scientific argument.

Table 66 PISA 2025, Competency 3: justifying decisions on contemporary issues.

No	Teacher skill	Teaching focus	Objective
3.3.1	Identifying decision-relevant content in the curriculum Identifies the curricular content that genuinely informs decisions, including contemporary issues and policies where data bears on real behaviour.	What to teach	[Understand] identify decision-relevant curricular content
3.3.2	Anticipating undifferentiated weighting in student decision reasoning Anticipates that students weigh all considerations equally, and helps them lay out the competing constraints that bear on a decision.	Where students struggle	[Analyse] anticipate undifferentiated weighting in student decisions
3.3.3	Modelling the move from evidence to justified decision Models the move from evidence to justified decision, connecting data to real behaviour and choices in the presence of values and trade-offs.	How to teach it	[Apply] model the move from evidence to justified decision
3.3.4	Framing decision-making as a legitimate scientific object of analysis Conveys that how a decision is justified is itself a legitimate object of scientific analysis, surfacing the competing scientific perspectives on an issue.	How to frame it	[Understand] frame decision-making as a scientific object of analysis
3.3.5	Designing assessments of decision-justification under explicit constraints Designs assessments in which students justify a decision under explicit constraints rather than assert a preference.	How to assess it	[Create] design assessments of decision-justification under constraints

Environmental science identity

Anchored in PISA 2025 Science · Box 11 — environmental awareness, concern and agency.

Building students' environmental awareness

PISA 2025 · Box 11: environmental awareness

Capability. Develop students' cumulative, systems-level awareness of environmental phenomena as genuine scientific content.

Table 67 PISA 2025, environmental awareness.

No	Teacher skill	Teaching focus	Objective
4.1.1	Identifying curricular sites that build cumulative environmental awareness Knows where the curriculum can carry Earth-system content cumulatively across years and topics — atmosphere, hydrosphere, biosphere, geosphere and their feedbacks — rather than as a single chapter or a one-off awareness day.	What to teach	[Understand] identify sites that build cumulative environmental awareness
4.1.2	Anticipating linear-causal default reasoning about Earth systems Anticipates that students reach for single-cause linear chains when asked about Earth-system phenomena, missing feedbacks, tipping points and leverage points, and detects this in student work.	Where students struggle	[Analyse] anticipate single-cause reasoning about Earth systems
4.1.3	Framing environmental phenomena as legitimate scientific content Conveys that environmental science is part of the curriculum's scientific work, held to the same scrutiny and evidential standards as any other scientific topic.	How to frame it	[Understand] frame environmental phenomena as rigorous scientific content

Developing students' evidence-based environmental concern

PISA 2025 · Box 11: environmental concern

Capability. Develop students' evidence-based concern for the environment, distinguishing constructive concern from anxiety or fatalism.

Table 68 PISA 2025, environmental concern.

No	Teacher skill	Teaching focus	Objective
4.2.1	Connecting environmental phenomena to students' lived experience Identifies points in the curriculum where environmental issues can be grounded in students' immediate experience — local data, fieldwork and comparison with familiar contexts.	What to teach	[Understand] ground environmental issues in students' lived experience
4.2.2	Diagnosing eco-anxiety and disengagement responses to environmental content Recognises that students encountering environmental content commonly respond with paralysing anxiety or disengaged fatalism, distinguishes these from constructive concern, and surfaces the distinction productively.	Where students struggle	[Analyse] distinguish eco-anxiety and fatalism from constructive concern
4.2.3	Cultivating evidence-based hope through documented success cases Integrates concrete documented cases of environmental success — ecosystem recovery, energy transition, effective policy — to build students' sense that change is possible, distinguishing evidence-based hope from naive optimism.	How to teach it	[Apply] build evidence-based hope through documented success cases

Developing students' environmental agency

PISA 2025 · Box 11: environmental agency

Capability. Develop students' sense of environmental agency, analysing action across personal, civic and structural scales.

Table 69 PISA 2025, environmental agency.

No	Teacher skill	Teaching focus	Objective
4.3.1	Anticipating individualisation of structural environmental problems Recognises that students reduce environmental crises to individual consumer choices, missing the structural and collective dimensions, and detects this pattern in their reasoning.	Where students struggle	[Analyse] detect individualisation of structural environmental problems
4.3.2	Teaching the analysis of action at different scales Equips students to analyse where action operates — personal, civic, collective, structural — and to distinguish tractable from intractable points of leverage.	How to teach it	[Apply] teach analysis of action across scales
4.3.3	Designing authentic impact-analysis tasks with structural framing Designs tasks such as carbon-footprint or life-cycle analysis that require students to integrate scientific evidence with structural and ethical analysis, with rubrics rewarding the quality of the action plan at multiple scales.	How to assess it	[Create] design impact-analysis tasks with structural framing

Annex III AI Skills Supplement

This annex extends the ImSTEAM Teacher Capability Framework (Annex II) to the intersection of AI literacy and scientific reasoning. It follows the same conventions as the main framework — categories of teacher skills, each with a teaching focus and a draft objective at its cognitive level — and describes what a teacher must be able to do to help students engage with and manage AI in scientific work.

This is a forward-looking, research-layer extension, not part of the project pilot. AI literacy is not a competency the project derives or delivers as a target; it is retained here for completeness, for follow-up funding, and for the ImSTEAM Alliance. The capabilities below have not been through the consortium prioritisation, so they carry no workshop-derived detail.

The annex covers only two of the four AI-literacy domains — engaging with AI and managing AI — and only at their intersection with scientific reasoning, modelling and the evaluation of evidence and claims. Creating with AI and designing AI are addressed by dedicated AI-focused Erasmus+ Teacher Academies, to which ImSTEAM defers. Within the two domains, the teaching dimensions are selectively populated, so some capabilities show fewer than five skills by design.



Table 70 Status within ImSTEAM.

ITEM	STATUS
Pilot implementation	NO
Research output	YES
Future exploitation	YES
Mandatory training	NO

Anchoring and references

This annex is anchored in the **AILit Framework** (OECD / European Commission). Because that framework is a preliminary review draft, references are given at **domain level only** (engaging with AI, managing AI) rather than to specific sub-competences, which may change as the framework is finalised.

As in Annex II, each skill is presented in a four-column row:

- **No** — sequential position of the capability and skill (for example 1.3: capability 1, skill 3).
- **Teacher skill** — the named skill and a short description of what enacting it involves.
- **Teaching focus** — which part of teaching the skill addresses: where students struggle, how to teach it, how to frame it, or how to assess it.
- **Objective** — a draft learning objective at the skill's cognitive level, with that level shown in brackets.

The five-dimension teaching arc described in Annex I applies here too; where a capability shows fewer than five skills, the absent dimensions are intentional, reflecting the framework's deliberately narrow scope rather than an omission.

Table 71 AI capability and anchoring reference.

No	Capability	Anchoring reference	Skills
1	Developing students' critical engagement with AI in science	AILit · Engaging with AI	4
2	Developing students' management of AI in scientific work	AILit · Managing AI	2

Capabilities

Anchored in the ALLit Framework (OECD/EC, review draft) — engaging with AI and managing AI, at their intersection with scientific reasoning.

Developing students' critical engagement with AI in science

ALLit Framework · Engaging with AI (domain level)

Capability. Develop students' critical engagement with AI in the scientific information environments they increasingly encounter.

Table 72 ALLit Framework, engaging with AI (domain level).

Nº	Teacher skill	Teaching focus	Objective
1.1	Anticipating uncritical acceptance of AI-generated scientific claims Recognises that students treat AI-generated scientific explanations as authoritative without distinguishing them from peer-reviewed sources, and surfaces the distinction productively.	Where students struggle	[Analyse] anticipate uncritical acceptance of AI-generated scientific claims
1.2	Teaching evaluation of AI-generated scientific outputs Develops in students the habit of cross-referencing AI-generated scientific claims against primary sources, identifying confabulation, and recognising the limitations of an AI's training data.	How to teach it	[Apply] teach evaluation of AI-generated scientific outputs
1.3	Framing AI as a generator of plausible-sounding claims to be checked against warranted sources Conveys that AI outputs are predictive completions to be verified against warranted sources, a distinction students must hold even when outputs appear scientifically credible.	How to frame it	[Understand] frame AI as a generator of plausible-sounding claims to be checked against warranted sources
1.4	Designing assessments of AI-output critique against scientific criteria Designs assessments in which students evaluate AI-generated scientific content against named criteria — source warranting, scope of training, identifiable failure modes — rather than against general plausibility.	How to assess it	[Create] design assessments of AI-output critique against scientific criteria

Developing students' management of AI in scientific work

ALLit Framework · Managing AI (domain level)

Capability. Develop students' deliberate management of AI as a partner in scientific reasoning tasks.

Table 73 ALLit framework, managing AI (domain level).

No	Teacher skill	Teaching focus	Objective
2.1	Teaching the decomposition of scientific tasks into AI-suitable and human-required parts Develops in students the habit of distinguishing the parts of a scientific task that AI can usefully assist — literature scanning, calculation, draft generation — from the parts that require human judgement, such as warranting, ethical evaluation and design choices.	How to teach it	[Apply] teach decomposition of scientific tasks into AI-suitable and human-required parts
2.2	Framing AI use in scientific work as a deliberate methodological choice Conveys that whether to use AI in a given scientific task is itself a methodological decision affecting the warrant of the result, rather than a frictionless productivity move.	How to frame it	[Understand] frame AI use in scientific work as a deliberate methodological choice

Annex IV Analysis of National Policies

STEAM Disciplines

STEAM disciplines are present in all five partner countries but consistently delivered as implicit integration rather than as an explicit, labelled STEAM framework. Slovenia has no standalone STEAM curriculum at all: content is distributed across subject syllabi and cross-curricular guidance documents. Croatia does not formally use the term 'STEAM competencies' in official documents, even where interdisciplinary learning is encouraged in practice. Italy's mainstream curriculum documents largely retain a subject-based STEM structure, although the more recent DM 221/2025 framework moves explicitly towards an integrated STEAM approach. Belgium embeds STEAM content within 16 key competences without naming it as such. Greece comes closest to a deliberate STEAM structure through its Skill Labs and Scientific Unit, but these remain time-limited and not systematically continued into upper-secondary subject provision.

