Final Evaluation of iLEGEND III

Intercultural Learning Exchange through Global Education, Networking and Dialogue

Joint Programme of the European Union and the Council of Europe, implemented by the North-South Centre of the Council of Europe (2022 to 2026)

FINAL EVALUATION REPORT

Raul Guerrero, independent external evaluator

2 September 2026

The opinions expressed in this report are those of the evaluator and do not necessarily reflect the official policy of the Council of Europe or of the European Union. The evaluation was conducted independently; the analysis and conclusions are the evaluator's own. Views gathered in interviews are reported in aggregate and are not attributable to individuals.

Contents

Executive summary 4

1. Introduction 7

1.1 Purpose, objectives and users 7

1.2 Scope 7

1.3 The intervention 7

1.4 The reconstructed theory of change 8

1.5 The results framework 9

2. Methodology 10

2.1 Approach and methods 10

2.2 Sources and sampling strategy 10

2.3 Limitations 11

3. Findings 12

3.1 Relevance 12

3.2 Coherence 13

3.3 Effectiveness 15

3.4 Efficiency 20

3.5 Sustainability 21

3.6 Added value 23

3.7 Human rights and gender equality (cross-cutting) 24

4. The theory of change, tested 25

5. Conclusions 26

6. Lessons learned 27

7. Recommendations 28

Annex 1: Logframe indicator record and quality assessment 32

Annex 2: Evaluation matrix 36

Annex 3: Documents reviewed 39

Annex 4: Data re-analysis 40

Acronyms

AUYD	African University on Youth and Development
CSO	Civil Society Organisation
DEAR	EU's Development Education and Awareness Raising programme
DIO	Council of Europe Directorate of Internal Oversight
EC	European Commission
EQ	Evaluation question
GCE	Global Citizenship Education
GE	Global Education
GEN	Global Education Network
GENE	Global Education Network Europe
GEW	Global Education Week
HEY	Human Rights Education for Youth
LTTC	Long-Term Training Course
NSC	North-South Centre of the Council of Europe
OVI	Objectively verifiable indicator
ToR	Terms of Reference
TOTC	Tutored Online Training Course
UYD	University on Youth and Development

Executive summary

Purpose, scope and approach

iLEGEND III (Intercultural Learning Exchange through Global Education, Networking and Dialogue) is a four-year, 2 million euro joint programme of the European Union and the Council of Europe: it is co-funded by the two organisations (1,500,000 and 500,000 euros respectively) and implemented by the North-South Centre (NSC) of the Council of Europe from December 2022 to December 2026. Through four connected strands (partnership and networking, advocacy and confidence building, capacity building, and awareness raising through the annual Global Education Week (GEW) campaign), it seeks to strengthen global education and youth engagement in European Union member states, candidate and potential candidate countries, and the NSC's Southern partner countries.

This final external evaluation, commissioned by the NSC in agreement with the European Commission, covers December 2022 to a July 2026 cut-off. The activities scheduled for the remainder of 2026 are treated as planned rather than observed; these activities include the 2026 Universities on Youth and Development, the online courses on the scope of the HEY programme, the final GEW edition, and the closing conference. The evaluation is theory-based, applying contribution analysis across six criteria and twelve evaluation questions, with human rights and gender equality as cross-cutting dimensions. It triangulates more than one hundred programme documents, a re-analysis of 46 monitoring and survey datasets, the three Interim Financial Reports and interviews or written contributions involving 21 people across six stakeholder groups; no finding rests on a single interview. Evidence above activity level is largely self-reported (three-month follow-up surveys were answered by an average of 23 percent of participants across activities, with respondents being self-selecting) and GEW reach cannot be verified below coordinator returns; this restricts claims about scale and causality, not the evidence on delivery, participant experience or the operating model.

Main findings

iLEGEND III is highly relevant and in exceptional demand (61 to 4,584 applications for 30 to 99 places per tutored online course), and it delivers reliably: satisfaction is consistently high, completion compares well with sector benchmarks, and recent online courses drew 55 to 74 percent of their participants from Africa at a fraction of the cost of residential provision. The GEW, reported in 45 countries in 2025 and increasingly recognised by national institutions, is the most visible national-level result, although its published participation totals are order-of-magnitude estimates rather than verified counts. Because the monitoring system does not measure behavioural or policy change, **higher-level results remain plausible rather than demonstrated**.

Financial management is sound and an external audit was passed without findings; **the binding efficiency constraint is an activity schedule too dense for the efficient use of the staff resources allocated to it**. Sustainability is strongest where results have been internalised by participants or embedded in institutions, as in Montenegro's curriculum and teacher career system, and weakest where they depend on individuals. **The quadrilogue described in the project documents does not operate as a four-way method in practice**: despite the NSC's efforts, parliamentary and local- and regional-authority participation remains very limited, and cooperation runs through context-specific configurations of civil society and government. **Human rights are strongly embedded in the programme's content but less so procedurally**; gender equality and vulnerable-group participation are not systematically integrated or measured.

Conclusions

Seven overarching conclusions follow:

1. The programme is relevant and delivers outputs reliably, but higher-level results remain plausible rather than proven, and its needs analysis is dated.
2. The formal quadrilogue does not reflect the programme's operating model: despite the NSC's efforts, parliamentary and local- and regional-authority participation remains very limited, and cooperation occurs mainly through context-specific civil-society and government configurations.
3. GEW is the flagship national result, but its measurement system does not adequately substantiate either its precise scale or its effects on attitudes and behaviour.
4. Durability remains dependent on individuals, although institutionalisation is demonstrably possible where the programme is embedded in national education systems.
5. The principal efficiency constraint is an activity schedule too dense for the efficient use of the programme's staff resources, not weak financial management.
6. The NSC's distinctive added value derives from its Global South mandate, neutral convening role and the authority of Council of Europe standards; complementarity with adjacent European global-education networks remains to be formalised.
7. Human rights are substantively integrated in content, though less so procedurally, while gender equality, vulnerable-group inclusion and accountability are not systematically addressed or measured.

Lessons learned

Four broader lessons emerge. Relationship-based networks can extend reach at low cost, but they remain fragile without signed roles, documented processes and institutional ownership. Online delivery can expand Southern access at a fraction of residential cost, but access is decided before enrolment: nomination timelines, language, outreach channels and who is invited to organise or nominate determine who is actually able to benefit. Small grants can better support empowerment and local adaptation than tightly specified contracts when proportionate safeguards are available. Finally, a programme that commits to a participatory convening model must resource that model or redefine the commitment.

Recommendations

The ten recommendations are mutually reinforcing and intended as a design agenda for the follow-on phase; each carries a priority level (high or medium):

1. Resolve the quadrilogue coherently (high priority). Either resource a genuine four-group model with defined parliamentary and local-authority roles, or adopt an explicit convening standard that formalises the institution-wide two-of-four minimum already applied in practice. The design should no longer promise a model the programme does not operate.

2. Institutionalise the network (high priority). Apply the coordinator terms of reference, formalise entry, succession and handover, and pursue the nationally anchored trajectory already visible in curriculum integration, accredited teacher training and domestic budget support.

3. Restore small grants with proven safeguards (medium priority). Reintroduce a modest grants window alongside service contracts, using the guidance, induction and proportionate compliance controls that worked in previous phases.

4. Leverage an alumni architecture (medium priority). Build on the former-participants database operational since the end of 2025: a community of practice, systematic routes from participation into

tutoring and facilitation, recognition for high-performing coordinators and teachers, and links to relevant international processes.

5. Strengthen measurement and data quality (high priority). Establish a verifiable GEW data chain with partner-completed returns, documented compilation, version control and checks of the largest entries; disclose denominators and response rates; report course completion against both selected and active participants; and introduce a small number of measurable outcome indicators with funded data-collection plans.

6. Optimise Southern access within the donor's participation requirements (medium priority). Within the geographic ratios set by the funding instrument, provide timely visa documentation, longer nomination windows, language-accessible editions of flagship courses where compatible with the required balance, diversified venues and organisers including Sub-Saharan actors, and travel arrangements that avoid unnecessary barriers; raise with the European Commission, in the design dialogue for the next phase, any measure that exceeds what the current requirements allow.

7. Consolidate the portfolio (high priority). Plan fewer and deeper activities, front-load the calendar, and match the number and range of activities to the efficient use of staff resources.

8. Agree a complementary division of labour with adjacent European global-education networks (medium priority). In coordination with the European Commission, agree and document complementary constituencies and convening roles, provide for reciprocal participation in key activities, and reflect the arrangement in next-phase documents.

9. Mainstream gender and strengthen the human rights approach (high priority). Use a gender marker for all activities, privacy-compatible disaggregation, at least one gender-responsive indicator per strand, and gender criteria for speakers, trainers and content; add an explicit objective, target and indicator for vulnerable-group participation; and introduce a participant-feedback and complaints channel beyond satisfaction forms.

10. Conduct a renewed needs analysis (medium priority). Update the evidence on who the programme reaches, where needs remain unmet and how needs differ across levels of national maturity, and use the findings to determine the geographic, thematic and delivery choices of the next phase.

1. Introduction

1.1 Purpose, objectives and users

This is the final external evaluation of iLEGEND III (Intercultural Learning Exchange through Global Education, Networking and Dialogue), commissioned by the North-South Centre (NSC) of the Council of Europe in agreement with the European Union (EU). In line with the Terms of Reference, it serves two purposes: accountability, by providing an independent assessment of what the programme achieved with the resources entrusted to it, and learning, by informing the design of the follow-on phase that is under preparation. Its specific objectives are to assess what the programme achieved and how, to understand how it contributed to change, and to identify what was learned and what can be improved in the next phase. Its primary users are the NSC, the Council of Europe Directorate of Programme Co-ordination, the Directorate of Internal Oversight, the EU's Development Education and Awareness Raising (DEAR) programme, and the members of the NSC Executive Committee. The timing gives the evaluation direct utility: its findings and recommendations arrive while the next phase is being designed.

1.2 Scope

The evaluation assesses the programme as a whole against its design (see Section 1.3) and its reconstructed theory of change (see Section 1.4). The thematic scope includes all four activity strands, the intervention mechanisms through which the programme delivers its support. It does not evaluate individual activities in isolation, although activity-level evidence is used to test the causal pathways.

The evaluation covers the whole geographic scope of the programme, including Europe (EU member states, candidates and potential candidates) and third countries mainly from the Southern Mediterranean region, the Middle East and Sub-Saharan Africa (see Section 2.1 for details on the sampling methodology).

The temporal scope covers all implementation years from December 2022 to the evaluation cut-off of July 2026. The 2026 edition of Global Education Week and the programme's closing events fall outside the evaluation window and are treated as planned, not observed, activities.[1]

The evaluation applies the six criteria specified in the Terms of Reference and answers the twelve evaluation questions agreed at the beginning of the evaluation; human rights and gender equality are treated as cross-cutting dimensions throughout (see Section 2).

1.3 The intervention

iLEGEND III is the third phase of a cooperation that began under a joint management agreement and continued through iLEGEND I and II. It is financed at 2,000,000 euros, of which the European Union contributes 1,500,000 euros (75 percent) under the DEAR budget line and the Council of Europe 500,000 euros (25 percent). The programme pursues one impact objective, the active engagement of citizens, especially young people, in sustainable development, through four strands. The partnership and networking strand runs the Universities on Youth and Development in Mollina, Spain, and their African edition (held in Cabo Verde in 2024), together with the youth networks and stakeholder platforms around

[1] The 2026 University on Youth and Development (September 2026), the final North-South Intercultural Conference (October 2026), the Global Education Week (GEW) (November 2026), the HEY Tutored Online Course on Environment in its English and French editions (November 2026), the HEY Cascade training on Intercultural/Interfaith Dialogue, and the African University on Youth and Development (December 2026, to be confirmed) all fall outside the evaluation period.

them. The advocacy and confidence-building strand sustains the Global Education Network of national coordinators, currently spanning more than thirty countries, organises its annual meetings and thematic seminars, and carries NSC contributions into the Council of Europe steering committees for education and for youth. The capacity-building strand delivers tutored online courses, realigned from late 2024 to the Human Rights Education for Youth methodology, together with residential and blended training and a training-of-trainers pathway. The awareness-raising strand coordinates the annual Global Education Week campaign, supported since 2024 by national service contracts with coordinators.

Global citizenship education, the concept the DEAR budget line exists to fund, is the umbrella under which people learn to understand global interdependence (development, justice, sustainability) and to act on it as citizens.[2] Human-rights education, intercultural and interfaith dialogue, media literacy and education for sustainable development are component strands or neighbouring approaches: each can serve global citizenship education when framed globally, and each has its own institutional homes and funding lines when pursued alone. The distinction matters to several findings in this report.