A particularly significant gap across all contexts is the lack of systematic integration of the arts as an epistemological resource within STEAM. Where arts appear in national curricula, they are associated with aesthetic or creative outcomes within art-specific subjects, not treated as a mode of inquiry capable of enriching scientific and mathematical learning. The Arts are not yet systematically integrated into science and mathematics education as an epistemological resource across the national curricula analysed.

Learning Objectives and Skills

Learning objectives are generally well formulated across all partner countries, structured through clear hierarchies of general goals, operational objectives, and knowledge standards. The Italian 2026 curriculum reform and the Greek IEP framework are among the more progressive examples, both explicitly positioning competence-based learning as the organising principle of STEAM education. Slovenian subject syllabi such as Chemistry and Electrical Engineering demonstrate particularly sophisticated approaches, with dedicated interdisciplinary links columns in operative objectives and an explicit orientation toward lifelong transferable skills.

However, none of the partner curricula explicitly align learning objectives with PISA 2025 competency descriptors. This means that teachers across all five countries are developing competences that are broadly aligned with PISA 2025 but cannot deliberately design for, or assess against, the specific dimensions that PISA measures: explaining phenomena scientifically, constructing and evaluating scientific enquiry designs, researching and using scientific information for decision-making, and agency in the Anthropocene.

Transversal and Key Competences

Critical thinking, creativity, collaboration, and problem-solving appear across all curricula, but their relationship to STEAM content is almost always implicit. They are described as outcomes of subject-based learning rather than as deliberately designed pedagogical targets. The Flemish framework is structurally the most explicit, listing these competences within its 16 key competences and



designating them as cross-curricular – but leaving their realisation entirely to school autonomy, which produces significant variability in practice. Slovenia's Chemistry and Physics curricula use the most operationally specific language, treating transversal skills as 'lifelong knowledge' and embedding them in specific task structures, but without connecting them to formal assessment criteria.

Systems thinking – a competence particularly central to GreenComp and to PISA's environmental competency – is mentioned in the Flemish framework and in the Greek Skill Labs but is not systematically operationalised in any partner curriculum as a taught and assessed competence.

STEAM Skill Assessment

STEAM skill assessment is the weakest dimension across the board, and the one where the gap between stated intentions and operational reality is most severe. Most curricula mention formative assessment principles, but none of the partner documents describe STEAM-specific assessment instruments or criteria. Belgium explicitly acknowledges that assessment approaches are left to school-level autonomy, with no national prescription. Italy states that there are no specific assessment approaches to STEAM competences whatsoever. Slovenia describes formative assessment as 'widely established' in practice but notes that it is not explicitly written into curricula.

The most advanced assessment approaches in the partner analyses are found in the Slovenian Chemistry and Physics & Ecology curricula, which use portfolios, process reports, peer evaluation, and continuous observation rather than traditional exams. These are excellent models but remain subject-specific and disconnected from any STEAM or PISA competency framework.

Identified Gaps

No Explicit STEAM Labelling or Unified Framework

Across all five countries, STEAM is an emergent property of how subjects are taught, not a designed and labelled framework. Teachers and students encounter STEAM integration only if individual educators or schools choose to pursue it, creating deep inconsistency within and across systems. Without explicit labelling, STEAM competences cannot be systematically taught, assessed, or recognised through formal certification.

The Arts Dimension is Systematically Marginalised

The 'A' in STEAM is the dimension most consistently absent or underdeveloped in partner curricula. Arts appear in national frameworks as aesthetic or cultural subjects, not as epistemological resources for scientific and technological learning. Design thinking, creative hypothesis formation, visual modelling, narrative argumentation, and speculative scenario building – all arts-based capacities that the ImSTEAM framework and the European LifeComp and EntreComp frameworks identify as integral to STEAM competence. These capacities are not systematically specified as components of science and mathematics learning across the five systems.

STEAM Assessment Does Not Exist as a Distinct Practice

No partner system has established a validated, system-wide assessment framework specifically designed for integrated STEAM competences – that is, the ability to integrate knowledge across



disciplines, apply creative and systemic thinking to scientific problems, and demonstrate competence in authentic, interdisciplinary contexts. Until STEAM competences are explicitly defined, labelled, and connected to PISA 2025 descriptors, assessment generally remains subject-specific and does not consistently capture integrated STEAM performance.

PISA Alignment is Implicit at Best

While several curricula mention competences that align with PISA 2025 dimensions, none of the partner documents explicitly map learning objectives to PISA competency descriptors. The implicit alignment is real but unusable: teachers cannot deliberately design for PISA competencies they cannot identify in their own curricula. This gap is most acute for the Epistemic Knowledge dimension and the Environmental competency, both of which are rarely made explicit or systematically mapped to PISA 2025 descriptors.

Immersive Technology Integration is Structurally Nascent

The Schoolmakers contribution from Belgium documents a particularly important systemic constraint: ICT, AI, and XR implementation trajectories in schools consistently span multiple years and require sustained professionalisation of a core team before broader teacher rollout is possible. This means that the integration of VR/AR into STEAM teaching cannot be treated as a short-term add-on. It requires a multi-year, staged implementation strategy with realistic timelines and institutional support structures.

Competence Development is Invisible at System Level

Across all partner contexts, the development of STEAM competences is not visible at system level. There are no data collection mechanisms, no shared metrics, no benchmarking against national or international standards, and no feedback loops that would enable teachers, schools, or policymakers to know whether STEAM competences are being developed, by whom, at what pace, and with what gaps. This is the problem that the ImSTEAM Skills Intelligence Hub (T2.4) is specifically designed to address.

Slovenia

Analyses conducted by: GoINNO Inštitut, Osnovna šola Solkan, Univerza na Primorskem – Pedagoška fakulteta.

Current State

The Slovenian national curriculum provides a solid disciplinary base for STEAM education, with STEAM-relevant subjects present across gymnasium and lower secondary programmes. STEAM disciplines are present but consistently delivered as implicit integration across subjects rather than through a unified STEAM curriculum: Slovenia has no standalone STEAM label anywhere in its official documentation. The relevant evidence is spread across programme frameworks, subject syllabi, cross-curricular guidance, and assessment documents.



The breadth of interdisciplinary integration is nevertheless notable. Spatial Modelling (technical grammar school) connects descriptive geometry, mathematics, informatics, and modern information technology. Microbiology links biology, chemistry, laboratory work, and sustainable development. Agriculture connects biology, chemistry, physics, and information technology including GPS, drones, and smart machinery. Environmental Education as Education for Sustainable Development is the only explicitly cross-curricular thematic area, connecting natural-science, technical, and social-science perspectives. In the arts dimension, Dance Composition (artistic grammar school) explicitly links dance with biology, informatics, mathematics, and physics – one of the few documented cases where Arts is given genuine epistemological status within a STEAM-relevant subject. The Drawing in Geometry and Technology curriculum provides what is arguably the most complete STEAM integration model in the entire dataset: it uses virtual 3D space as a common platform where mathematical precision, technological tools, and artistic creativity meet to solve engineering problems, and its assessment approach already incorporates product and process evaluation, peer assessment, and standards-based criteria.

The Chemistry Elective curricula (Chemistry Experiments, Chemistry in the Environment, Chemistry in Life) are equally significant. Every operative objective is accompanied by a dedicated Interdisciplinary Links column explicitly connecting chemistry with physics, biology, mathematics, fine arts, and computer science. Students investigate natural dyes and pigments, studying their chemical structure alongside their historical use in art and aesthetic impact on society. Transversal competences are explicitly described as 'lifelong knowledge': problem-solving, information and communication literacy, complex thinking, collaborative work, self-initiative, independence, responsibility, and leadership. Portfolios, peer and self-evaluation, and continuous observation are the primary assessment tools.

At the national policy level, the Recovery and Resilience Plan (RRP 2021) commits to updating 216 curricula to integrate digital competences, computer science content, sustainability competences, and financial literacy by Q4/2025. The Common Objectives framework (2023) defines GreenComp-aligned sustainability competences including systems thinking, critical thinking, problem formulation, futures literacy, adaptability, exploratory thinking, and collective action – the most operationally detailed sustainability competence framework in the Slovenian curriculum dataset. The school-level curricula for Biology and Chemistry (upper secondary) confirm high-quality STEAM-relevant design, explicitly positioning biology as an increasingly interdisciplinary and quantitative science.

The Physics and Ecology Projects curriculum exemplifies STEAM education working well at classroom level: students learn fundamental physics through observation of natural phenomena, with a motivational domain explicitly addressing environmental values. Assessment moves away from traditional testing toward mandatory process reports and presentations. Formative assessment has become widely established in Slovenian practice, even though it is not explicitly written into curricula or official recommendations – terms such as 'monitoring of progress accompanied by appropriate feedback' are used instead.



Gaps

The central structural gap is the absence of explicit STEAM labelling and a cross-system competence framework. Good practice – the Chemistry interdisciplinary links column, the Drawing in Geometry and Technology peer assessment model, the Physics and Ecology motivational domain – is invisible at system level and non-transferable. The motivational domain objectives in Physics and Ecology (valuing nature, recognising environmental impact) are not connected to PISA C_ENV competency descriptors or GreenComp frameworks. Systems thinking is not explicitly addressed in Physics and Ecology. The Arts dimension, while present in isolated cases such as Dance Composition and the Chemistry–fine arts connection, is not operationalised as a component of science or mathematics teaching at system level. Immersive technology is structurally absent: no curriculum integrates VR/AR as a pedagogical modality, despite the obvious fit for molecular visualisation, virtual laboratories, and ecological system simulation. Formative assessment, though widely practised, is not formally codified, which means its quality and consistency cannot be guaranteed or evidenced at the system level.

Desired State and Intervention Areas

The desired state for Slovenia is a nationally recognised STEAM framework that overlays existing subject structures with explicit STEAM labelling, cross-subject competence mapping, and PISA 2025 alignment indicators. The most urgent intervention is the production of a cross-reference index mapping Slovenian national curriculum objectives to PISA 2025 competency clusters, as a key output of T2.1. The Interdisciplinary Links column format used in the Chemistry curriculum should be proposed as a standard feature for STEAM subject design across all partner countries. The Drawing in Geometry and Technology and Chemistry curricula should be formally documented as best-practice STEAM integration cases and disseminated across the consortium through WP6. A VR/AR-enhanced scenario for Physics and Ecology should be developed as a flagship T2.3 output, enabling students to interact with simulated ecological systems and map findings to PISA C_ENV descriptors. The portfolio-based assessment approach from Chemistry should be aligned with PISA 2025 descriptors and connected to the ImSTEAM Open Badge certification system. Osnovna šola Solkan's common goals – Digital Literacy, CS integration, Creative Educational Community, Sustainability-Oriented Education – should be operationalised into specific, assessable STEAM learning objectives cross-referenced with PISA 2025 and GreenComp, and the school should serve as the primary anchor institution for ImSTEAM pilot activities in Slovenia.

"Slovenia has no standalone STEAM curriculum; the relevant evidence is spread across programme frameworks, subject syllabi, cross-curricular guidance, and assessment documents." (GoINNO Inštitut Desk Analysis)

"In Slovenia, formative assessment has become widely established, even though it is not explicitly written into the curricula or official recommendations." (Osnovna šola Solkan Desk Analysis)

"Students should participate in the assessment of tasks." (Drawing in Geometry and Technology Curriculum, 2012)



Sources (active links)

- [The official grammar school curricula, multiple years \(2008/09, 2012, 2015/16\). Professional Council of the Republic of Slovenia for General Education / Ministry of Education and Sport / National Education Institute Slovenia.](#)
- [Curriculum: Mathematics. Gymnasium. Adopted 14 February 2008. Ministry of Education and Sport / National Education Institute Slovenia.](#)
- [Common Goals and Their Integration into Curricula and Knowledge Catalogues. Ljubljana, 2023. National Education Institute Slovenia.](#)
- [Vocational Technical Education – Electrical Technician \(2019\) – Catalogue of Knowledge – Electrical Measurements. Institute of the Republic of Slovenia for Vocational Education and Training \(CPI\).](#)
- [Official grammar school curriculum portal](#)
- [Ministry of Education: National Programme Document for Development Planning until 2033](#)
- [Government of the Republic of Slovenia – Recovery and Resilience Plan \(RRP\), 2021.](#)
- [Ministry of Education / National Education Institute Slovenia: Extended Programme Curriculum in Primary School.](#)
- [Ministry of Education: Primary School Act \(Article 2: Objectives of Education\).](#)
- [Ministry of Education / National Education Institute Slovenia: Starting Points for the Revision of Curricula in Primary School and Gymnasium.](#)
- [Ministry of Education: Common Objectives and Their Integration into Curricula and Catalogues of Knowledge.](#)
- [Ministry of Education / Ministry of Digital Transformation: Digital Education Action Plan \(ANDI\) 2021–2027.](#)
- [Curricula for biology, chemistry, science, physics, mathematics, and computer science for general upper secondary education \(Ministry of Education, National Education Institute Slovenia\).](#)

Croatia

Analysis conducted by: Faculty of Mathematics, University of Rijeka (UNIRI), DANTE, ADRIA.

Current State

The reviewed Croatian curriculum documents indicate a strong orientation towards interdisciplinary learning, analytical reasoning, scientific thinking, digital competence, and broader transversal competences. The National Curriculum Framework organises generic competences into three broad categories (ways of thinking, ways of working and using tools, and personal and social development) which encompass problem-solving, decision-making, critical thinking, metacognition, creativity, and innovation, intended to be developed across all levels, curriculum areas, subjects, and cross-curricular themes. The framework describes these competences as forming 'a unique and



**Co-funded by
the European Union**

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor the EACEA can be held responsible for them.

comprehensive whole,' an ambitious integration agenda that, in practice, lacks pedagogical scaffolding at subject level.

On assessment, the Croatian report provides one of the more detailed analyses in the partner dataset. The national curricular framework describes three assessment approaches: assessment for learning (formative), assessment as learning (self-assessment), and assessment of learning (summative). In mathematics, emphasis is placed on assessing conceptual understanding, mathematical communication, and problem-solving, not merely procedural recall. Arts-related assessment may include process-oriented and qualitative dimensions within separate arts curricula; however, the sources reviewed in this section do not provide sufficient evidence to claim that arts-based assessment is systematically connected to STEM or STEAM learning. The ICT cross-curricular theme focuses on digital literacy, critical information searching, and digital content creation, assessed through practical activities. Overall, the reviewed Croatian curriculum documents contain several elements that are relevant for STEAM-oriented education, including interdisciplinary learning, mathematical reasoning, problem-solving, digital competence, creativity, and innovation. However, this alignment is implicit rather than explicit: the documents do not use STEAM terminology and do not define STEAM competences as a distinct curricular category.

Gaps

The central gap is terminological and structural: Croatian curricular documents do not use the term 'STEAM competencies,' and interdisciplinary integration does not translate into subject-level design or assessment instruments. STEAM-relevant competences can be inferred from broader curricular aims, subject curricula, and cross-curricular themes, but they are not explicitly defined as a distinct STEAM competence framework. Their integration therefore depends largely on interpretation by schools and teachers. The reviewed documents do not explicitly operationalise the Arts dimension as an epistemological component of science, mathematics, technology, or interdisciplinary STEAM learning. The reviewed documents also do not identify explicit PISA 2025-aligned assessment instruments for STEAM competences in the Croatian context.

Desired State and Intervention Areas

The desired state is the development of a Croatian STEAM competence mapping framework that makes explicit the connections between existing curricular aims, subject-specific outcomes, cross-curricular themes, and PISA 2025 competency descriptors. Rather than introducing STEAM as a completely new policy layer, this framework should clarify where STEAM-relevant competences already exist in Croatian curriculum documents and where additional pedagogical guidance, examples, and assessment tools are needed. UNIRI, as a higher education partner, should lead the Croatian strand (T2.1) of this work by producing an annotated curriculum cross-reference map connecting Croatian curricular objectives with PISA 2025 competency dimensions. In addition, the ImSTEAM pilot in Croatia could focus on developing and testing adaptable, PISA-informed assessment tools for STEAM-oriented learning activities, with particular attention to interdisciplinary problem-solving, mathematical reasoning, digital competence, creativity, and the integration of Arts-based approaches where pedagogically meaningful.



"The reviewed Croatian curriculum documents contain several STEAM-relevant elements across subjects and cross-curricular themes, although these are not explicitly framed as a distinct STEAM competence framework." (UNIRI Desk Analysis)

Sources (active links)

- [Ministry of Science and Education: National Curriculum Framework \(proposal after public consultation\), December 2017.](#)
- [Ministry of Science and Education: National Curriculum for High School Education \(proposal after public consultation\), December 2017.](#)
- [Decision on the adoption of the Curriculum for Mathematics for primary and secondary schools in the Republic of Croatia, January 2019. Narodne novine, NN 7/2019.](#)
- [Decision on the adoption of the Curriculum for the Cross-Curricular Theme: Use of Information and Communication Technology for Primary and Secondary Schools in the Republic of Croatia, January 2019. Narodne novine, NN 7/2019.](#)

Italy

Analyses conducted by: I.C. 'Cassarà-Guida', Palermo; Università degli Studi di Palermo, UNIPA, CESIE ETS.

Current State

Italy presents a complete gap profile in the analysis and also a significant policy opportunity in its most recent curriculum documents. The picture is one of profound structural tension between policy ambition and operational reality – a tension that ImSTEAM is well positioned to address.

The most recently published Italian curriculum document in the dataset is the Nuove Indicazioni Nazionali (DM 221/2025), analysed by I.C. Cassarà-Guida. This framework implements a pedagogical shift explicitly described as 'New Humanism,' treating STEAM not as isolated technical subjects but as an integrated language where scientific rigour meets creative design and ethical reflection. Transversal skills are explicitly described as 'the vehicle through which scientific content is delivered.' Assessment focuses on the student's ability to apply knowledge in new and complex contexts and involves the evaluation of the process, not just the final product. Formative assessment tools explicitly recommended include e-portfolios, self-assessment rubrics, and systematic observation. The 'trial and error' method is promoted as a valuable learning experience – one of the most progressive assessment formulations in the entire five-country dataset.

The 2023 STEM Guidelines (DM 184/2023), analysed by both UNIPA and CESIE, represent the most progressive STEAM-relevant policy document in Italian education more broadly. They explicitly promote interdisciplinary learning, state that 'the interdisciplinary and multidisciplinary approach, combined with the integration of theory and practice, therefore forms the core of STEM education,' and reference the 4Cs framework (Critical thinking, Communication, Collaboration, Creativity). Authentic tasks and systematic observation are described as the primary means for assessing STEAM competence acquisition. The PNRR investment strand 'New Skills and New Languages' (Mission 4, Component 1, Investment 3.1) provides the policy window within which ImSTEAM can operate,



promoting the integration of STEM, digital, and innovation skills across all school cycles with a fully interdisciplinary approach.