1.4 The reconstructed theory of change

The evaluation reconstructed the programme's theory of change from the Description of the Action and the logframe, refined it with the inception briefing and the Executive Committee synopses, and used it to organise the evaluation. Its central proposition is that sustained engagement of Europe's citizens, especially young people, in sustainable development depends on acting simultaneously on policy dialogue, the global-education competences of practitioners and youth, public awareness, and interregional partnerships between Europe and partner countries. The logframe pairs each immediate outcome with one intermediate outcome, giving four causal pathways with a distinct outputs-layer between activities and outcomes, as Figure 1 shows.

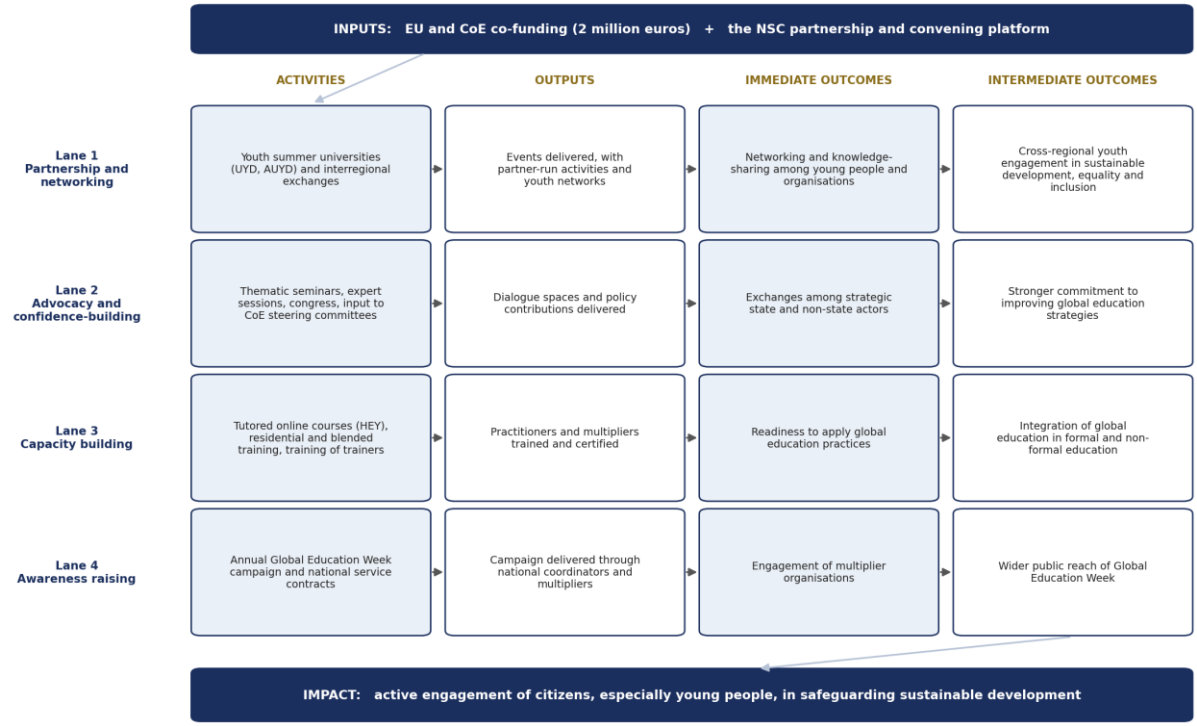


Figure 1. *The reconstructed iLEGEND III results chain: four causal pathways from activities to intermediate outcomes, between shared inputs and a shared impact objective. Cross-cutting mechanisms: the quadrilogue, the human rights approach and the interregional dimension.*

[2] https://dearprogramme.eu

Three cross-cutting mechanisms carry the whole construction: (i) the quadrilogue, the four-way partnership of governments, parliamentarians, local and regional authorities and civil society, which the project documents present as a principal source of the programme's added value; (ii) the human rights approach (HRA), as it is termed in the Council of Europe project management methodology; and (iii) the interregional dimension connecting Europe with Southern partner countries. The reconstruction also made the programme's assumptions explicit, among them sustained political will, open civic space, trained participants acting as multipliers, feasible travel and visas, timely co-funding, and results outliving the funding period. At inception, the evaluation identified the uptake step, between immediate and intermediate outcomes, as the point where contribution claims would be most fragile. Section 4 tests this logic against the evidence.

1.5 The results framework

The iLEGEND III logframe is structured around a single impact statement, four intermediate outcomes, four immediate outcomes and four outputs, with each vertical chain corresponding to one of the programme's four pillars: partnership and networking (chain 1), advocacy and confidence building (chain 2), capacity building (chain 3) and awareness raising (chain 4). At impact level, the project aims to increase the active engagement of citizens (in particular young people) of EU member States, candidate and potential candidate countries in safeguarding sustainable development and confronting global challenges and inequalities. Beneath this, the four intermediate outcomes express the intended changes in behaviour and practice of the programme's main constituencies: increased engagement of young people and youth civil society actors in promoting sustainable development, equality and inclusion between regions; strengthened commitment of strategic state and non-state actors to exchange recommendations on global education strategies; better integration of global education practices into formal and non-formal education; and increased outreach of Global Education Week. Each is paired with an immediate outcome capturing the precursor condition (networking and knowledge sharing, exchange of recommendations, practitioners' readiness to apply global education practices, and engagement of multiplier entities) and with an output describing the opportunities, capacities and resources the NSC delivers directly.

The design logframe contains 26 indicators, with baselines referenced to 2022 and targets set for 2026. The quality of the framework as a measurement and management tool is assessed under evaluation question E1 (Section 3.3), and the indicator-by-indicator record is presented in Annex 1.

2. Methodology

2.1 Approach and methods

The evaluation is theory-based, utilization-focused and participatory, and applies contribution analysis rather than experimental attribution. The findings in the following section are organised by six criteria (relevance, coherence, effectiveness, efficiency, sustainability and added value) and, within them, by the twelve evaluation questions of the evaluation matrix (see Annex 2).

Criterion	Code	Evaluation question (verbatim, ToR)
Relevance	R1	To what extent is the intervention design relevant to the NSC mandate and to the EC DEAR programme as defined in the logical framework / Description of the Action?
	R2	To what extent do the intervention objectives address identified needs of the target group(s) and are they adapted to the institutional, human and financial capacities of key stakeholders?
Coherence	C1	Are the different activities of the intervention logically aligned and mutually reinforcing? (internal coherence)
	C2	To what extent is the intervention coherent with broader policy frameworks? (external coherence)
	C3	To what extent does the intervention complement, add value to, and avoid duplication with relevant DEAR / Council of Europe initiatives, other international networks (GENE, Bridge 47, CONCORD) and national policies?
Effectiveness	E1	What progress has the project made in achieving its objectives, and to what extent has it integrated and advanced human-rights principles (non-discrimination, participation, empowerment, accountability) throughout implementation?
	E2	What are the reasons for the achievement or non-achievement of results?
	E3	How effective and valuable has the Global Education Week been in raising awareness and promoting global education, and what elements of NSC support contributed most to its achievements?
Efficiency	EF1	To what extent were financial and other resources used as intended and in an efficient manner to achieve the project's results?
Sustainability	S1	Which achieved results can be expected to continue after the project ends? What are the reasons, and what could be done to enhance sustainability?
	S2	To what extent was the management of relationships with regional partners and among stakeholders effective in establishing real partnership and ownership?
Added value	AV1	What is the added value of the NSC in its implementation and methodology (the quadrilogue approach)?

Table 1. *The evaluation is structured around six criteria and twelve evaluation questions.*

2.2 Sources and sampling strategy

Four sources of evidence were triangulated. First, a documentary corpus of more than one hundred programme and activity documents, including the Description of the Action and design logframe, the three annual progress reports, the three Interim Financial Reports and the March 2025 budget-revision explanatory note, activity concept notes and final reports, Executive Committee synopses, mission reports, the Global Education Week reports and their underlying survey files, etc. (see Annex 3).

Second, the re-analysis of 46 monitoring and survey datasets covering application records, participant lists, immediate evaluation forms, three-month follow-up questionnaires and campaign surveys; indicators for completion, satisfaction, gender and regional composition were recomputed from the source files rather than taken from the reports (see Annex 4).

Third, semi-structured interviews and written contributions involving 21 people (12 women and 9 men) across six stakeholder groups: national coordinators of the Global Education Network (5), trainers and content experts (4), implementing partner organisations (4), the NSC Secretariat (4), the programme-coordination functions of the Council of Europe and the European Commission (2), and the Executive Committee governance layer (2).

Stakeholder group	Number		
	TOTAL	Women	Men
National coordinators of the Global Education Network	5	4	1
Trainers and content experts	4	3	1
Implementing partner organisations	4	1	3
NSC Secretariat	4	2	2
Programme-coordination functions of the Council of Europe and the EC	2	2	0
Executive Committee governance layer	2	0	2
TOTAL	**21**	**12**	**9**

Table 2. *The evaluation involved 21 key informants (12 women and 9 men) across six stakeholder groups.*

Interviewees were selected purposively[3] and findings are reported only where at least two independent sources converge. In line with the confidentiality assured to interviewees, and given the small number of interviews per stakeholder group, this report does not identify interviewees, does not associate countries with individual informants except where a fact about a country is itself the finding, and does not attribute views by role.

Fourth, the analytical anchors: the findings were mapped against the recommendations of the final evaluation of iLEGEND II (October 2022), the quality and reporting record of the design logframe, and the theory of change reconstructed at inception.

2.3 Limitations

Five limitations qualify the findings and are reflected throughout this report. First, outcome data above the activity level are largely self-reported. Three-month follow-up questionnaires are sent to all participants and every return received is used, but only 4 to 37 percent of each cohort responds (average 23 percent); because responding is voluntary, respondents are self-selecting (typically the most engaged participants) so the results cannot be read as representative of the whole cohort. Second, the GEW reach figures cannot be verified below the level of coordinator self-reports; the report treats them as order-of-magnitude estimates. Third, behavioural and attitudinal change is not measured anywhere in the programme's monitoring system, so contribution claims above immediate outcomes rest on convergent testimony and documented individual trajectories. Fourth, some key stakeholders were unavailable during the evaluation period. Fifth, matters shared with the evaluator in confidence are used only as unattributed structural observations, in agreement with the informants. Within these limits, the evidence is dense: no finding in this report rests on a single interview.

[3] Interviewees were selected purposively against the sampling plan in the Inception Report, with deviations agreed with the evaluation manager, including replacement of unavailable interviewees and additional written governance contributions; the parliamentary and local-authority seats were not pursued after the Secretariat confirmed that their involvement is very limited despite the NSC's efforts, which is itself reported as a finding.

3. Findings

3.1 Relevance

R1. *To what extent is the intervention design relevant to the NSC mandate and to the EC DEAR programme as defined in the logical framework / Description of the Action?*

iLEGEND III responds to the Strategic Framework 2022-2025 (Strategic Priorities 3, 4, 5, 11 and 12) and the Youth Sector Strategy 2030 (Priority 3, "Living together in peaceful and inclusive societies"), and promotes CM/Rec(2011)4 on education for global interdependence and solidarity, CM(2016)25 (Building Inclusive Societies), the White Paper on Intercultural Dialogue (CM(2008)30), the Reference Framework of Competences for Democratic Culture (RFCDC) and the Charter on Education for Democratic Citizenship and Human Rights. The programme also aligns with the 2030 Agenda, with emphasis on SDGs 4, 5, 10, 13 and 16.

The fit with the NSC mandate is substantive rather than nominal. The NSC's statutory mission is to promote Council of Europe values (human rights, democracy and the rule of law) beyond Europe through dialogue and education, and each of the programme's four strands operationalises part of that mission. The universities create the North–South spaces for exchange that the NSC exists to convene; the network and committee work bring global education into intergovernmental settings; the courses translate Council of Europe standards, including the Reference Framework of Competences for Democratic Culture, into practitioner capacity; and the campaign gives the agenda an annual public expression.

On the EU side it supports the DEAR programme, the 2017 European Consensus for Development and the NDICI-Global Europe instrument 2021-2027. The design aligns with the DEAR programme's objective of engaging European citizens with global interdependence by working through multipliers and civil society, consistent with the delivery model supported by DEAR. Third, it is well matched to the joint programme instrument, combining the Council of Europe's convening legitimacy with EU resources in a field where the Commission does not itself bring ministries and civil society together in the same way. Interviews with actors involved in coordination and governance consistently confirmed this institutional fit, and the 2022 evaluation reached a similar conclusion regarding the predecessor design.

Across the interviews, the practical added value of the international framework was described in three main terms: opportunities for peer exchange across countries; the international legitimacy that helps open doors at national level; and access to training and modest financial support that is often unavailable domestically.