The mainstream national guidelines – Indicazioni nazionali per i Licei, Linee guida Istituti Tecnici, and Linee guida per l'istruzione professionale – tell a very different story. Learning objectives and competences are formulated for each STEAM discipline separately, often presented in a fragmented manner. Interdisciplinary approaches are not operationalised into concrete learning objectives. Transversal competences are connected to individual STEAM disciplines but not addressed as a cross-disciplinary agenda. Most critically, no specific assessment approaches to STEAM competences are discussed anywhere in these guidelines. The only STEAM-related skills assessed through nationally standardised tests relate to mathematics. The Arts dimension is entirely absent from both the STEM Guidelines and the PNRR framework.

Gaps

Italy lacks assessment instruments that explicitly operationalise integrated STEAM competences within the mainstream national curriculum, despite the presence of broader formative-assessment tools in recent policy documents. The gap between the 2026 curriculum framework's vision of Arts-Science interdisciplinary teaching and actual classroom practice is significant, as implementation depends on school leadership and on structures for cross-departmental teacher collaboration. The Arts dimension is absent from both the STEM Guidelines and the PNRR framework, even while the 2026 curriculum explicitly references it. The Scuola Futura platform and the AIRDM and UMI-CIIM networks represent additional resources for Italian STEAM teacher education that are not yet connected to a validated competence framework.

Desired State and Intervention Areas

The desired state is that Italy produces validated STEAM formative assessment instruments – a gap so severe that Italian schools should be prioritised as the primary validation site in the ImSTEAM assessment toolkit pilot (WP4). UNIPA should lead the development of a STEAM competency cross-reference map for Italian secondary education as a key output of T2.1, and produce an Italian STEAM policy brief – building on the PNRR framework and the ImSTEAM evidence base – making the case for explicit STEAM labelling, Arts inclusion, cross-disciplinary learning objectives, and STEAM-specific assessment. CESIE should lead the development and validation of assessment instruments for the Italian context, drawing on its evaluation expertise and PNRR-connected networks. I.C. Cassarà-Guida should develop two or three Arts-Science interdisciplinary unit templates – drawing on the ImSTEAM classification templates (T2.2) – and serve as the Italian anchor school for ImSTEAM pilot activities. The school's formative assessment practices should be formalised into rubrics aligned with PISA 2025 descriptors and connected to the ImSTEAM Skills Intelligence Hub.

"The objective is to promote an integrated approach to the teaching of STEM disciplines, enriched by the 'A' of Arts, to foster critical thinking and innovation through the contamination of languages."
(Nuove Indicazioni Nazionali, DM 221/2025)



"No specific assessment approaches to STEAM competences are discussed in the guidelines for upper secondary schools. The only STEAM-related skills assessed through nationally standardized tests relate to mathematics." (UNIPA Desk Analysis)

"The interdisciplinary and multidisciplinary approach, combined with the integration of theory and practice, therefore forms the core of STEM education." (Italian STEM Guidelines DM 184/2023, p. 12)

Sources (active links)

- [Linee guida per le discipline STEM, Ministerial Decree No. 184, 15 September 2023, Ministry of Education and Merit \(MIM\).](#)
- [Nuove Indicazioni Nazionali per il curricolo della scuola dell'Infanzia e del Primo Ciclo di istruzione, DM 221/2025, Ministry of Education and Merit \(MIM\).](#)
- [Indicazioni nazionali per i Licei, Ministry of Education and Merit \(MIM\), 2010.](#)
- [Linee guida Istituti Tecnici, Ministry of Education and Merit \(MIM\), 2010.](#)
- [Linee guida per favorire e sostenere l'adozione del nuovo assetto didattico e organizzativo dei percorsi di istruzione professionale, Ministry of Education and Merit \(MIM\).](#)
- [National Recovery and Resilience Plan \(PNRR\) – Mission 4: Education and Research.](#)
- [Scuola Futura platform – STEAM educational resources and programmes.](#)

Belgium – Flemish Community

Analyses conducted by: UC Limburg, UCLL; Katholiek Secundair Onderwijs Maaseik-Kinrooi, Mosa-RT, Schoolmakers.

Current State

Education in Belgium is organised according to the constitutional principle of freedom of education, enshrined in Article 24 of the Belgian Constitution. This principle guarantees both the freedom to establish and organise schools and the freedom of parents to choose a school for their children. Consequently, schools are organised by both public authorities and private organisations, provided they comply with legal quality requirements. The Catholic education network is by far the largest educational provider in Flanders, accounting for approximately two-thirds of all pupils in compulsory education.

Since the Belgian state reforms, responsibility for education has largely been transferred to the Communities. Within the Flemish Community, the Flemish Parliament defines the minimum educational goals (attainment targets), while educational umbrella organisations translate these into detailed curricula. These curricula are developed per subject and per educational stage (covering two school years), providing part of the vertical curriculum progression. Individual school boards and schools retain considerable autonomy regarding pedagogy, curriculum implementation and instructional methods.

The Flemish curriculum framework is one of the most explicit educational frameworks in Europe regarding transversal competences. The sixteen key competences that structure the attainment



targets embed STEAM across multiple domains rather than within separate subjects. Competence in mathematics, science and technology, together with digital competence and media literacy, contains the core STEM content, while other competences contribute inquiry, creativity, collaboration, critical thinking and problem solving. Culture and Expression provides a structural place for the Arts, while systems thinking is explicitly included within the broader competence framework.

Within this competence-based architecture, educational umbrella organisations develop curricula that specify learning objectives, content sequences and pedagogical frameworks. In addition to specialised STEM programmes, explicit STEM learning objectives are embedded within compulsory subjects such as mathematics, science, technology and geography, ensuring that every pupil develops basic STEM competences regardless of their educational pathway. A distinctive feature of the Flemish curriculum is the use of concordance tables, which connect related STEM learning objectives across subjects and educational stages. These allow teachers to identify common STEM goals and strengthen interdisciplinary coherence across the curriculum.

Teacher education has also increasingly embedded STEM. At the request of the Flemish Minister of Education, the STEM for Basic Education initiative collaborated with representatives of all Flemish teacher education institutions to develop a shared reference framework for STEM in initial teacher education. This framework reflects a broad consensus across higher education while leaving sufficient flexibility for institutions to emphasise their own educational vision.

A distinctive contribution of the Flemish context concerns the implementation of digital innovation rather than curriculum design itself. Following the COVID-19 pandemic, the Flemish Government launched the Digisprong programme, providing substantial investments in digital devices, infrastructure and pedagogical support through the Knowledge Centre Digisprong. This was complemented by the Flemish XR Action Plan (2022), which stimulated experimentation with immersive technologies and collaboration between education, research and industry.

Experiences from these initiatives consistently demonstrate that successful implementation of ICT, artificial intelligence and extended reality requires a multi-year implementation process. Sustainable innovation depends on the gradual professional development of a core group of early adopters before broader teacher adoption can occur. Success factors include a shared educational vision, strong leadership, phased implementation, pedagogical alignment with authentic learning goals, opportunities for professional dialogue and a culture of experimentation supported by psychological safety.

Gaps

The challenges are structural rather than conceptual. The strong pedagogical autonomy granted to schools—rooted in the constitutional freedom of education—prevents governments from prescribing instructional methods and results in considerable variation in the coherence and quality of STEAM implementation across schools.

Although Arts education is structurally embedded through the key competence Culture and Expression, explicit interdisciplinary integration with science, technology and engineering remains limited.



While the curriculum includes STEM learning objectives across compulsory subjects and provides concordance tables that facilitate interdisciplinary curriculum design, no dedicated national framework currently exists for assessing integrated STEAM competences. Educational inspection monitors the attainment of educational goals, and central assessments are gradually being introduced, but comparable system-level evidence on students' integrated STEAM competences remains limited.

Although Digisprong substantially improved schools' digital infrastructure, its pedagogical ambitions have only been partially realised. Many teachers continue to experience difficulties integrating digital technologies meaningfully into everyday teaching because their own pedagogical-technological expertise remains insufficient. As a result, technology availability does not automatically lead to effective technology-enhanced learning.

Finally, the sustainable implementation of immersive technologies often requires a considerably longer timeframe than typical innovation projects or funding cycles allow.

Desired State and Intervention Areas

The desired state is a Flemish educational system in which existing curriculum structures are complemented by practical resources that support teachers in implementing integrated STEAM education within the current competence-based framework.

Rather than introducing additional policy layers, intervention should focus on translating the existing curriculum into concrete interdisciplinary teaching materials, classroom scenarios and assessment instruments that enable teachers to connect STEM disciplines with creativity, systems thinking and authentic societal challenges.

The recently developed STEM framework for initial teacher education provides an important foundation for strengthening future teachers' competences. Building upon this shared reference framework, teacher education institutions could further integrate interdisciplinary STEAM pedagogy, digital technologies and immersive learning environments into both initial teacher education and continuous professional development.

Future professionalisation initiatives should explicitly adopt the phased implementation model emerging from the Digisprong and XR experiences: developing a core group of teacher leaders first, followed by gradual school-wide implementation supported by coaching, peer learning and professional learning communities.

Existing infrastructures such as the Knowledge Centre Digisprong and the broader Flemish educational support networks provide valuable dissemination channels for sharing evidence-informed practices across schools. Continued collaboration between educational umbrella organisations, teacher education institutions and schools can further strengthen coherence while respecting the pedagogical autonomy that characterises the Flemish educational system.

Finally, future initiatives should prioritise the development of practical assessment frameworks for integrated STEAM competences, aligned with both the Flemish key competences and international frameworks such as PISA 2025, enabling schools to monitor interdisciplinary competence development more systematically.