Two aspects of the programme's design nevertheless reduce this overall fit. First, the realignment of the capacity-building strand with the Human Rights Education for Youth methodology following the Council of Europe's 2023 strategy change strengthened coherence with the Council of Europe but, in the view of interviewees within the network, diluted the programme's global citizenship education identity and therefore its alignment with the specific rationale of the DEAR funding line, as analysed under C1. Second, the results framework through which the Description of the Action articulates the programme's ambitions is weaker than the underlying programme design it is intended to capture, as discussed under E1 (Section 3.3).

R2. *To what extent do the intervention objectives address identified needs of the target group(s) and are they adapted to the institutional, human and financial capacities of key stakeholders?*

The programme's objectives respond to real and well-documented needs, although their adaptation to differing capacities is only partial and the underlying needs assessment is now dated. Evidence of

demand is unequivocal: course records show between 61 and 4,584 applications for each tutored online course, for between 30 and 99 available places, while the self-paced course is open access and requires no application. Coordinators in countries without a national funding line for global education also described the programme as the only structured support available to them. Content was deliberately adapted to non-European contexts, as can be verified directly in the course materials through the inclusion of regional jurisprudence, case studies and readings, rather than relying solely on the testimony of those who developed them. The courses are also delivered at a cost per participant that is compatible with the capacities of the organisations involved.

Three qualifications emerge from the evidence. First, the programme's targeting model operates primarily at organisational level: it trains representatives of organisations that are, in turn, expected to work with vulnerable groups, rather than targeting vulnerable individuals directly. This is a defensible model, but it is communicated unclearly (several informants, including some within the programme's own network, misunderstood it).

Second, a uniform programme design does not adequately serve countries at earlier stages in the development of global education, where coordinators expressed a need for more differentiated support. The gap between mature, established networks and coordinators operating in contexts with no national funding was described consistently across interviews.

Third, the question of who is ultimately reached remains only partly resolved. Public calls for online courses remain open for three to four weeks and a no-repeat rule prevents any participant from being selected more than once; both measures mitigate concerns about access within the online strand. The more clearly documented constraint concerns nomination windows for residential activities, which are channeled through partner organisations and may be considerably shorter. In one documented case, a ministry declined to participate because the available timeframe would effectively have restricted selection to candidates based in the capital. More generally, short call periods leave limited scope for opportunities to be disseminated beyond already well-connected audiences. In addition, the renewed, target-group-specific needs analysis recommended by the 2022 evaluation was not carried out.

3.2 Coherence

C1. *Are the different activities of the intervention logically aligned and mutually reinforcing? (internal coherence)*

The four strands reinforce one another by design and, increasingly, in practice. Course participants become tutors through the training-of-trainers pathway and contribute to the annual campaign; the multi-partner intercultural-dialogue courses were genuinely co-designed, as partner organisations independently confirmed; and participants in the online courses have progressed through the training-of-trainers pathway into the NSC pool of trainers and subsequently trained others. At least one such trajectory, fully traced by the evaluation, illustrates the intended cascade.

At national level, however, the four strands cohere primarily around the Global Education Week. Coordinators report limited national resonance of the university work and courses beyond individual participation. Internal coherence is therefore strong at the centre but remains dependent on hubs at the periphery.

A second risk to conceptual coherence emerged from the interviews and warrants careful consideration, as it can initially appear contradictory. Following the 2023 Council of Europe strategy change, several courses were renamed: tutored online courses on global education and a thematic area, such as media or dialogue, became Human Rights Education for Youth courses on those same themes. The syllabi

remained substantially unchanged; what shifted was the identity placed in the foreground. This explains why interviewees could state both that human rights education had been brought to the front and that the course content had remained the same. The change is therefore real, but it is primarily one of framing and programme identity rather than curriculum.

The implications are nevertheless significant. Global citizenship education, as defined in Section 1.3, is the concept funded through the DEAR line and the network's principal unifying identity, while neighbouring educational strands have their own institutional homes and funding sources. As the programme's public identity moves towards those strands, network members perceive the field's only dedicated funding line as increasingly carrying agendas supported elsewhere, while the concept intended to unify the network becomes, in one interviewee's words, "foggy".

Conceptual identity is itself a dimension of coherence. The next phase should therefore make an explicit strategic choice about the programme's core educational concept and articulate that choice consistently across its positioning, terminology and delivery.

C2. *To what extent is the intervention coherent with broader policy frameworks? (external coherence)*

Coherence is strong overall and became more pronounced following mid-course adjustments. The programme sits squarely within the Council of Europe's education and youth frameworks, including the Reference Framework of Competences for Democratic Culture, which underpins the course structure. Its alignment with Council of Europe priorities was further strengthened after the 2023 Reykjavik Summit through a renewed emphasis on human rights education and the introduction of thematic content on artificial intelligence and the environment.

On the EU side, the programme serves as an instrument of the DEAR agenda, while the campaign's annual themes are deliberately linked to broader European and global priorities. National coordinators emphasise that this connection is central to making the annual motto relevant and adaptable at country level; conversely, a theme perceived as too vague reduces its usefulness.

At national level, coherence is achieved primarily through alignment rather than formal integration. The campaign and courses are compatible with national curricula and development-education policies where these exist, and in at least one documented case the campaign is explicitly linked to the national curriculum framework (Montenegro). In several countries, however, the absence of a national global-education policy means that the programme moves beyond complementing existing policy and, in practice, partly substitutes for it.

Coherence with African policy frameworks is also evident, but remains under-exploited. The courses reference African Union (AU) instruments, and the African edition of the university was aligned with the AU's Year of Education. Nevertheless, informants working in the African context consistently called for stronger and more explicit links to the AU, both in programme content and in institutional partnerships.

C3. *To what extent does the intervention complement, add value to, and avoid duplication with relevant DEAR / Council of Europe initiatives, other international networks (GENE, Bridge 47, CONCORD) and national policies?*

The programme occupies a distinctive niche in an ecosystem that has thinned considerably, increasing both its relevance and the responsibility to define its role clearly. As several member platforms confirmed during the interviews: CONCORD, the European confederation of development NGOs, has withdrawn from global citizenship education; Bridge 47 no longer provides an alternative infrastructure; and the DEAR budget line has become the field's principal source of funding. Against this backdrop, the

evaluation found no material duplication in the programme's course offer. The environment course, for example, complements rather than replicates the Council of Europe's legal-professional training by adding content on climate change, litigation and indigenous rights. Similarly, the programme's multi-partner dialogue courses (the Tutored Online Course on Intercultural/Interfaith Dialogue and its Human Rights Education for Youth successor editions) are complementary by design, with each partner contributing a distinct institutional mandate and area of expertise.

Two issues of complementarity nevertheless require attention. First, the relationship with GENE, the ministries-based European network identified as a partner in the iLEGEND III application, remained limited during the implementation phase despite increasing convergence between the two networks' mandates. The EC has signalled to both organisations the need to avoid duplication. A clearer, preferably written, division of labour therefore appears necessary, built around the NSC's distinctive reach into civil society, youth constituencies and countries beyond Europe, and supported by reciprocal participation in key events and activities. No duplication of activities was identified during this phase; the finding is taken up in Conclusion 6 and Recommendation R8.

Second, complementarity could be strengthened at event level. Interviews indicated that the platforms co-located at the Mollina university coordinate only to a limited extent, leaving potential operational and substantive synergies unrealised. Programme management has already identified the relationship with GENE as one of the two most urgent issues to address in the next cycle, underscoring the strategic importance of resolving these questions.

3.3 Effectiveness

E1. *What progress has the project made in achieving its objectives, and to what extent has it integrated and advanced human-rights principles (non-discrimination, participation, empowerment, accountability) throughout implementation?*

At output and immediate-outcome level, progress is strong and documented; above that level it is plausible but undemonstrated. The table below summarises the delivery record of the online courses as recomputed from the source files.

Indicator	Value	Source
Satisfaction with online courses (1 to 6 scale)	4.6 to 5.6 across all editions	Evaluation forms, re-analysed
Completion, share of selected participants	55 to 83 percent by course	Participant lists and course reports
Certification rate, 2025 tutored courses	77.5 percent	Year 3 progress report (OVI 21)
Applications per online course	61 to 4,584, for 30 to 99 places	Course reports
African participants, recent online courses	55 to 74 percent	Participant lists, re-analysed
Residential events, satisfaction	90 to 94 percent	Event evaluation forms, re-analysed
Self-reported application or sharing of learning at three months	61 to 97 percent of respondents (response rates 4 to 37 percent)	Follow-up surveys, re-analysed

Table 3. *Delivery and learning record of the training strand. All figures recomputed from source datasets except where a report is cited.*

Completion requires contextual interpretation, and the different delivery formats should be distinguished. For the tutored online courses, which select a limited number of participants through open calls, completion among selected participants ranges from 55 to 83 percent by course. When calculated only among participants who became active, the rate rises to around 75 percent or more (77.5 percent in 2025 under the programme's reporting convention). These figures capture different dimensions of performance and should therefore be reported together. For the self-paced course, which is open access and involves no participant selection, the indicator record shows completion rates of approximately 43

to 45 percent. Both sets of results compare favourably with benchmarks for their respective formats: academic reviews of open online courses generally report completion rates below 13 percent, with the most recent large-scale analyses reporting median completion of around 3 percent among non-paying learners and around 46 percent among verified learners, while tutored professional courses typically achieve rates of 40 to 70 percent.[4] The reasons for non-completion are discussed under E2.

The programme's own indicator record also corroborates delivery at output level: events were held, courses delivered, the network expanded, and the campaign grew. Above the immediate-outcome level, however, the evidence base becomes much weaker. The logframe's two impact indicators, both of which identify this evaluation as their data source, were never reported against, while the only outcome indicator with a reliable measurement remained below target on the one occasion it was properly recorded.

The quality of the results framework conditions what the programme can demonstrate. The design logframe contains 26 indicators, with baselines referenced to 2022 and targets set for 2026; as a design tool, the framework is formally complete (each indicator carries a baseline and a defined endline target) but managerially weak, offering limited support for tracking progress or informing decisions over the life of the project. This weakness was compounded in Year 1 reporting, when a broken link on the Project Management Platform caused an outdated version of the logframe (containing 38 indicators) to be used by mistake. As a result, the first year of reporting was conducted against a superseded framework rather than the 26-indicator design version, introducing inconsistencies between the indicator set that was reported on and the one that should have governed monitoring.

Most indicators are specific and time-bound, but measurability is limited in practice. Eleven are activity or reach measures, while nine rely on self-reported satisfaction or intentions collected through anonymous forms that, for three consecutive years, prevented the disaggregation originally envisaged. Only six indicators point towards behavioural or policy change; of these, the two impact-level indicators identify this evaluation as their data source and were never reported against during implementation.

The targets are also generally modest. Two simply reproduce their baselines, while most percentage thresholds were highly likely to be exceeded given the reliance on self-selected respondents. The one genuinely ambitious target (620,000 campaign participants) is also the one supported by data that could not be independently verified.

[4] One of the most persistent challenges highlighted in both academic research and the popular press is the consistently high dropout rate among learners enrolled in massive open online courses (MOOCs). Although MOOCs often attract thousands of participants, completion rates for many courses remain below 13%. Onah, Sinclair, and Boyatt (2014) examine MOOC attrition from several perspectives, with particular attention to learner behavioural patterns and the factors that may contribute to disengagement and non-completion. Reference: Onah, Daniel F. O., Sinclair, Jane and Boyatt, Russell (2014) Dropout rates of massive open online courses: behavioural patterns. In: 6th International Conference on Education and New Learning Technologies, Barcelona, Spain, 7-9 Jul 2014. Published in: EDULEARN14 Proceedings pp. 5825-5834. ISBN 9788461705573. ISSN 2340-1117. More recent large-scale evidence reports still lower completion among non-paying learners and around 46 percent among verified learners: Reich, Justin and Ruipérez-Valiente, José A. (2019) The MOOC pivot. Science 363(6423), pp. 130-131. Available at: https://wrap.warwick.ac.uk/id/eprint/65543/#:~:text=Although%20many%20thousands%20of%20participants,most%20courses%20is%20below%2013%25

Indicators´ dimension	Assessment	Details
Specific	Partial	Most indicator texts are precise about what to count, but key constructs are undefined ("initiative at interregional level", "early follow-up", "political initiative"), which enabled the OVI 3 construct-substitution in Y3 and the contradictory OVI 4/5 claims across years. Indicator wording itself drifts between reports (OVI 19's three variants). The Y2 report even carries swapped targets (OVI 18 shows a satisfaction target, OVI 20 shows the participant-count target) and "TBD (Ref. year XXXX)" placeholders in year two of implementation.
Measurable	Weak	Three data classes: (i) verifiable counts (meetings, certificates, language versions, countries) are genuinely measured; (ii) self-reported percentages from end-of-activity and 3-month forms are collected, but Ns and response rates are not disclosed, forms are anonymous so the promised gender/age/region disaggregation fails repeatedly ("disaggregation not available", Y2 OVIs 9/10; "not possible", Y3 OVI 20), and respondents are self-selected satisfied completers; (iii) GEW reach (OVI 8) is self-reported by coordinators through the "chain of trust" with "no direct monitoring or verification mechanisms at the NSC level". The impact OVIs are not measured at all.
Achievable	Yes, arguably to the point of becoming a weakness	The target architecture is generally undemanding. Satisfaction and intention thresholds of 70-80% are exceeded by 10-25 percentage points each year, largely by construction of the respondent pool; two targets simply reproduce their baselines (OVIs 15 and 22, both at 80%); and the impact-level ambition amounts to just one additional political initiative over four years. The target for countries participating in GEW was set at 30, only five above the baseline of 25, and was ultimately exceeded by 15 countries. The one clearly stretching target (reaching 620,000 people through GEW, roughly twice the baseline) is also the one that rests on data that could not be independently verified.
Relevant	Output- and perception-heavy	Of 26 indicators, 11 are activity/reach counts, 9 are satisfaction/intention self-reports, and only 6 point towards behaviour or policy change (OVIs 1, 2, 3, 4, 5, 6); of which two are unmeasured, one was construct-switched, one is "Not achieved", and whether iLEGEND III contributed to the instruments reported under OVI 5 (including the RFCDC-ESD guidance document) has not been verified. Nothing measures whether the quadrilogue convened, whether gender outcomes moved (no gender indicator survives, only disaggregation clauses that fail), or whether training changed practice beyond a 3-month self-report.
Time-bound	Formally yes	Reference years, including 2026, are specified in the PMM. However, the Year 2 "TBD" targets and ambiguity over whether targets are annual or cumulative make it difficult to determine what achievement "on time" actually means. For example, OVI 8 is explicitly labelled "Annual", yet Year 3 treats progression from 430,000 to 620,000 as "On track"; conversely, OVIs 14, 17 and 18 are defined as cumulative but are reported as annual figures without running totals.

Table 4. *Most of the programme's indicators are specific and time-bound, but measurability is limited in practice.*

The programme relies on a formally complete but managerially soft results framework: internally coherent on paper, but populated largely with the more readily available half of the data originally envisaged. Its modest targets make positive performance ratings relatively easy to achieve, while the outcome and impact layer is either not measured, affected by construct drift, or effectively deferred to the final evaluation. Yet the evaluation was neither designed nor resourced to generate the data required for these indicators. Much of this evidence should have been collected progressively throughout implementation, leaving the evaluation to assess and synthesise results rather than reconstruct the missing evidence base retrospectively.

The indicator-by-indicator reporting and achievement record is presented in Annex 1 and analysed further under E2 below.

Human-rights integration is strongest where content is centrally produced. Recent courses are grounded in European human-rights law and jurisprudence; the environment course frames environmental protection explicitly as a rights issue and was deliberately designed to avoid a Western-centric perspective; and safeguarding practices at residential events were described in interviews as rigorous. Of the four principles identified in the evaluation question, participation is visibly reflected in the programme's convening practices, while non-discrimination is embedded in participant-selection

rules. Empowerment was weakened during this phase by the replacement of small grants with deliverable-bound contracts (see EF1 and S2). Accountability is the least developed principle: no grievance mechanism or participant-feedback loop beyond satisfaction forms was documented, and human-rights outcomes are not systematically tracked. These are process dimensions of the human rights approach; they are examined further in Section 3.7 and taken up in Conclusion 7 and Recommendation R9. The gender dimension is assessed separately in Section 3.7.

E2. *What are the reasons for the achievement or non-achievement of results?*

The enabling factors are consistent across sources: residential interaction, skilled facilitation, dedicated resources, iterative course design that improved across successive editions, and, for the campaign, the introduction of national service contracts in 2024. **The constraining factors are equally consistent and are largely documented in the programme's own reporting**. Most are fundamentally efficiency constraints with consequences for effectiveness; this report therefore classifies them under efficiency while noting their causal effects here.

A senior post remained vacant from the end of 2022 until mid-2024, as recorded in the Year 2 progress report. The consequences extended beyond increased workload. Because the network depended heavily on personal relationships, the vacancy disrupted continuity with coordinators and partners, delayed the establishment of framework contracts, and left agreed follow-up from the stakeholder platform without a clear owner. Event preparation windows were also consistently short. Facilitators reported receiving contracts only weeks before events and being allocated limited paid preparation time, conditions they associated with reduced pedagogical depth and, in some cases, near-burnout.

The causes of course non-completion are similarly identifiable. Participants are more likely to disengage during self-paced phases when balancing employment or facing unstable internet connectivity, while some selected participants never become active at all (hence the programme's use of an active-participant denominator). Course teams also documented substantial participant use of artificial-intelligence tools in assessed activities. Management has responded by adapting the pedagogical design of the online courses, but corresponding controls have not yet been extended to all assessed components, including the final evaluation forms. Travel and visa constraints caused further documented attrition, disproportionately affecting participants from Sub-Saharan Africa.

Finally, a substantial share of activities and expenditure is concentrated in the final months of the programme in 2026, a scheduling risk that management itself has acknowledged as a mistake. As the closing activities fall outside the evaluation period, the completeness and quality of final-year delivery cannot be assessed within this evaluation.

E3. *How effective and valuable has the Global Education Week been in raising awareness and promoting global education, and what elements of NSC support contributed most to its achievements?*

The campaign is the programme's most visible result at national level, and its value is perhaps clearest where the quantitative results are relatively modest. Progress reports document activities in 26 countries in 2023, 28 in 2024 and 45 in 2025, with nine new organisations joining in the final documented year. Across interviews, the campaign was consistently identified as the centrepiece of national-level work. In some countries, it has been sustained for more than a decade and has gained recognition from national ministries.

This growth reflects identifiable enabling factors rather than momentum alone. The national service contracts introduced in 2024 transformed coordination from largely voluntary work into a resourced

function, supporting dedicated staff time, communication and school outreach, and the production of printed materials. Over successive editions, the accumulation of tools and templates also reduced the effort required to re-mobilise the campaign each year, while peer exchange within the network enabled practices to spread between countries.

Montenegro illustrates what the campaign can achieve when its national institutional anchor provides access to public systems. There, coordination is located within the Ministry of Education, creating opportunities to link campaign participation to the teacher career-promotion system and the new curriculum framework. Comparable recognition has not systematically emerged, largely because most national coordinators are civil-society organisations without equivalent access to education-system structures. In this respect, Montenegro encapsulates the wider quadrilogue finding: the campaign's potential for institutionalisation depends substantially on who holds the national coordinating role and the systems to which that actor has access.

According to those responsible for delivering the campaign, the NSC support elements that contribute most are the national service contracts, annual thematic framing when it aligns with wider European agendas, editable communication materials, and the peer network among national coordinators.

For the analysis that follows, "reach" refers to the number of people whom coordinators report as having participated, either in person or online, in registered campaign activities. It is therefore a measure of activity participation, not of changes in attitudes or behaviour. The published reach figures also require careful interpretation, as summarised in the table below.

Year	Published claim	What the underlying data support	Verdict
2023	312,667 participants; 7,000 registered activities; 26 countries	No reach dataset exists for 2023; the planning document contains pre-event entries only; the activities figure has no documentary basis the evaluation could locate	Not verifiable
2024	Over 300,000 participants; about 6,000 activities; 28 countries	Coordinator survey (30 responses, mostly banded estimates); band minima support roughly 240,000, with one country supplying over 80 percent	Order of magnitude only
2025	430,717 participants; 5,698 activities; 45 countries	Coordinator survey (36 responses); the workbook's own country sheet sums to 397,380; the largest entry was later documented at well below the figure used	Order of magnitude only; totals likely overstated

Table 5. *GEW published reach against the underlying data. Sources: progress reports, campaign reports and survey files, evaluation re-analysis, and documentation obtained during the evaluation.*

The NSC's own campaign evaluation notes that no activity-level data are collected beyond estimated participation totals, while the Year 3 progress report acknowledges that decentralised figures rely on reporting by national coordinators and are not independently verified at NSC level. **The most defensible conclusion is therefore that the campaign mobilises hundreds of thousands of people across dozens of countries; the apparent precision of six-digit totals is not supported by the underlying data**. At the same time, the compilation process is traceable and the largest revision to the figures is documented. The issue is thus one of verification and version control, and it could be addressed through relatively simple and low-cost procedural improvements (see recommendation R5).

3.4 Efficiency

EF1. *To what extent were financial and other resources used as intended and in an efficient manner to achieve the project's results?*

Financial resources were used as intended, under sound and externally validated management. The table below summarises execution.

Item	Value	Source
Budget	2,000,000 euros (EU 75 percent, Council of Europe 25 percent)	Contribution agreement
Spent, end-2023	336,098 euros (17 percent)	Interim Financial Report 1
Spent, end-2024	823,483 euros (41 percent)	Interim Financial Report 2
Spent, end-2025	1,307,549 euros (65 percent; about 68 percent with commitments)	Interim Financial Report 3
Budget revision (approved 2025)	Travel raised by 27 percent and services by 4 percent, offset by savings on staffing circumstances, the local-office line and national workshops; total unchanged	Explanatory note to the EC, 14 March 2025
Estimated direct cost per participant	Roughly 125 to 150 euros (online course) against up to approximately 2,000 euros (residential, including board and lodging, tutors, travel and related participant costs)	Management estimates
External audit	One EC-commissioned audit, passed without findings	Finance function

Table 6. *Financial execution and efficiency markers.*

The underspend recorded in the first year is largely attributable to documented delays in Council of Europe recruitment. The finance function nevertheless reports that expenditure in the final year remains on track, supported by planned costs associated with the closing conference, this evaluation, and ongoing campaign contracts.

The March 2025 budget revision, submitted alongside the second financial report, is significant not only in financial terms but also for what it reveals about implementation choices. The increase in the travel budget reflects both higher-than-anticipated airfares during an inflationary period and a deliberate decision to facilitate the participation of a larger number of young people from Africa. Part of the budget reallocation therefore directly supported the inclusion of participants from the Global South, an issue examined elsewhere in this report in relation to access. Conversely, the reduction in the allocation for national workshops reflects the fact that only 26 of the 40 originally envisaged national providers could be selected, as documented in the project records.

Value for money varies considerably by delivery modality. The online strand enables participation from the Global South at a substantially lower cost than residential activities, while the evidence gathered for this evaluation consistently identifies in-person interaction as the strongest single enabler of results.

Staff costs account for approximately 40 percent of direct costs, a proportion broadly consistent with previous phases of the programme and with comparable Council of Europe projects.

Two structural observations complete the assessment, each with important implications for the programme's efficiency and sustainability. First, under the contribution agreement, the NSC is required to provide a 25 percent co-financing share against a largely fixed core budget, compared with co-financing rates of around 5 to 10 percent in comparable Council of Europe projects. This means that the programme absorbs a substantial share of the NSC's limited discretionary resources, leaving little capacity to finance activities not covered directly by the project, including coordinator travel, follow-up between events, and broader institutional development. At this level, co-financing no longer simply complements the NSC's capacity; it places a significant claim on it.

Second, **staff time is budgeted below the effort that implementation actually requires**, as staff across functions consistently reported. This creates a form of hidden in-kind contribution, understates the true cost of implementation, and contributes to workloads that programme staff themselves describe in terms of fragmentation and exhaustion. Three internal sources independently reached the same conclusion: the number and range of activities in the portfolio is not matched to the efficient use of the staff resources allocated to it. This convergence of evidence provides the basis for Recommendation R7.

In terms of delivery mechanisms, the current phase replaced the small-grants model used under iLEGEND II with service contracts tied to defined deliverables. While this approach is administratively lighter, most stakeholders consider it less empowering for recipients. This trade-off is particularly relevant given that the NSC has an established track record in grant management, supported by tested safeguards and only one recorded case of non-compliance across previous grant schemes.

The main remaining efficiency risk is the concentration of activities and expenditure in 2026, which management itself has acknowledged as a planning weakness. This compression increases pressure on staffing, procurement, financial absorption, and implementation quality, and may compound the structural capacity constraints identified above.

3.5 Sustainability

S1. *Which achieved results can be expected to continue after the project ends? What are the reasons, and what could be done to enhance sustainability?*

What can reasonably be expected to continue is primarily what has already been institutionalised or internalised; most other results remain dependent on continued funding and on particular individuals. The campaign is the most durable outcome. Coordinators report that they would continue it, albeit at reduced intensity, without external support; in several countries it has become an established tradition; and in Montenegro it is embedded in both the curriculum framework and the teacher career-promotion system. This provides the clearest evidence that institutionalisation is both possible and sustainable. Acquired competences and freely available materials are also likely to persist by their nature, while some individual trajectories (notably alumni who have gone on to become tutors, trainers and multipliers) now continue independently of the project.