Sources (active links)

- [Secondary education attainment targets: Minimum goals and key competences \('Onderwijsdoelen'\). Flemish Government / AHOVOKS, 2025.](#)
- [The Common Foundational Curriculum \(GFL\) in relation to complementary curricula \(14/03/2023, Katholiek Onderwijs Vlaanderen\).](#)
- [Educational Goals in Secondary Education \(Onderwijsdoelen secundair onderwijs\). Current official framework. Flemish Government / AHOVOKS.](#)
- [Schoolmakers: Implementation of ICT, AI and XR in education: success factors and pitfalls. Jan Royackers, 31 March 2026.](#)

Greece

Analyses conducted by: Academy of Athens; University of Patras, UPatras.

Current State

Greece is among the partner systems that come closest to a deliberately structured STEAM approach, through two institutional mechanisms. The Skill Labs – introduced nationwide in 2021 by the Ministry of Education and the Institute of Educational Policy (IEP) and mandatory for all students in Kindergarten, Primary, and Lower Secondary education – represent the most structurally innovative approach to STEAM competence development in the five-country dataset. Each Skill Lab consists of 5 to 7 individual workshops lasting 5 to 7 weeks, targeting the 4Cs (Critical Thinking, Communication, Collaboration, Creativity) as explicit pedagogical objectives rather than secondary by-products of subject teaching. The 2025–2026 cycle focuses on four themes: Live Better – Well-being; Care for the Environment; Care and Act – Social Awareness and Responsibility; Create and Innovate – Creative Thinking and Initiative. The assessment model within Skill Labs is among the most advanced in the dataset: because the 4Cs are qualitative rather than quantitative, the Ministry has shifted away from standard exams toward Descriptive Assessment and Authentic Evaluation. Teachers use IEP-provided Descriptive Assessment Rubrics. The Student Portfolio serves as the primary assessment tool. Self-assessment and peer-assessment through Reflective Journals are explicitly integrated.

The Technology Course curriculum (7th, 8th, and 9th grades) introduces the STEAM content approach as an explicit activity structure, requiring students to participate in interactive inquiry-based learning related to real-world problems through the holistic approach of STEAM and the engineering design process. It also introduces the Greek acronym ΦΥ.Τ.ΕΜ.ΜΑ.Γ for physics education – explicitly described as a variation of STEM where the letter 's' (Science) is replaced by 'ΦΥ' (Physics), and an extension of STEM into STEMΓ where Γ refers to Language – a nationally specific STEAM extension that directly connects to the ImSTEAM Arts integration rationale.

The 2023 Mathematics curriculum for the Gymnasium and Physics curriculum for the Lyceum (analysed by UPatras) are among the most recently issued official curriculum documents in the five-country dataset. Mathematics formulates learning objectives as Expected Learning Outcomes specifying what students should be able to do, with transversal competences – reasoning, modelling,



communication, reflection, critical understanding, and data analysis – explicitly stated and recurring throughout thematic fields. Physics connects the subject with technology and STEM-type integration in ways that closely align with PISA 2025 competency clusters C1 (explaining phenomena scientifically) and C2 (constructing and evaluating scientific enquiry). Physics explicitly promotes two advanced assessment modes: Laboratory Reports (students write reports similar to those of scientific journals) and Argumentative Writing (students practise thinking like scientists, developing arguments and producing scientific discourse). Mathematics adopts 'the formative approach to assessment, and more specifically assessment for learning.'

Gaps

Skill Labs remain a time-limited addition to mainstream teaching rather than a systemic redesign. Their 4Cs framework is not aligned with PISA 2025 competency descriptors, limiting their use for the deliberate development and assessment of competencies defined in the PISA 2025 framework. The 4Cs are developed primarily in lower secondary and are not systematically continued into upper secondary education, creating a developmental discontinuity. The implicit PISA alignment in Physics and Mathematics is real but unusable: UPatras produces teachers who may be well-aligned with PISA 2025 by virtue of their disciplinary training, but who cannot articulate that alignment or design deliberately for PISA competency development in their future students. There is no integration of immersive technology (VR/AR) into the learning objectives for Physics and Mathematics, despite the subjects' obvious potential for virtual laboratory experimentation and 3D modelling.

Desired State and Intervention Areas

The desired state is that the Greek Skill Labs become the primary institutional anchor for ImSTEAM pilot activities in Greece, their 4Cs framework formally aligned with PISA 2025 competency descriptors and extended into upper secondary education through the ImSTEAM curriculum design process. The Academy of Athens should produce a bilingual (Greek/English) resource mapping the Skill Labs framework onto PISA 2025 competencies, and develop PISA-aligned Skill Lab activity templates demonstrating how to use PISA scenarios as prompts for 4C development. UPatras should lead the development of annotated versions of Greek Physics and Mathematics objectives showing PISA 2025 alignment and distribute this resource to Greek secondary schools through the ImSTEAM platform. UPatras should design and pilot at least two VR/AR-enhanced Physics and Mathematics scenarios for the internal pilot and integrate PISA-aligned assessment instruments into its pre-service teacher education programme.

"The acronym ΦΥ.Τ.ΕΜ.ΜΑ.Γ constitutes: (a) a variation of STEM, where the letter 's' (Science) is replaced by 'ΦΥ' (Physics), and (b) an extension of STEM into STEMΓ, where Γ refers to Language." (Lyceum Physics Curriculum, 2023, p. 2)

"In the present curriculum, emphasis is placed on the use of formative assessment, as a pedagogical function dynamically integrated into teaching practice, aiming at the continuous monitoring of the achievement of the expected learning outcomes." (Lyceum Physics Curriculum, 2023, p. 7)



Sources (active links)

- [Institute of Educational Policy \(IEP\): 'Nea Programmata Spoudon' / New Curricula \(2021–2023\).](#)
- [Mathematics Curriculum 2023 – Gymnasium, National Gazette B, 235/2023.](#)
- [Physics Curriculum 2023, Lyceum, National Gazette B, 2913/2023.](#)

The national analyses presented in this annex were synthesized across four dimensions that are particularly relevant for the design of the ImSTEAM intervention: the degree of STEAM integration, the maturity of assessment practices, alignment with the PISA 2025 competency architecture, and readiness for XR-supported educational innovation.

Table 74 Country findings for STEAM integration, assessment, PISA alignment and XR readiness.

Country	STEAM Integration	Assessment	PISA Alignment	XR Readiness
Slovenia	Strong subject-level interdisciplinary integration. STEAM elements embedded across curricula (e.g., Chemistry, Geometry & Technology, Environmental Education), but not explicitly labelled as STEAM.	Advanced formative assessment practices (portfolios, peer/self-assessment, continuous observation), but not systematised as STEAM assessment.	Implicit alignment with PISA competencies; significant overlap with scientific inquiry and environmental literacy, but no explicit mapping.	Moderate potential. Strong fit for virtual laboratories, molecular visualisation and ecological simulations, but XR is not formally embedded in curricula.
Croatia	Interdisciplinary learning promoted at policy level through transversal competencies, but weak subject-level operationalisation. STEAM concept not explicitly used.	Assessment-for/as/of-learning framework exists, but no STEAM-specific instruments or assessment guidance.	Implicit alignment through critical thinking, problem solving and digital competence, but no direct PISA-based framework or tools.	Low readiness. XR and immersive technologies are largely absent from curriculum objectives and implementation guidance.

Italy	Strong policy discourse supporting integrated STEAM and transversal skills, but significant gap between policy intentions and classroom implementation. STEAM integration remains uneven.	Progressive assessment principles (e-portfolios, authentic tasks, self-assessment), yet no dedicated STEAM assessment instruments within mainstream provision.	Partial and implicit alignment. Competencies resonate with PISA constructs but are not systematically mapped or assessed through PISA descriptors.	Low-moderate readiness. National policies encourage innovation and digital transformation, but immersive technologies are not yet systematically integrated into teaching practice.
Belgium (Flemish Community)	Most explicit competence-based architecture. STEAM-related competencies embedded within attainment targets and key competences, although implementation is delegated to schools.	Assessment approaches determined locally; no national STEAM assessment framework. Considerable variability across schools.	Implicit alignment. Systems thinking, inquiry and digital competence provide strong links to PISA, but no explicit competency mapping.	High readiness. Extensive experience with ICT, AI and XR implementation. Clear evidence on organisational requirements, leadership, teacher professionalisation and staged adoption pathways.
Greece	Most structured STEAM-oriented provision through Skill Labs and Technology curriculum. Strong focus on inquiry, problem solving and interdisciplinary learning.	Advanced authentic and formative assessment practices (rubrics, portfolios, reflective journals), though not aligned to STEAM or PISA frameworks.	Closest alignment to PISA among partner countries, particularly in Mathematics and Physics curricula, but still largely implicit.	Moderate readiness. Strong pedagogical fit for immersive learning, but XR is not currently embedded in curriculum objectives or assessment approaches.

The tables reported below provide a comparative summary of the main findings from the analysis of partner education systems across the four key project dimensions: STEAM integration, assessment practices, alignment with PISA frameworks, and XR readiness. The first table offers a cross-cutting overview of systemic trends and key challenges, highlighting recurring patterns as well as differences across dimensions. The second table presents a country-level comparison, relating the development of each dimension to the main structural gaps that limit implementation. Together, these findings provide an analytical basis for interpreting the current state of STEAM education across partner contexts and for informing the subsequent stages of framework development.

Table 75 Overall findings across the dimensions of STEAM integration, assessment, PISA Alignment and XR readiness.