By contrast, the networks and international dimension that participants value most remain fragile. The coordinator network was built over two decades largely through the personal contacts of one long-serving officer, and his retirement at the end of 2022 resulted in a documented loss of continuity. A structured alumni mechanism was absent for most of the phase; a database of former participants became operational at the end of 2025 and is beginning to support follow-up in the capacity-building strand, but it does not yet amount to the fuller alumni architecture the evidence calls for (see Recommendation R4). Similarly, the stakeholder taskforce agreed at the 2024 university, which was intended to meet every two months and operate through a shared calendar, did not become functional: only two meetings were held, the planned platform was never established, and staff turnover further disrupted continuity.

Draft terms of reference for coordinators have existed since 2021 but remain unused. While their formal adoption belongs in the recommendations, the evidence points towards the difference they could make to be stated clearly. Signed terms of reference would turn network membership from a largely personal understanding into an institutional relationship. They would define reciprocal expectations, including what coordinators are expected to contribute and what support they can expect in return; provide continuity through personnel changes on both sides; offer an induction reference for new coordinators;

establish a basis for minimum standards in campaign reporting; and make the coordinator role more legible to the institutions in which it is situated, particularly where a ministry or national platform must justify that role internally. The trade-off is the degree of formalisation that management has so far hesitated to introduce. The evidence from this phase, however, suggests that the costs of informality are higher (dependence on individuals, uneven continuity and weak institutional memory).

Further gains in sustainability are already articulated by the governance layer: moving from a co-funded project model towards a formalised and nationally budgeted component of member states' education systems; embedding the work through accredited teacher-training modules; developing a self-sustaining network; and consolidating national relays with ministries. The sub-granting mechanism foreseen for the follow-on phase also provides a natural vehicle for developing the alumni and follow-up structures whose absence is one of the clearest gaps identified by the evidence.

S2. *To what extent was the management of relationships with regional partners and among stakeholders effective in establishing real partnership and ownership?*

The programme was effective in establishing genuine partnership and ownership where engagement was sustained, but less so where relationships were more peripheral or intermittent. Co-ownership of the intercultural-dialogue strand (the Tutored Online Course on Intercultural/Interfaith Dialogue and its Human Rights Education for Youth successor editions) is both genuine and well established. Partner organisations independently described content being co-designed across successive editions, joint selection decisions, regular information-sharing and mutual respect for institutional mandates. Partners also treat the course as part of their own programming, indicating that ownership extends beyond participation in delivery.

The UN partnership on the environment course was productive, but the relationship was less close and sustained. The partner's substantive design inputs were taken up, demonstrating genuine influence over the course, but the relationship itself was experienced as distant: no debrief took place, and continuation remains uncertain. The pan-African partnership generated institutional effects extending beyond the programme which interviewees attributed specifically to the NSC, including mediation between youth platforms, reconciliation with continental structures and the accession of a new member state. However, the agreed mechanisms intended to sustain follow-through subsequently lapsed.

A similar pattern is visible within the wider network. Ownership is strongest where partners are engaged continuously: coordinators who implement the campaign describe a genuine sense of ownership over it. Participatory governance, however, has weakened. The campaign working group was not convened in the final documented year, a thematic choice made by the network was overridden, and interviews suggest that expansion has increasingly been prioritised over depth of engagement.

Government ownership follows the pattern analysed under AV1: it is strong where a ministry anchor or pre-existing institutional relationship exists, but limited elsewhere. Overall, relationship management is therefore a programme strength where sustained investment has been made. Its principal weakness is not the quality of those relationships but their uneven depth and continuity, which have been further affected by the staffing discontinuities documented under E2.

3.6 Added value

AV1. *What is the added value of the NSC in its implementation and methodology (the quadrilogue approach)?*

The NSC has a clear and distinctive added value, recognised by external partners, but the evidence suggests that this added value lies less in the quadrilogue approach than project documents imply. Partner organisations highlighted the NSC's Global South mandate, embedded in its institutional identity, as unusual among European institutions and as a response to the recurring risk that European capacity-building could be perceived as extractive by those on the receiving end. One interviewee contrasted its role starkly: "We see other institutions as bigger, but the Council of Europe as more effective." Interviews also credited the NSC with a mediation role in Africa-Europe youth cooperation that no other organisation was seen to perform.

Across informants, three further sources of added value recur: the Council of Europe institutional anchor; its legally binding human-rights framework and standards recognised across 46 member states; and the NSC's neutrality as a convenor. Together, these give the NSC a combination of legitimacy, normative authority and relational reach that larger organisations do not necessarily replicate. This finding is also consistent with the independent benchmark literature. The UNDP Independent Evaluation Office's 2025 synthesis of youth programming finds that platforms create particular value when they connect young people to sustained follow-up support and meaningful roles in decision-making—areas in which this programme is strongest at its core, but less consistent at its margins.[5]

The quadrilogue requires a more qualified assessment. Project documents present it as a working method through which governments, parliaments, local and regional authorities, and civil society act together, and identify it as a principal source of the programme's added value. The evidence collected across stakeholder groups indicates that the model does not operate in practice as described. Management applies an institution-wide standard under which an event is considered satisfactory when two of the four stakeholder groups are represented, and confirmed that, despite continued efforts to engage them, the parliamentary and local- and regional-authority components are not effectively involved in iLEGEND III. Interviews at donor and coordination levels identified no observable quadrilogue practice beyond this reduced standard. Governance contributions describe the same imbalance in more measured terms: civil society is the most consistently mobilised component; government engagement is uneven; and parliaments remain weakly involved in taking up the outputs generated by civil society. Interviewees across the network reached a similar conclusion, including the observation that the model "works in theory but in practice still needs to be proven". The 2022 evaluation identified the same difficulty and described it as a perennial challenge.

The issue is therefore less one of delivery failure than of a mismatch between the methodology formally presented and the model actually implemented. The operative standard was largely achieved, but it is not the four-way model set out in the project documents. The evidence instead points to several more pragmatic configurations that can generate results. In Morocco, cooperation works through relationships between the civil-society coordinator and ministries that pre-date the programme. In Montenegro, ministry-based coordination, combined with a civil-society counterpart, has helped translate programme participation into institutional embedding. In Serbia, engaged ministry officials have supported civil-society-led campaign activity. These cases suggest that the nature of the national anchor

[5] https://www.undp.org/sites/g/files/zskgke326/files/2026-06/youth_en.pdf

strongly shapes which stakeholder groups are reached and how ownership develops. However, a functioning four-way quadrilogue was not observed in any of the documented country cases.

The NSC's added value therefore lies primarily in its institutional position, convening legitimacy, normative framework and capacity to bridge actors across regions, rather than in demonstrated application of the quadrilogue as a four-way governance methodology. Management itself identified this issue, together with the relationship with GENE, as one of the most urgent points to address in the new cycle.

3.7 Human rights and gender equality (cross-cutting)

Human rights are an organising principle of the centrally produced content rather than an add-on. Recent courses are grounded in European human-rights law and jurisprudence; the environment course frames environmental protection as a rights issue and was co-developed with UN expertise; and safeguarding practices at events were consistently described in interviews as rigorous. **Gender equality, by contrast, is not systematically embedded where delivery is decentralised, and the evidence of this gap is cumulative**.

Viewed through the human rights approach of the Council of Europe project management methodology, these strengths are primarily normative: they concern standards and content. The procedural dimension is thinner. As documented under E1, participation is visible in convening practices and non-discrimination in selection rules, but no grievance mechanism or participant-feedback loop beyond satisfaction forms was identified, and human-rights outcomes are not systematically tracked. These process aspects are reflected in Conclusion 7 and addressed by Recommendation R9.

Across interviews and written contributions, questions about how gender is addressed in the programme repeatedly produced little substantive evidence. Reporting templates do not provide for sex-disaggregated data, while the anonymity of participant evaluation forms prevented the disaggregation promised in the logframe for three consecutive years, as the progress reports themselves acknowledge. Management notes that disaggregated reporting is a standing requirement of the DEAR programme and that the relevant terminology was reviewed in 2026 to reconcile the Council of Europe and European Commission project-management platforms; the underlying gap, which management itself acknowledges, is that the gender dimension has not been addressed with the depth the programme's own commitments require. The design logframe's indicators do not address human rights and inclusion. The only such indicator appeared in the 2023 report by mistake (see Section 3.3, E1) and was below target. The 2022 evaluation had already recommended systematic gender mainstreaming supported by indicators, but this was not implemented. Recent feedback received by the evaluation from the governance layer calls, in closely similar terms, for gender to be treated as a transversal principle across governance, speakers, content, activities and evaluation. The evidence suggests that the gap extends beyond reporting to programme requirements, implementation expectations and measurement.

On the participation of vulnerable groups specifically, the programme cannot currently demonstrate either their inclusion or their exclusion. No objective, target or indicator tracks vulnerable-group participation (the sole inclusion indicator appeared by mistake in the 2023 report); the organisational targeting model, which management clarified during the evaluation (see R2), relies on partner organisations to reach such groups without systematically recording who is reached; and the forms that might capture relevant information are anonymous. This is not evidence that vulnerable groups are absent. It is evidence that the programme does not currently know, and cannot demonstrate, the extent to which they participate. Recommendation R9 addresses this gap through the introduction of an explicit target and indicator.

Access presents a further equity challenge. The evidence documents visa failures affecting participants from Sub-Saharan Africa and North Africa, as well as at least one partner expert. Nomination windows for residential activities were sometimes too short to reach candidates beyond capital cities who already held passports, whereas public calls for courses generally remained open for three to four weeks. Flagship courses also continue to be delivered only in English. Interviews raised a broader concern about who is able not only to attend but also to shape delivery, noting in particular that no Sub-Saharan organisation currently organises activities at the Mollina university.

These findings should be read against an important design dimension. The funding instrument limits non-European participation to no more than half of participants and gives priority to European participants, reflecting DEAR's status as a European programme; the limited involvement of Sub-Saharan organisations in shaping delivery is likewise conditioned by these rules rather than by NSC choices, and under the next DEAR cycle the required ratio is expected to be set at 80 percent European to 20 percent non-European participants. Language choice has also been used deliberately to manage this balance: English-language delivery has helped maintain high European participation after French-language editions attracted predominantly African applicants. This clarification is analytically important as it shows that some access constraints are not simply implementation oversights but arise from the programme's underlying funding and geographic design. That does not remove the equity question; it sharpens it and makes it a matter for explicit consideration in the design of the next phase.

There are also important countervailing strengths. Council of Europe invitation letters can facilitate visa access when issued sufficiently early, and participants and partners repeatedly described the NSC's practical support in navigating access barriers as exceptional. Management also retains selection records documenting the range of countries and participant profiles reached. Publishing these data in aggregate, as recommended by the evaluation, would provide a stronger evidence base for assessing who the programme reaches and where access gaps remain.

4. The theory of change, tested

Mapping the findings against the reconstructed theory of change yields a clear conclusion: the causal logic is strongest at the lower levels and becomes progressively weaker toward the top. The table below summarises this assessment.

Link or assumption	Verdict	Basis
Inputs and activities to outputs and immediate outcomes (all four pathways)	Validated	Delivery, completion, satisfaction and network formation documented across sources (sections 3.3, 3.5)
Individual learning to wider application (the uptake step)	Plausible, unmeasured	Traceable individual trajectories; no follow-up structures or measurement (E1, E2, S1)
Convening to four-group quadrilogue engagement	Not validated as designed	Two groups not involved; two-of-four standard operated; working two-actor configurations documented (AV1)
Campaign to public reach	Partially validated	Real, growing national embedding; reach quantifiable to order of magnitude only (E3)
Public reach to attitude change	Untestable	No attitude or behaviour measurement exists (E1, E3)
Capacity to institutionalisation	Demonstrated once	Curriculum and career-system embedding in Montenegro; dependent on ministry anchor (S1)
Assumptions: political will, visas, timely resourcing, multipliers	Partially breached	Political-context failures in two countries; documented visa losses; staffing gap; multiplier assumption never resourced (E2, S1, 3.7)

Table 7. *The reconstructed theory of change against the evidence.*

The implication for the next phase concerns programme design: the application step needs to be designed rather than assumed, institutionalisation should be elevated from an implicit expectation to an explicit result with corresponding indicators, and the inputs required to achieve these outcomes should be realistically costed.

5. Conclusions

This section presents seven analytical conclusions that synthesise across the findings in Sections 3 and 4. The conclusions are organised around cross-cutting analytical threads rather than by evaluation criterion, because the substantive analytical work of the evaluation operates across criteria rather than within them.

Conclusion 1. The programme is relevant and delivers its outputs reliably, while results beyond the activity level remain plausible but unproven (R1, R2, E1).

Delivery, participant satisfaction, completion, and reach are documented beyond reasonable doubt at the output level, and individual trajectories provide credible evidence that the intended results cascade operates in practice. However, the monitoring system does not capture whether learning is translated into application at scale, and no data were generated for the logframe's impact-level indicators. The theory of change is therefore well supported at its base but becomes progressively weaker at the transition from individual learning to broader application. Relevance itself rests on an ageing evidence base: the renewed, target-group-specific needs analysis recommended by the 2022 evaluation was not carried out, and a relatively uniform offer serves less well countries at earlier stages of global-education development (R2).

Conclusion 2. The quadrilogue envisaged in the project documents does not reflect the programme's actual operating model, and this gap has persisted across phases (AV1).

Two of the four stakeholder groups are not engaged at all, while the de facto internal standard requires the presence of only two. In practice, the programme operates through a range of productive two-actor configurations shaped by national contexts rather than through a consistent four-party model. After three phases, this ambiguity persists and the programme should either invest in making the full quadrilogue operational or formally reframe its model and commitments to reflect how it actually works.

Conclusion 3. The Global Education Week is the programme's flagship national-level result, but its measurement system does not adequately capture or substantiate its significance (E3).

The campaign's growth and increasing national institutionalisation are real and well documented. However, the published reach figures are best understood as order-of-magnitude estimates rather than precise counts, despite being reported with a level of precision the underlying data cannot support. A result of this scale and value warrants a transparent and robust data chain capable of withstanding scrutiny.

Conclusion 4. The programme's durability currently depends heavily on individual commitment, while institutionalisation remains achievable but only partially realised (S1, S2).

The network's history, unused terms of reference, absence of alumni structures, and unsuccessful taskforce all point to the same underlying pattern: continuity has not yet been secured through formal institutional arrangements. Montenegro demonstrates that a more durable model is possible, and the governance layer has already defined the intended end state: integration into national education systems with dedicated domestic budget support.

Conclusion 5. The programme's principal efficiency constraint is an activity schedule too dense for the efficient use of its staff resources rather than weak management (EF1, E2).

The evidence is consistent: an eighteen-month vacancy, staff time budgeted below the effort the portfolio actually requires, a challenging co-financing contribution, and a heavily back-loaded final year. As detailed under EF1, the consequences are a delivery model that sometimes draws on existing capacity rather than adding to it, substantial but largely invisible in-kind co-financing, operational dispersion, and staff exhaustion. These pressures were foreseeable: the 2022 evaluation had already concluded that the staffing structure represented the minimum required. Within these constraints, however, financial management has been sound and subject to satisfactory audit.

Conclusion 6. The programme generates distinctive added value that is clearly recognised by external partners, but this value derives from the NSC's Global South mandate, its neutral convening role, and the normative authority of Council of Europe standards rather than from the quadrilogue model itself (AV1, C3).

Partners with access to multiple potential partners specifically credited the programme's ability to bring actors together on equal terms and to mediate across institutional and political divides. This convening credibility is therefore a strategic asset and should be made an explicit foundation of the next phase. Consolidating it also requires clarifying the division of labour with adjacent European global-education networks whose mandates are converging with the NSC's; management itself has identified this relationship as one of the most urgent issues for the next cycle (C3), and Recommendation R8 addresses it.

Conclusion 7. Human rights are strongly embedded in the programme's substantive content from a normative perspective, but less so procedurally, and gender equality is not systematically integrated into its delivery; these gaps are consistently documented across sources and phases (E1, Section 3.7, cross-cutting).

The absence of inclusion indicators, three consecutive years of unsuccessful data disaggregation, repeated absence of gender considerations across reporting sources, and the non-implementation of a 2022 recommendation together point to a persistent institutional weakness rather than an isolated monitoring gap. The same applies to the procedural dimension of the human rights approach: accountability and feedback mechanisms beyond satisfaction forms are absent, and human-rights outcomes are not tracked (E1). A future phase should address this explicitly through clear accountability, measurable commitments, and systematic gender-responsive monitoring.

6. Lessons learned

This section distils four transferable lessons from iLEGEND III for wider application. They differ from the recommendations in Section 7 in two respects: first, they extend beyond decisions concerning Phase III and beyond NSC-iLEGEND itself, with relevance for comparable initiatives; second, they identify broader patterns and principles rather than programme-specific actions.

Lesson 1. Relationship-based networks can expand reach at low cost, but without institutional scaffolding they remain fragile.

The cost of relying on personal relationships often becomes visible only later, particularly during staff turnover and handovers. Signed roles, documented processes, institutional ownership, and overlap between outgoing and incoming personnel are not bureaucratic burdens; they are the infrastructure required for continuity.

Lesson 2. Online delivery can extend access to the Global South at a fraction of the cost of residential provision, but access is decided before enrolment: nomination timelines, language, outreach and who is invited to organise or nominate determine who can actually benefit.

These upstream choices can matter as much as who ultimately applies; only deliberate inclusion measures convert potential reach into actual participation.

Lesson 3. For empowerment objectives, small grants can be more effective than tightly specified deliverable-based contracts when appropriate safeguards are in place.

They create space for local ownership, adaptation, and initiative that conventional contracting can constrain. Removing a mechanism that the programme's own evaluation had identified as effective therefore risks trading mission effectiveness for administrative convenience.

Lesson 4. A convening model that is not matched by adequate resourcing loses credibility the longer the discrepancy persists.

Where a programme commits to a specific participatory mechanism, it must either provide the resources needed to make that mechanism operational or redefine the commitment to reflect the standard it can realistically sustain. An arrangement that meets an internal minimum while repeatedly falling short of the public promise serves neither clarity nor credibility.

7. Recommendations

This section presents ten recommendations. Unless otherwise stated, the addressee is the NSC. Each recommendation is assigned a priority level: high priority for issues that are structural or time-critical for the design of the follow-on phase, and medium priority for important improvements that can be sequenced more flexibly; possible actions that did not reach at least medium priority were not retained as recommendations. Recommendations bearing on the design of the follow-on phase are marked (next phase) and are time-critical, as that design is currently under negotiation. Each recommendation identifies the findings and conclusions on which it is based.

Recommendation 1. Resolve the quadrilogue coherently.

Addressee:	**Priority:**	**Timing:**	**Basis:**
NSC	*High*	*Next phase; repeat*	*AV1; Conclusion 2*

The next phase should make an explicit choice about the programme's operating model and reflect that choice consistently in its design documents. Either the quadrilogue should be resourced as a genuine four-group model, with defined roles and adequate provision for parliamentary and local-authority participation, or the programme should adopt an explicit convening standard that formalises the two-of-four minimum the NSC already applies institution-wide, based on the configurations that demonstrably work, including civil-society and government partnerships, ministry-based or dual coordinators, and engaged ministry technical staff. The programme should no longer commit in its design documents to a model it does not operate in practice.

Recommendation 2. Institutionalise the network.

Addressee:	**Priority:**	**Timing:**	**Basis:**
NSC	*High*	*Immediate and next phase*	*S1, S2*

The NSC should apply coordinator terms of reference, which have existed in draft since 2021, and formalise arrangements for entry, succession, handover, and continuity. In parallel, it should pursue the institutional trajectory already articulated by the programme's governance layer: accredited teacher-

training modules, curriculum integration where a ministry anchor exists, and progression towards nationally budgeted status. The existing cases should be documented and used as a practical model of what this trajectory can look like (e.g., Montenegro).

Recommendation 3. Restore small grants with the safeguards already proven to work.

Addressee:	Priority:	Timing:	Basis:
NSC	*Medium*	*Next phase; repeat*	*EF1, S1, E1*

The next phase should reintroduce a small-grants window alongside service contracts, at the modest scale coordinators describe as decisive, ranging from hundreds to low thousands of euros. The mechanism should retain the safeguards already tested in previous phases, including written guidance on recurrent compliance errors and an induction session at the point of award. The programme's own record, with only one non-compliant grantee across past schemes, indicates that small grants can be managed with proportionate controls while preserving the flexibility required for empowerment objectives.

Recommendation 4. Build an alumni architecture.

Addressee:	Priority:	Timing:	Basis:
NSC	*Medium*	*Next phase*	*S1, E1, C1*

The next phase should create a structured follow-up offer that gives participation a pathway beyond the initial activity, building on the database of former participants operational since the end of 2025. This should include an alumni community of practice, systematic progression of strong participants into tutoring and facilitation roles, recognition mechanisms for outstanding coordinators and teachers, and routes into relevant international processes. Demand for such continuity is documented across stakeholder groups, and it represents a comparatively low-cost means of multiplying programme effects.

Recommendation 5. Strengthen measurement and data quality.

Addressee:	Priority:	Timing:	Basis:
NSC with DIO	*High*	*Immediate; repeat*	*E3, E1, Annex 1*

The programme should establish a verifiable campaign data chain based on partner-completed returns, a documented compilation method, version control, verification of the largest reported entries, and transparent disclosure of denominators and response rates. It should also foresee an inclusion indicator, report completion against both selected and active-participant denominators, and introduce two or three realistic outcome-level indicators supported by an explicit measurement plan. The next evaluation should not inherit another results framework whose higher-level indicators remain unfed.

Recommendation 6. Optimise Southern access within the donor's participation requirements.

Addressee:	Priority:	Timing:	Basis:
NSC	*Medium*	*Next phase*	*Section 3.7, E2, EF1*

The geographic participation ratios are set by the funding instrument, not by the NSC; under the next DEAR cycle the required balance is expected to be 80 percent European to 20 percent non-European participants. Within these parameters, the next phase should improve the quality of access for participants from the Global South: timely visa support letters; extension of the existing three-to-four-week public-call standard to nomination windows for residential events; language-accessible editions of flagship courses where compatible with the required balance, with ratios managed through selection rather than through language barriers; greater diversification of venues and organisers, including Sub-Saharan organisers within the universities; and travel arrangements that avoid requiring African

participants to leave the continent unnecessarily in order to return to it. Where a desirable measure exceeds what the donor's requirements allow, the NSC should raise it with the European Commission in the design dialogue for the next phase rather than abandon it. Implementation should be tracked through participation data.

Recommendation 7. Consolidate the portfolio.

Addressee:	Priority:	Timing:	Basis:
NSC	*High*	*Next phase*	*EF1, E2*

The next phase should plan fewer, deeper activities, supported by front-loaded calendars and staffing assumptions that reflect the effort the portfolio actually requires. The evidence found by the evaluation is consistent: the number and range of activities is not matched to the efficient use of the staff resources allocated to them, internal sources converge on operational dispersion and exhaustion, and the back-loading of the final year has been acknowledged as a planning error. Consolidation should therefore be treated as a design principle rather than allowing the same pressures to recur.

Recommendation 8. Agree a complementary division of labour with adjacent European global-education networks.

Addressee:	Priority:	Timing:	Basis:
NSC with the EC	*Medium*	*Before the next cycle*	*C3; Conclusion 6*

In coordination with the European Commission, the NSC should agree and document a complementary division of labour with the adjacent European global-education networks whose mandates are converging with its own, clarifying respective constituencies and convening roles. The arrangement should recognise the NSC's comparative reach among civil society, youth actors and partners beyond Europe, provide for reciprocal participation in relevant events, and be reflected in the organisations' next-phase documents. The evaluation found no duplication of activities to date, but the donor has signalled the need to avoid it, and management itself has identified this relationship as one of the most urgent issues for the next cycle (C3, Conclusion 6).

Recommendation 9. Mainstream gender, strengthen the procedural human rights approach and introduce an explicit vulnerable-group lens.

Addressee:	Priority:	Timing:	Basis:
NSC	*High*	*Immediate; repeat*	*Section 3.7; E1; Conclusion 7*

The programme should introduce a gender marker for all activities, data systems that permit meaningful disaggregation while protecting respondents, at least one gender-responsive indicator for each programme strand, and gender criteria for the selection of speakers, trainers, and content. In parallel, it should establish an explicit objective, target, and indicator for the participation of vulnerable groups, working through relevant partner organisations. To strengthen the procedural dimension of the human rights approach, it should also introduce a participant-feedback and complaints channel beyond end-of-activity satisfaction forms and track human-rights-related outcomes systematically (E1, Section 3.7).

Recommendation 10. Conduct a renewed needs analysis for the next phase.

Addressee:	Priority:	Timing:	Basis:
NSC	*Medium*	*Next phase; repeat*	*R2; Conclusion 1*

The 2022 evaluation recommended a renewed, target-group-specific needs analysis; it was not carried out during this phase, and the consequences are visible in the findings: a relatively uniform offer that serves less well countries at earlier stages of global-education development, an organisational targeting model that is communicated unclearly, and limited knowledge of who is ultimately reached (R2, Section

3.7). The NSC should therefore commission a renewed needs analysis explicitly linked to target groups and differentiated by levels of national maturity, establishing who the programme currently reaches, where unmet needs remain, and how needs vary across country contexts. The findings should directly inform the next phase's geographic and thematic choices, so that future priorities are grounded in current evidence rather than carried forward primarily through continuity.