Dimension	Overall finding
STEAM Integration	Present in all systems but generally implicit, fragmented, and rarely identified as a coherent STEAM framework.
Assessment	The weakest dimension across countries. Formative practices exist, but dedicated STEAM assessment instruments are largely absent.
PISA Alignment	Alignment is mostly implicit and therefore difficult for teachers to operationalise in curriculum design and assessment.
XR Readiness	Belgium demonstrates the highest implementation readiness; Slovenia and Greece show strong pedagogical potential; Italy and Croatia remain at an earlier stage of adoption.

Table 76 Gap analysis by country.

Country	STEAM Integration	Assessment	PISA Alignment	XR Readiness	Main Gap
Slovenia	High	Medium-High	Medium	Medium	Lack of explicit STEAM and XR integration
Croatia	Medium	Low	Low-Medium	Low	Absence of operational STEAM implementation
Italy	Medium	Low	Medium	Low-Medium	Assessment and implementation gap
Belgium (FL)	High	Medium	Medium	High	Variability due to school autonomy
Greece	High	Medium-High	High	Medium	Limited continuity and XR integration



Annex V Online Survey Questionnaire

This annex reproduces, in English, the master version of the online survey instrument reported in Chapter 5. The instrument was deployed as five national versions — Italian, Greek, Dutch, Slovene and Croatian — each translated from this master and adapted to the terminology of the national education system; the reproduction below neutralises those national adaptations (for example, references to a specific national curriculum are given here as "your national curriculum"). Each closed item used a five-point response scale; the anchors of each scale are stated with the relevant section below, and the mapping of every national version's anchors onto the common 1–5 metric used in the analysis is given in the scale-mapping table at the end of this annex. The instrument opens with a study-information and consent statement, followed by a respondent profile and ten question sections.

Introduction and Consent

ImSTEAM is building immersive teacher training for science across five countries. This survey gathers teachers' views on how the science curriculum is emphasised and taught, on the constraints under which science is taught, and on where immersive activities — virtual, augmented and mixed reality — could most help. There are no right answers; we want your honest view. The survey is anonymous and takes about eight minutes.

Consent statement: "I have read the above information and consent to participate." (Participation required affirmative consent; no response was recorded without it.)

Respondent Profile

Table 77 Respondents profile collected so that needs can be understood across countries, subjects and settings.

Field	Response options
Country	Single choice from a list of European countries, with an "Other" option
Subject(s) taught	Free text
Level taught	Lower secondary · Upper secondary · Both
Years teaching science	0–3 · 4–10 · 11+
Age group	Under 30 · 30–44 · 45–59 · 60+
Gender	Female · Male · Prefer not to say
School type	Public · Private · (national variants, e.g. subsidised free schools)
Geographic area	Urban · Suburban · Rural



Question Sections

Q1 Emphasis of your national curriculum

How strongly does your national curriculum emphasise each of the following?

Table 78 Survey Q1.

#	Item
1	Explaining how and why phenomena happen, using models
2	Explaining how and why phenomena happen, using graphs
3	Making predictions based on scientific reasoning
4	Designing experiments
5	Making sense of data
6	Evaluating evidence and constructing arguments
7	Making evidence-based decisions
8	Using science to understand and respond to environmental and climate change

Response scale — intensity: *Not at all · A little · Moderately · Strongly · Very strongly*

Q2 What you are actually able to teach in depth

For the same eight areas, how deeply are you actually able to teach each?

Items as in Q1.

Response scale — frequency: *Never · Rarely · Sometimes · Often · Always*

Q3 What limits your science teaching

How strongly does each of the following limit your science teaching?

Table 79 Survey Q3.

#	Item
1	Large classes
2	Limited laboratory space or equipment
3	Few digital devices or a weak internet connection
4	Updated teaching materials / textbooks not available in our language
5	Digital resources not available in our language
6	Examination pressure
7	An overloaded curriculum
8	Little professional development available for teachers
9	Available teaching time
10	Institutional or administrative support

Response scale — intensity: *Not at all · A little · Moderately · Strongly · Very strongly*



Q4 Helping students reason and work as scientists

How difficult is it to help students do each of the following?

Table 80 Survey Q4.

#	Item
1	Explain phenomena on a scientific basis
2	Form hypotheses and design experiments to test them (variables, controls, and so on)
3	Read and interpret data — recognising what is evident and what is latent
4	Build or use a model to explain or predict
5	Construct an argument based on evidence
6	Use evidence and data to make informed everyday decisions (health, diet, environment)
7	Use scientific understanding to address wider societal issues (climate change, public health)
8	Use science to address a real problem or design a solution (health, energy, environment)

Response scale — difficulty: *Very easy · Easy · Moderate · Difficult · Very difficult*

Q5 What makes these competencies hard to convey

How often does each of the following make teaching difficult?

Table 81 Survey Q5.

#	Item
1	Things too small or too large to see
2	Processes too slow or too fast to observe
3	Invisible mechanisms — forces, fields, reactions
4	Experiments too dangerous or costly to run in class
5	Real places, events or situations I cannot bring into the classroom

Response scale — frequency: *Never · Rarely · Occasionally · Often · Very often or always*

Q6 Where immersive activities would help most

How much would immersive activities help in each of the following?

Table 82 Survey Q6.

#	Item
1	Seeing processes that are invisible, too small, too large, or too fast/slow for normal observation
2	Safely running experiments that would be dangerous or costly in a real laboratory
3	Entering a model or phenomenon to understand it from the inside
4	Visiting real places, events or situations that cannot be brought into the classroom
5	Changing variables interactively and observing the effects in real time

Response scale — intensity: *Not at all · A little · Moderately · A lot · Very much*

Open extension: Is there anything else you would like to add about where immersive activities could help? (free text)

Q7 What such an activity would need to be usable for you

How important is each of the following for you to be able to use an immersive activity?

Table 83 Survey Q7.

#	Item
1	Can be carried out within the lesson time available
2	Works with my class size and the devices I have
3	Is ready to use, without demanding preparation
4	Is consistent with the curriculum and examinations
5	Comes with training for teachers
6	Is available in our language

Response scale — importance: *Not important · A little · Moderately · Very important · Essential*

Q8 How prepared you feel

How prepared do you feel on each of the following?

Table 84 Survey Q8.

#	Item
1	Knowledge of immersive tools and what they can do
2	Ability to select and integrate immersive activities into lesson design
3	Technical skills to use immersive tools in class
4	Confidence in managing the class while using immersive tools
5	Ability to assess student learning through immersive activities

Response scale — preparedness: *Not at all · A little · Moderately · Fairly · Very*

Q9 Your experience so far

Which best describes your experience with immersive tools? (single choice)

Table 85 Survey Q9.

#	Item
1	I have never used them
2	I have tried them once or twice
3	I use them occasionally
4	I use them regularly
5	I use them often

Q10 Anything else

Is there anything else you would like to add — for example, the science topics where immersive activities would help your students most? (free text)

Scale-Mapping Table

Each closed item used a five-point scale whose anchors were translated equivalently across the five national versions. The table below records the English master anchors for each scale type and the common 1–5 value assigned to each in the analysis reported in Chapter 5. The national-language anchors of each version map onto the same 1–5 values.

Table 86 Scale-mapping.

Scale (sections)	1	2	3	4	5
Intensity (Q1, Q3, Q6)	Not at all	A little	Moderately	Strongly / A lot	Very strongly / Very much
Frequency – emphasis (Q2)	Never	Rarely	Sometimes	Often	Always
Difficulty (Q4)	Very easy	Easy	Moderate	Difficult	Very difficult
Frequency – sources (Q5)	Never	Rarely	Occasionally	Often	Very often / always
Importance (Q7)	Not important	A little	Moderately	Very important	Essential
Preparedness (Q8)	Not at all	A little	Moderately	Fairly	Very
Experience (Q9)	Never used	Once or twice	Occasionally	Regularly	Often

For Q4 (difficulty), a higher value denotes greater difficulty; for all other scales a higher value denotes greater intensity, frequency, importance, preparedness or experience. The direction of each scale is preserved consistently in the analysis, so that, for example, a high Q4 value and a high Q6 value both indicate a stronger case for immersive support.

Annex VI Online Survey Analysis

This annex presents the complete item-level results of the online survey reported in Chapter 5. It is organised by question section. Sections whose cross-national shape is the object of interest are presented as figures; sections that serve principally as a numerical reference are presented as tables; and a master descriptive table at the end gives the values underlying the figures, so that every figure in this annex and in Chapter 5 can be traced to the item on which it rests. All values are computed on the valid responses to each item under the reading rules of Section 5.1: the country is the unit of analysis, no pooled or weighted average is formed, and the Slovenian and especially the Croatian results are read for direction rather than magnitude. In every table, each cell reports the item mean on the 1–5 scale followed, in parentheses, by the top-two-box percentage — the proportion of respondents choosing one of the two highest categories. For Q4 a higher value denotes greater difficulty; for all other sections a higher value denotes greater emphasis, depth, severity, benefit, importance or preparedness, as set out in the scale-mapping table of Annex V.

Q1 & Q2 · Curriculum emphasis and depth of teaching

The two paired sections are presented together, as the relationship between them is the finding. Figure 11 shows, for each competency and each country, the difference between the depth to which teachers report teaching it and the emphasis they judge the curriculum to place on it. Bars to the right indicate teaching to greater depth than emphasised; bars to the left, the reverse. In four of the five systems the differences are predominantly positive, the exception being Belgium. The underlying means and top-two-box values for both sections are given in the master table (Table 90).

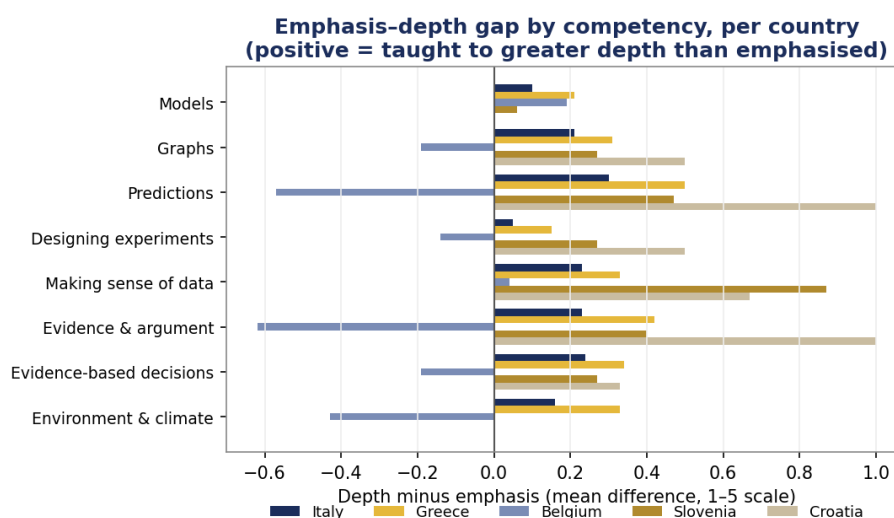


Figure 11 Difference between depth of teaching (Q2) and curricular emphasis (Q1) for each competency, by country.

Q3 · Constraints on science teaching



Table 87 reports the severity of each constraint by country; the cross-national comparison of these values appears as Figure 6 in Chapter 5. Examination pressure, curriculum overload and teaching time are the most severe constraints in Greece and Croatia, institutional support is the single most severe constraint in Belgium, and the language-availability items are severe only in Greece.

Table 87 Constraints on science teaching (Q3): mean severity and, in parentheses, top-two-box percentage, by country.

Item	Italy	Greece	Belgium	Slovenia	Croatia
Large classes	2.98 (39)	4.08 (75)	3.43 (57)	3.93 (73)	3.33 (50)
Lab space/equipment	3.33 (50)	3.88 (62)	3.48 (48)	3.47 (47)	4.00 (67)
Devices/internet	2.97 (37)	3.31 (44)	2.38 (19)	2.40 (7)	2.17 (0)
Textbooks not in language	2.11 (16)	3.62 (58)	2.38 (19)	2.07 (0)	1.83 (0)
Digital resources not in language	2.13 (16)	3.29 (46)	2.10 (14)	2.67 (20)	2.17 (0)
Exam pressure	2.67 (27)	4.48 (92)	2.43 (24)	3.13 (40)	4.17 (67)
Overloaded curriculum	3.00 (36)	4.48 (90)	3.29 (48)	3.60 (67)	4.17 (67)
Little PD available	2.66 (22)	2.85 (31)	2.38 (19)	2.40 (13)	2.67 (33)
Teaching time	3.30 (46)	4.38 (90)	3.48 (48)	3.93 (73)	3.83 (67)
Institutional support	2.98 (36)	3.46 (50)	3.95 (71)	2.60 (13)	2.00 (0)

Q4 • Difficulty in engaging students as scientists

Figure 12 shows the difficulty teachers report in helping students carry out each scientific practice. The difficulty is greatest in Greece across every practice, and is concentrated in every system in forming and testing hypotheses and building and using models. Values appear in the master table (Table 90).

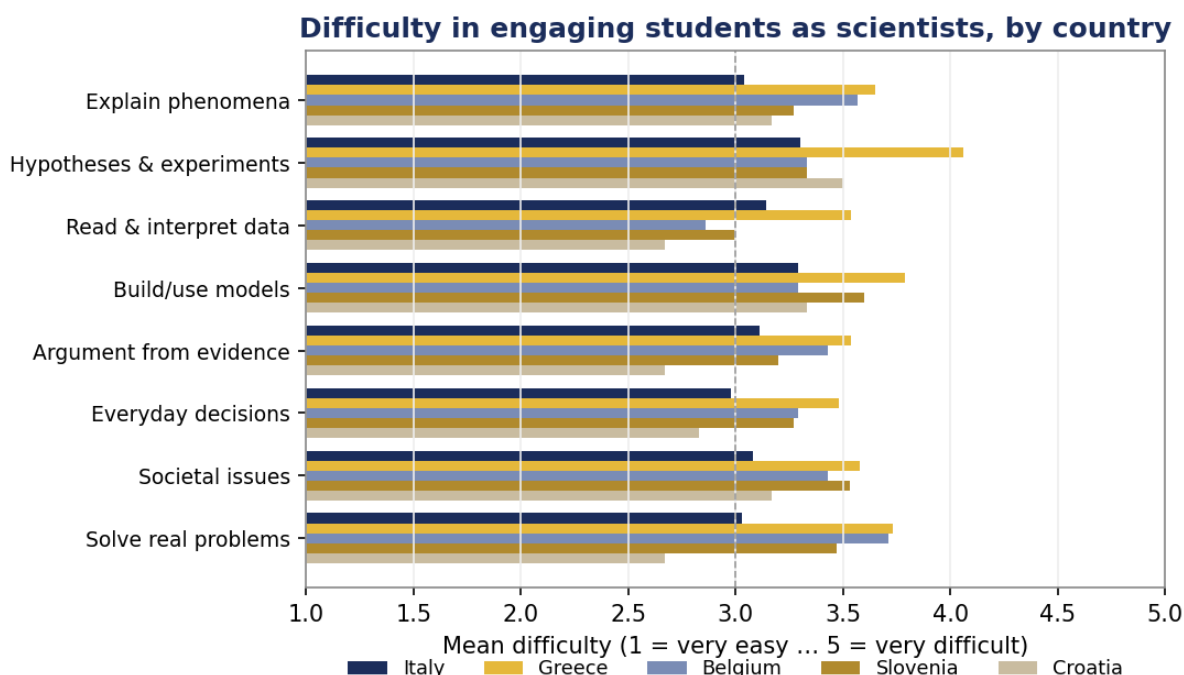


Figure 12 Difficulty in engaging students as scientists (Q4), by practice and country.

Q5 & Q6 · Sources of difficulty and matched immersive affordances

Figure 13 sets, for each country, the reported frequency of each classroom source of difficulty (Q5) against the anticipated benefit of the immersive affordance intended to answer it (Q6). In every panel the benefit bars exceed the difficulty bars across the paired affordances, the relationship reported in Chapter 5 as holding within each national context as well as in aggregate. The item values appear in the master table (Table 90).

Reported difficulty (Q5) vs anticipated benefit (Q6), by paired affordance and country

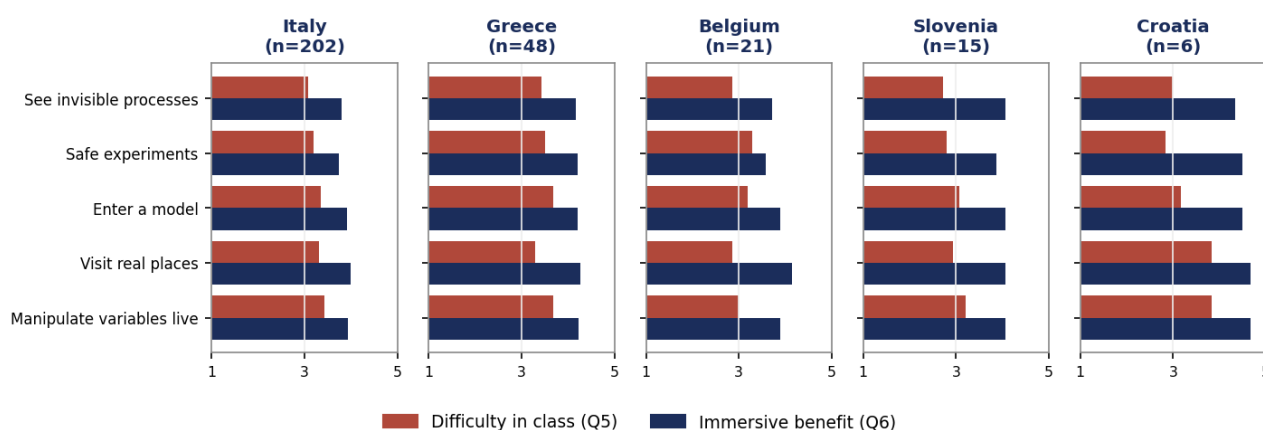


Figure 13 Reported difficulty (Q5) and anticipated immersive benefit (Q6) for each paired affordance, by country.

Q7 · Conditions required for usability



**Co-funded by
the European Union**

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor the EACEA can be held responsible for them.

Figure 14 shows how strongly teachers require each condition for an immersive activity to be usable. Compatibility with class size and devices and with lesson time are the most strongly required conditions in most systems; curriculum and examination alignment is required most strongly in Greece and Croatia. Values appear in the master table (Table 90).