Annex 1: Logframe indicator record and quality assessment

This annex presents the design-quality assessment of the results framework and the indicator-by-indicator reporting record across the three progress reports, substantiating the findings under E1 and E2 (Section 3.3).

Quality assessment

The framework is formally complete and internally coherent on paper, but managerially weak in practice. Indicators were populated only partially and asymmetrically during implementation. Data actually collected included end-of-activity evaluation forms, which consistently captured satisfaction and stated intention; some three-month follow-up surveys, introduced from Y2 onwards; certification lists; attendance figures; and self-reported data from GEW coordinators. In Y1, OVI 3 was reported as "N/R — no data collected yet".

Several data elements foreseen in the framework were never adequately generated:

- **Disaggregation:** Anonymous evaluation forms prevented disaggregation by gender, age and region throughout the first three years. The Y3 report states that a new Privacy Notice will address this "in future reporting", meaning effectively only during the final year of implementation.
- **Response rates and denominators:** These were never disclosed. Percentages such as "94% of respondents" are not interpretable without the number of respondents and the response rate.
- **Verifiable GEW data:** Management described the reporting system as a "chain of trust". The 2023 dataset consists of pre-event planning tables rather than evidence of actual participation. The Y1 figures of 312,667 participants and 7,000 activities therefore have no documentary basis. Even where data existed, version control was weak: the published 2025 figure of 430,717 participants differed from the country-sheet total of 397,380, while the France data shifted between estimated and actual values without a clear audit trail.

Taken together, these weaknesses mean that the framework provided a formally structured basis for reporting, but not a sufficiently reliable basis for assessing performance. It tracked implementation activity and participant perceptions reasonably well, but could not substantiate several of its headline results or demonstrate whether outputs translated into behavioural, institutional or policy change. Against this background, its quality against the SMART criteria is assessed below.

Specific: partially. Most indicators are sufficiently precise about what should be counted. However, several central constructs are undefined, including "initiative at interregional level", "early follow-up" and "political initiative". This created room for construct substitution under OVI 3 in Y3 and contributed to contradictory claims concerning OVIs 4 and 5 across reporting years. Indicator wording also changed between reports: OVI 19 appears in three different formulations. The Y2 report further contains apparently transposed targets, with OVI 18 showing a satisfaction target and OVI 20 showing a participant-count target, as well as unresolved placeholders such as "TBD (Ref. year XXXX)" during the second year of implementation.

Measurable: weak in practice. The indicators fall into three broad categories:

1. Verifiable counts, such as meetings, certificates, language versions and participating countries, which were generally measured adequately.

2. Self-reported percentages derived from end-of-activity and three-month follow-up forms. Although these data were collected, the reports never disclosed sample sizes or response rates. The forms were anonymous, meaning that the planned gender, age and regional disaggregation repeatedly failed ("disaggregation not available", Y2, OVIs 9 and 10; "not possible", Y3, OVI 20). Respondents were also a self-selected group of participants who completed the activities, creating a strong likelihood of positive-response bias.
3. GEW reach data under OVI 8, which were self-reported by national coordinators through the "chain of trust". The NSC itself acknowledged that there were "no direct monitoring or verification mechanisms at the NSC level" (Y3, p. 26, Challenges).

The impact-level indicators were not measured.

Achievable: yes, but insufficiently demanding. Satisfaction and intention thresholds of 70-80% were exceeded by 10-25 percentage points each year, partly because they were measured among self-selected activity completers. Two targets merely reproduced their baselines: OVIs 15 and 22 both set an 80% target against an 80% baseline. The impact target required only one additional political initiative over four years. The target of 30 GEW countries was only five above the baseline of 25 and was exceeded by 15 countries. The only clearly stretching target (620,000 people reached through GEW, approximately twice the baseline) is also the indicator supported by the least verifiable data.

Relevant: predominantly output- and perception-oriented. Of the 26 indicators, eleven concern activities or reach, nine concern satisfaction or stated intention, and only six relate to behavioural or policy change: OVIs 1-6. Of these six, two were not measured, one underwent construct substitution, one was reported as "Not achieved", and whether iLEGEND III contributed to the standard-setting instruments reported under OVI 5 (including the RFCDC-ESD guidance document) was not verified. The framework does not measure whether the quadrilogue was effectively convened, whether gender-related outcomes improved, or whether training led to changes in practice beyond participants' self-reported responses after three months. No substantive gender indicator remained in the framework; only disaggregation requirements were included, and these were not fulfilled.

Time-bound: formally, but not always operationally. The PMM identifies 2026 as the reference year for final achievement. However, unresolved "TBD" targets in Y2 and ambiguity between annual and cumulative reporting weaken the time-bound character of the framework. OVI 8 is labelled "Annual", yet Y3 assesses a result of approximately 430,000 against the 620,000 target as "On track", without clarifying whether the target is annual or cumulative. Similarly, OVIs 14, 17 and 18 are cumulative indicators but are reported as annual figures without presenting running totals. This makes it difficult to determine whether progress is genuinely on schedule.

Indicator inventory (26-OVI framework; numbering as used in Y2/Y3 reports)

Values from the Y3 Annex 1 (pp. 33-53) were cross-checked against the main report texts: Y1 (pp. 6-11), Y2 (§2.2 tables), and Y3 (pp. 7-17). "Status" reflects the report's own assessment, based on Y3 unless otherwise indicated. "Achievability" reflects the evaluation's assessment of whether the target can be reached by the end of 2026.

OVI	Indicator (abridged)	Baseline / Target	Y1 (2023)	Y2 (2024)	Y3 (2025)	Status per reports	Assessment of achievability by Dec 2026
1	Impact: Nº political initiatives by decision-makers	2 / 3	—	—	—	never reported	Unmeasured; assessment devolves to this evaluation
2	Impact: Nº initiatives by youth/CSOs	0 / 6	—	—	—	never reported	Unmeasured; same circularity
3	SO1: % developing interregional initiative post-participation	0 / 70%	N/R "no data collected yet"	65.3%	"92%"	Y2 Underachieved; Y3 Achieved	Unknowable; Y3 measured a different construct ("shared knowledge/skills", not "developed an initiative"). Only clean data point (65.3%) is below target
4	SO2: Nº recommendations with early follow-up	0 / 1	narrative claim	"1" (seminar conclusions → Council of Europe WG)	0	Y2 Achieved; Y3 Not achieved	Possible but contradictory across years; hinges on 2026
5	SO2: Nº standard-setting instruments proposed by Council of Europe steering committees	0 / 1	"1: Education Strategy 2024-2030"	NA	"1: RFCDC-ESD Guidance document"	Y3 Achieved	Achieved on paper, but attribution of either instrument to iLEGEND is untested
6	SO3: % practitioners applying GE competences (3-month survey)	70% / 80%	74% (82%/66% by course)	80%	78%	Y3 "At risk of slightly underachieving"	Borderline; self-report only
7	SO3: Nº training institutions certifying courses	0 / ≥1	UCV partnership	"concrete steps" for ECTS	same + Helsinki for 2026	On track (3 yrs running)	At risk: "concrete steps taken" ≠ certification; identical wording since Y2
8	SO4: (Annual) Nº people reached in GEW	300,000 / 620,000	312,667 ("not exhaustive, 6 countries missing")	">300,000 in person"	">430,000" (5,600 activities, 45 countries)	Y2 Achieved(!); Y3 On track (p. 15)	Not credibly achievable/verifiable: needs +44% on a figure our re-analysis already shows overstated (430,717 vs 397,380 country sheets; France entry; 2023 figure without documentary basis)
9	Oc1: % willing to cooperate interregionally	0 / 70%	83.3%	95.9% (no disaggregation)	93%	Achieved	Achieved but see selection bias below
10	Oc1: % expanded network	0 / 80%	66%	89.7%	94%	Achieved	Achieved (Y2 value printed "89,7,8%")
11	Oc2: Nº NSC interventions relaying results at CDEDU/CMJ	1 / 9	2	4	2	On track	Achievable (cumulative approx. 8; verifiable in minutes)
12	Oc3: % declaring intention to use GE competences	70% / 80%	85%	90.5%	95%	Achieved	Achieved (intention only)

OVI	Indicator (abridged)	Baseline / Target	Y1 (2023)	Y2 (2024)	Y3 (2025)	Status per reports	Assessment of achievability by Dec 2026
13	Oc4: Nº new organisations joining GEW preparations	4 / 7 (≥3 new)	1 first-time country	"ongoing discussions"	9 new orgs (listed)	Y3 Achieved	Achieved
14	Op1: Nº events supported (universities)	0 / 6	≥2 (UYD, AUYD 2023)	2	1	On track	Achievable if 2026 UYD/AUYD run (back-loaded year risk)
15	Op1: % satisfied with events	80% / 80%	96.3%	94.2%	90%	Achieved	Achieved every year: target = baseline
16	Op1: % exploring synergies/good practices	0 / 70%	93.5%	98.5%	92%	Achieved	Achieved
17	Op2: Nº events organised/coordinated by NSC	6 / 10 (3 seminars, ≥5 online talks, GEN AMs, 1 Congress)	4 (incl. 2 online talks)	3	1	On track	Borderline: cumulative approx. 8; headline 10 possible via 2026 congress, but sub-targets (≥5 online talks: only 2 traced) look unmet
18	Op2: Nº participants state/non-state events	310 / 670 (30% state)	314 (incl. UYD's 237)	102	43 (excl. UYD's 210)	On track	Doubtful and incoherently counted: Y1 counted UYD participants, Y3 did not; cumulative approx. 460 needs a very large 2026 congress; state-actor share far off 30%
19	Op2: Nº CDEDU/CDEJ-CMJ meetings attended	2 / 10	"2 plenary/committee"	2 (Y2 self-labelled Underachieved)	2	On track	Achievable (approx. 8 + 2026); note target text drifted: "attended" – "facilitated by" – "with the contribution of"
20	Op2: % rating dialogue high quality	70% / 80%	78%	92%/95%	83-87%	Achieved	Achieved; disaggregation impossible: "form was completed anonymously" (Y3 p. 11)
21	Op3: % trained certified	70% / 80%	76.6%/83.3%	79.3%/81.6%/100%; MOOC 43/71	77.5% online; ToT 88%; MOOC 44.5%	On track	Borderline: tutored courses hover just under 80%; MOOC far below and excluded from the headline
22	Op3: % satisfaction with courses	80% / 80%	95% (100%/90%)	94%	96%	Achieved	Achieved: target = baseline
23	Op3: Nº language versions of tools	1 (EN) / ≥2 (EN+FR)	3 (SPOT EN/FR/AR)	GEG EN/FR/ES(+SC/AR pending)	multiple	Achieved	Achieved for GEG/MOOC; flagship tutored courses remain EN-only (ICD/IFD, Environment, Y3 p. 14)
24	Op3: Nº schools/NFE facilities involved	2 / 10	11 (GEW school debates)	12 (course schools)	3	Y3 "On track"	Achieved cumulatively (approx. 26) but the counting basis changed each year (GEW debates vs course enrolment)
25	Op4: Status of GEW webpage	operational / annually updated	updated	updated	updated	Achieved	Achieved (trivially)
26	Op4: Nº countries with GEW activities	25 / 30	26	28	45	Overachieved	Achieved but approx. 15 of the 45 (Brazil, Colombia, Haiti, Chad, DRC…) sit outside the action's EU/candidate-country impact population, so overachievement partly measures scope drift, not depth

Annex 2: Evaluation matrix

The matrix below was approved at inception. It was implemented as designed, with the following notes: two written contributions replaced the planned governance interviews; the parliamentary and local-authority interviews foreseen under AV1 were not pursued after the Secretariat confirmed that they are not effectively involved in the programme; and the data re-analysis (Annex 4) supplied the quantitative evidence foreseen under E1, E3 and EF1.