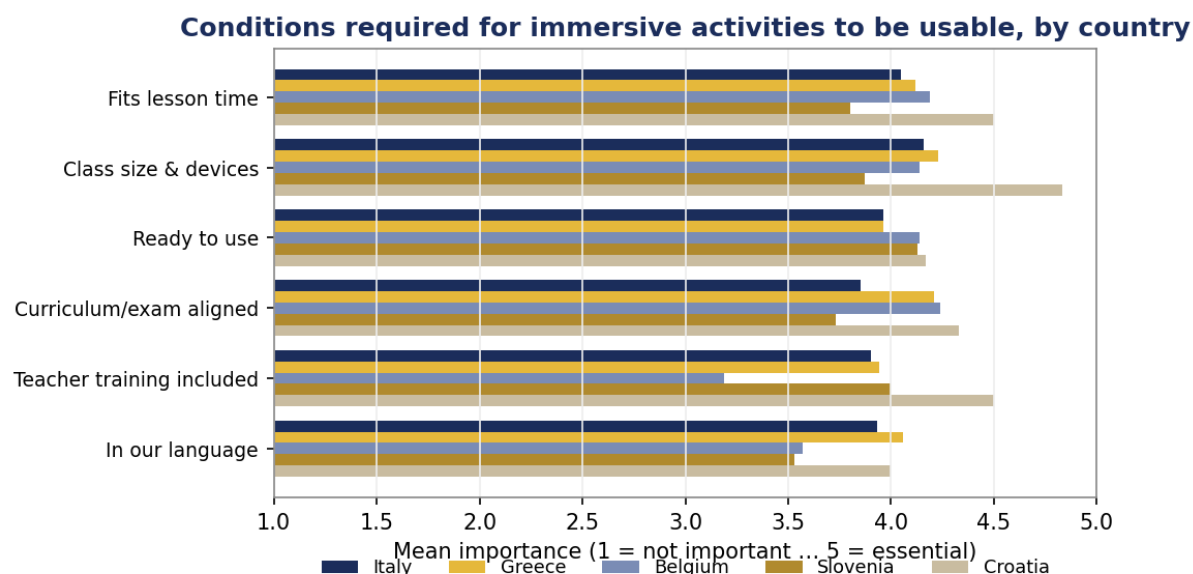


Figure 14 Conditions required for immersive activities to be usable (Q7), by condition and country.

Q8 • Self-assessed readiness

Table 88 reports self-assessed readiness by country; the cross-national comparison appears as Figure 9 in Chapter 5. In every system the pedagogical dimensions sit above the technical dimensions, and knowledge of tools is the weakest item throughout.

Table 88 Self-assessed readiness for immersive teaching (Q8): mean and, in parentheses, top-two-box percentage, by country.

Item	Italy	Greece	Belgium	Slovenia	Croatia
Knowledge of tools	2.81 (25)	3.12 (42)	2.86 (48)	2.93 (27)	2.50 (0)
Integration in design	2.97 (33)	3.42 (48)	3.29 (62)	3.20 (40)	2.83 (17)
Technical skills	2.82 (28)	3.23 (35)	3.33 (52)	2.93 (33)	2.83 (33)
Classroom management	3.16 (37)	3.77 (58)	3.62 (52)	3.53 (53)	3.00 (33)
Assessing learning	3.14 (37)	3.44 (46)	3.05 (48)	2.93 (20)	2.67 (17)

Q9 • Prior experience with immersive tools

Table 89 reports the distribution of prior experience by country; the same distribution appears as Figure 10 in Chapter 5. In every system a majority have either never used immersive tools or tried them only once or twice, and regular or frequent use is rare.

Table 89 Prior experience with immersive tools (Q9): percentage of respondents in each category, by country, with valid n

Country	Never	Once/twice	Occasionally	Regularly	Often	n
---------	-------	------------	--------------	-----------	-------	---

Italy	44%	25%	25%	2%	4%	202
Greece	40%	21%	33%	2%	4%	48
Belgium	33%	24%	38%	5%	0%	21
Slovenia	40%	33%	27%	0%	0%	15
Croatia	33%	33%	33%	0%	0%	6

Master Descriptive Table

The table below gives the mean and, in parentheses, the top-two-box percentage for every item of the figure-based sections (Q1, Q2, Q4, Q5, Q6 and Q7), by country. Together with Tables 87 to 89 it constitutes the complete numerical record of the survey's closed items.

Table 90 Master descriptive table: item mean (top-two-box percentage) for sections Q1, Q2, Q4, Q5, Q6 and Q7, by country.

Item	Italy	Greece	Belgium	Slovenia	Croatia
Q1 — Curriculum emphasis					
Models	3.23 (40)	2.62 (17)	3.00 (29)	3.27 (33)	3.00 (17)
Graphs	3.19 (40)	2.94 (31)	3.19 (33)	2.93 (33)	3.00 (33)
Predictions	3.20 (40)	2.81 (27)	3.00 (38)	3.13 (27)	2.50 (17)
Designing experiments	3.06 (37)	2.58 (15)	2.95 (33)	3.13 (33)	2.83 (17)
Making sense of data	3.36 (46)	2.98 (33)	3.86 (67)	3.00 (27)	3.50 (67)
Evidence & argument	3.32 (46)	2.81 (23)	3.38 (48)	3.33 (33)	2.83 (33)
Evidence-based decisions	3.25 (41)	2.83 (23)	3.24 (38)	3.33 (33)	3.50 (50)
Environment & climate	3.33 (50)	2.73 (25)	3.00 (43)	3.40 (47)	3.00 (33)
Q2 — Depth actually taught					
Models	3.33 (46)	2.83 (23)	3.19 (62)	3.33 (53)	3.00 (17)
Graphs	3.40 (49)	3.25 (40)	3.00 (52)	3.20 (33)	3.50 (33)
Predictions	3.50 (58)	3.31 (48)	2.43 (29)	3.60 (53)	3.50 (50)
Designing experiments	3.11 (37)	2.73 (15)	2.81 (38)	3.40 (47)	3.33 (33)
Making sense of data	3.59 (60)	3.31 (46)	3.90 (86)	3.87 (73)	4.17 (67)
Evidence & argument	3.55 (53)	3.23 (42)	2.76 (38)	3.73 (60)	3.83 (67)
Evidence-based decisions	3.49 (50)	3.17 (44)	3.05 (52)	3.60 (60)	3.83 (67)
Environment & climate	3.49 (56)	3.06 (40)	2.57 (38)	3.40 (47)	3.00 (17)
Q4 — Difficulty engaging students					
Explain phenomena	3.04 (20)	3.65 (46)	3.57 (52)	3.27 (33)	3.17 (17)
Hypotheses & experiments	3.30 (35)	4.06 (79)	3.33 (48)	3.33 (33)	3.50 (50)
Read & interpret data	3.14 (27)	3.54 (48)	2.86 (19)	3.00 (13)	2.67 (0)
Build/use models	3.29 (37)	3.79 (62)	3.29 (48)	3.60 (60)	3.33 (33)
Argument from evidence	3.11 (26)	3.54 (52)	3.43 (57)	3.20 (27)	2.67 (17)
Everyday decisions	2.98 (23)	3.48 (46)	3.29 (43)	3.27 (33)	2.83 (0)
Societal issues	3.08 (28)	3.58 (52)	3.43 (52)	3.53 (47)	3.17 (17)

Item	Italy	Greece	Belgium	Slovenia	Croatia
Solve real problems	3.03 (28)	3.73 (56)	3.71 (57)	3.47 (47)	2.67 (0)
Q5 — Sources of difficulty in class					
Too small/large to see	3.08 (35)	3.42 (50)	2.86 (24)	2.73 (13)	3.00 (33)
Too slow/fast to observe	3.19 (41)	3.50 (58)	3.29 (48)	2.80 (20)	2.83 (33)
Invisible mechanisms	3.36 (50)	3.67 (62)	3.19 (43)	3.07 (33)	3.17 (50)
Dangerous/costly experiments	3.32 (52)	3.29 (48)	2.86 (24)	2.93 (27)	3.83 (67)
Inaccessible real places	3.43 (55)	3.67 (62)	3.00 (33)	3.20 (40)	3.83 (67)
Q6 — Anticipated immersive benefit					
See invisible processes	3.80 (67)	4.17 (79)	3.71 (67)	4.07 (80)	4.33 (100)
Safe experiments	3.74 (65)	4.21 (79)	3.57 (57)	3.87 (73)	4.50 (100)
Enter a model	3.92 (71)	4.21 (81)	3.90 (71)	4.07 (80)	4.50 (100)
Visit real places	4.00 (75)	4.27 (83)	4.14 (76)	4.07 (80)	4.67 (100)
Manipulate variables live	3.94 (71)	4.23 (83)	3.90 (71)	4.07 (80)	4.67 (100)
Q7 — Usability conditions					
Fits lesson time	4.05 (76)	4.12 (88)	4.19 (90)	3.80 (67)	4.50 (100)
Class size & devices	4.16 (83)	4.23 (90)	4.14 (90)	3.87 (67)	4.83 (100)
Ready to use	3.96 (74)	3.96 (71)	4.14 (95)	4.13 (73)	4.17 (100)
Curriculum/exam aligned	3.85 (68)	4.21 (79)	4.24 (81)	3.73 (60)	4.33 (83)
Teacher training included	3.90 (68)	3.94 (75)	3.19 (33)	4.00 (80)	4.50 (100)
In our language	3.93 (72)	4.06 (77)	3.57 (67)	3.53 (60)	4.00 (83)

References

- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman.
- Bianchi, G., Pisiotis, U., & Cabrera Giraldez, M. (2022). *GreenComp: The European Sustainability Competence Framework*. Publications Office of the European Union, Luxembourg.
- Caena, F., & Redecker, C. (2019). Aligning teacher competence frameworks to 21st century challenges: The case for DigCompEdu. *European Journal of Education*, 54(3), 356–369.
- Carlson, J., & Daehler, K. R. (2019). The Refined Consensus Model of Pedagogical Content Knowledge in Science Education. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning PCK in Teachers' Knowledge for Teaching Science* (pp. 77–92). Springer.
- Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054.
- OECD (2023). *PISA 2025 Science Framework (Second Draft)*. OECD Publishing, Paris.
- OECD (2025). *OECD Teaching Compass*. OECD Education Policy Perspectives. OECD Publishing, Paris.
- Redecker, C. (2017). *European Framework for the Digital Competence of Educators: DigCompEdu*. Publications Office of the European Union, Luxembourg.
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15(2), 4–14.
- Webb, N. L. (1997). *Criteria for Alignment of Expectations and Assessments in Mathematics and Science Education*. CCSSO / NISE Research Monograph No. 6.
- Wiggins, G., & McTighe, J. (2005). *Understanding by Design* (2nd ed.). ASCD.