Criterion	Evaluation question	Sub-questions & judgement criteria	Indicators / measures	Data-collection instruments	Data sources	Data analysis
Relevance	R1. To what extent is the intervention design relevant to the NSC mandate and to the EC DEAR programme (per the logframe / DoA)?	Alignment of design with the NSC Statute and mandate, the Council of Europe Strategic Framework and its three pillars (human rights, democracy, rule of law), Council of Europe GE policy and pedagogical frameworks (CM/Rec(2011)4, Global Education Guidelines, RFCDC), and the EC DEAR objectives / results framework. Internal consistency of objectives with the iLEGEND III logframe.	Degree of documented alignment with (i) NSC mandate, (ii) DEAR results framework, (iii) Council of Europe Strategic Priorities 3, 4, 5, 11, 12 and SDGs 4, 5, 10, 13, 16. Explicit cross-references in design documents.	Document review; KIIs	DoA; logframe; Council of Europe Strategic Framework; CM/Rec(2011)4; SG/Inf(2020)34; ExCom synopses; DEAR docs & C(2024)8167 [request]; iLEGEND II evaluation [request]. NSC Secretariat (Gulletta); EC DG INTPA.	Qualitative alignment mapping; triangulation.
Relevance	R2. To what extent do the objectives address identified needs of target groups and the capacities of stakeholders?	Adequacy of needs analysis; appropriateness of targeting (youth 18–30/35, practitioners, CSOs, state actors); responsiveness to capacity constraints; relevance to EU candidate / potential-candidate and partner countries (MENA, Sub-Saharan Africa).	Evidence of needs analysis informing design; perceived relevance among target groups (existing survey data); share of activities / participants reaching priority audiences and partner countries (disaggregated by sex, age, country group, disability).	Document review; secondary analysis of existing participant surveys; KIIs	Activity reports (UYD, AUYD, HEY, LTTC/TOTC); participant lists & existing Microsoft Forms datasets [request]; DoA. GEN coordinators; trainers; youth multipliers.	Descriptive statistics; thematic synthesis.
Coherence	C1. Are the different activities logically aligned and mutually reinforcing? (internal coherence)	Coherence among the four pillars (partnership building and networking, advocacy and confidence-building, capacity building, and awareness-raising incl. sub-granting); reinforcement of outputs across components; HEY tutor-cascade and sub-grant linkages.	Mapping of inter-pillar linkages; share of participants / organisations engaged in more than one pillar; evidence of cumulative effects across iLEGEND I–III.	ToC reconstruction; document review; KIIs	Logframe; DoA; narrative reports (Y1–Y3); activity reports. Programme team (Gulletta, Giulio); GEN.	Contribution / ToC analysis.
Coherence	C2. To what extent is the intervention coherent with broader policy frameworks? (external coherence)	Coherence with EU NDICI-Global Europe, the European Consensus for Development, the Council of Europe RFCDC, the Council of Europe Youth Sector Strategy 2030 and the 2030 Agenda; consistency with Council of Europe strategic priorities.	Documented references / synergies; consistency of objectives with EU and Council of Europe strategic frameworks.	Document review; KIIs	Council of Europe Strategic Framework; RFCDC; CM/Rec(2011)4; CM(2016)25; DEAR / C(2024)8167 [request]. EC DG INTPA; Council of Europe statutory contacts (CDEDU, CDEJ/CMJ).	Qualitative policy mapping.
Coherence	C3. To what extent does it complement, add value to, and avoid duplication with relevant DEAR / Council of Europe initiatives, other networks and national policies?	Complementarity with GENE, Bridge 47, CONCORD, UNESCO, OECD and UNAOC; synergy via the Council of Europe–UNAOC MoU; avoidance of duplication; links with other DEAR-funded initiatives.	Documented joint actions / synergies; absence of duplication confirmed by partners; references in partners' documents.	Document review; KIIs with external partners	Council of Europe–UNAOC MoU [request]; benchmark docs (GENE, Bridge 47, CONCORD); mission reports. GENE, Bridge 47, CONCORD, UNESCO, OECD, UNAOC focal points.	Benchmarking; triangulation.
Effectiveness	E1. What progress has the project made towards its objectives, and how has it integrated human-rights principles (non-discrimination, participation, empowerment, accountability)?	Achievement of outputs and immediate / intermediate outcomes against logframe targets; integration of the HRA in design, implementation, governance and monitoring; participation of rights-holders (youth, CSOs) in governance bodies.	Achievement rates vs logframe indicators — e.g. Impact: Nº political initiatives (baseline 2 → target 3 by 2026); Nº youth/CSO initiatives (0 → 6); IO1: ≥70% of respondents developing an interregional initiative. Cumulative reach vs iLEGEND I+II (1,050+); share of activities applying HRA; sex/age/region disaggregation.	Document review; secondary analysis of monitoring & existing surveys; KIIs	Narrative reports (Y1–Y3); logframe monitoring data; scoreboards; activity post-evaluations / existing surveys; activity reports. Participants; trainers; GEN.	Contribution analysis; descriptive statistics; thematic synthesis; triangulation. NB scope limit (2026 GEW & Congress excluded); outcome data

Criterion	Evaluation question	Sub-questions & judgement criteria	Indicators / measures	Data-collection instruments	Data sources	Data analysis
						uneven across clusters.
Effectiveness	E2. What are the reasons for achievement or non-achievement?	Internal and external enablers / inhibitors: political will and xenophobic/populist discourse; civic-space restrictions; COVID legacy and format changes; partner capacities; NSC resources and co-funding timeliness; multiplier commitment; quadrilogue balance.	Frequency and salience of enabling / hindering factors identified through triangulation; correspondence with the DoA risk factors.	KIIs; document review; survey open-ended responses (where available)	Narrative-report challenge sections; IFRs & scoreboards (timeliness; 68% of €2m spent and committed by end-2025, IFR-3); ExCom synopses (member withdrawals, budget & strategy context). NSC Secretariat; programme team; GEN; partners.	Qualitative causal analysis; triangulation.
Effectiveness	E3. How effective and valuable has the Global Education Week (GEW) been in raising awareness and promoting GE?	Reach and depth of GEW; contribution of NSC support (toolkit, campaign manager, info sessions); perceived value among national coordinators; trend across the 2023–2025 editions.	Nº of countries, activities and participants in GEW; communication reach / campaign metrics; satisfaction and usefulness scores; trend vs the IO4 outreach baseline.	Secondary data analysis; KIIs; existing survey data	GEW reports 2023 / 2024 / 2025; communication & campaign metrics [request]. GEN coordinators.	Descriptive statistics; trend analysis; thematic synthesis. NB 2026 GEW outside scope.
Efficiency	EF1. To what extent were financial and other resources used as intended and efficiently to achieve results?	Budget execution vs results; cost-efficiency of modalities (online vs residential); timeliness of delivery; quality of management and M&E; use of the NSC quadrilogue structures.	Annual burn rate / implementation rate (68% of €2m spent and committed by end-2025, IFR-3 — to explain); unit cost per beneficiary by modality; share of activities delivered on time; quality of the M&E system.	Financial data analysis (annual snapshots); document review; KIIs	IFRs 1–3; scoreboards (6); annual financial records [request]; narrative reports; ExCom synopses (Programme & Budget 2024–2027). Programme team; NSC finance (Monica).	Descriptive financial analysis; modality benchmarking. (€2m confirmed as total envelope.)
Sustainability	S1. Which achieved results can be expected to continue after the project, and what could enhance sustainability?	Continuation of GEN dynamics, practitioner capacities, advocacy outputs and sub-grantee initiatives; institutional anchoring within the NSC and partner organisations; resourcing of the follow-up cycle (2025–2027).	Evidence of uptake by national coordinators / ministries; integration of GE in national policies; commitments to the follow-up cycle; alumni and network continuity.	KIIs; document review; existing survey data (sustainability items)	Narrative reports; iLEGEND I/II continuity evidence. GEN coordinators; line ministries (via dual reps); CSOs / sub-grantees; EC DG INTPA; NSC governance.	Thematic synthesis; contribution analysis.
Sustainability	S2. To what extent was the management of relationships with regional partners and stakeholders effective in building real partnership and ownership?	Quality of partnership management; ownership by partner-country stakeholders; sustainability of South–North and South–South linkages; inclusiveness across the actor categories the quadrilogue method seeks to involve (note lighter non-civil-society participation).	Partners' perceived ownership and trust; continuity of joint actions beyond project funding; balance of participation across those categories.	KIIs with partner-country stakeholders	Network on Youth & Global Citizenship reports; AUYD report; partnership records; ExCom synopses (ownership & engagement). Partners in Algeria, Cabo Verde, DRC, Lebanon, Liberia, Morocco, Tunisia; NSC partnership focal points.	Thematic synthesis; triangulation.
Added value	AV1. What is the added value of the NSC in its implementation and methodology (the quadrilogue approach)?	Comparative advantage of the NSC's quadrilogue method (its convening of governments, parliamentarians, local and regional authorities and civil society in its activities) relative to other actors; the practical reality of the method (civil-society-led; lighter governmental, parliamentary and local/regional participation). The NSC's own governance via the Executive Committee is assessed as a related but distinct dimension.	Stakeholder perceptions of the NSC's unique contribution; outcomes plausibly attributable to the quadrilogue; balance of participation across the four layers.	KIIs across the relevant actor categories and governance stakeholders; document review (ExCom synopses); comparative / benchmark analysis	ExCom synopses (29th–35th); benchmark docs (GENE, Bridge 47, CONCORD); Council of Europe–UNAOC MoU; mission reports (CDEDU, CDEJ/CMJ). NSC statutory & Executive Committee bodies; GEN; partners.	Benchmarking; contribution analysis; triangulation.

Annex 3: Documents reviewed

The evaluation reviewed more than one hundred programme and activity reports as well as governance and benchmark documents:

- Council of Europe Evaluation Guidelines (December 2021)
- Description of the Action and design logframe
- Budget-revision explanatory note and revised budget workbook (March 2025)
- Interim Financial Reports (2023, 2024, 2025)
- Annual progress reports (2023, 2024, 2025)
- NSC Executive Committee synopses (29th, 30th, 32nd, 33rd, 35th meetings)
- Mission reports to the Council of Europe steering committees for education and youth (2024 to 2026)
- Activities: concept notes, final reports, participant lists
- Immediate evaluation forms and three-month follow-up files for the Universities on Youth and Development (2023, 2024, 2025), the African UYD (2024), the LTTC on Global Education and Education for Sustainable Development, the tutored online courses on Global Education and Media (2023, 2024) and on Global Education and Intercultural/Interfaith Dialogue (2023, 2024), and the HEY courses on Media and Information Literacy (2025, 2026), Intercultural/Interfaith Dialogue (2025) and Environment (2025)
- HEY Intercultural/Interfaith Dialogue 2026 final report (draft; three-month results pending), participant list and application-form export; Tutored Online Course on Global Education and Intercultural/Interfaith Dialogue 2024 application-form export; final participant lists for HEY Intercultural/Interfaith Dialogue 2025 and the Tutored Online Course on Global Education and Media 2024
- Global Education Network annual meetings and thematic seminars (2023, 2024)
- Network on Youth and Global Citizenship and Africa-Europe stakeholders reports (2024, 2025)
- Global Education Week concept notes, reports, planning document (2023), coordinator survey files (2024, 2025) and analysis workbook (2025)
- Framework contracts and deliverables for Global Education Week services (2024 to 2026)
- Consolidated course evaluation-form exports 2023 to 2026, including the French-cohort MIL-EMI 2025 evaluation, the 2026 ICD-IFD course evaluation and the MIL-EMI 2026 three-month follow-up
- iLEGEND II final evaluation report (October 2022)
- Council of Europe and EU policy documents including CM/Rec(2011)4, CM(2016)25, SG/Inf(2020)34 and DEAR programme documentation
- UNDP Independent Evaluation Office, Youth and development, Reflections series (2025)

Annex 4: Data re-analysis

This annex documents the re-analysis of the programme's monitoring and survey data and substantiates the quantitative statements made under E1 (Table 3) and E3 (Table 5). It presents the data inventory, the indicators recomputed for the online courses and the residential activities, the verifiability assessment of the Global Education Week reach figures, and the limitations of the underlying data.

The corpus comprises the 46 monitoring and survey datasets received with the NSC document package (application records, participant lists, immediate end-of-activity evaluation forms, three-month follow-up questionnaires and campaign survey files). All figures in this annex were recomputed from the source files; where an export does not correspond to the final cohort or is missing, figures are taken from the course reports and marked with an asterisk (*). Only aggregates are reported; no personal data are reproduced.

Instruments and conventions

- Satisfaction is the mean of responses to the item “To what extent has this course met your expectations/needs?”, scored 1 (low) to 6 (high) in all immediate evaluation forms. The share of respondents scoring 5 or 6 is reported alongside each mean.
- Completion is the number of certified participants divided by the number selected. The programme's reports also use an active-participant denominator, which yields higher rates (77.5 percent in 2025); both conventions are discussed under E1.
- Three-month outcomes are the shares of respondents answering yes to the questions on having shared the knowledge acquired and on having implemented, or at least planned, a follow-up activity. The 2023 questionnaires carry a single question (application), so no sharing figure exists for the 2023 courses.

Data inventory

The table below summarises the data that exist per activity. It is the evidence base for the observation, discussed under E1 and reflected in recommendation R5, that monitoring is strong at course level, weaker at campaign level and absent at outcome level.

Activity	Participant list	Immediate evaluation (n)	Three-month follow-up (n)	Final report
HEY Environment 2025	Yes (46)	Yes (28)	Yes (12)	Yes
HEY MIL-EMI 2025 (EN + FR)	EN only (39 of 76)	Yes (39: 21 EN, 18 FR)	Yes (20, anonymised)	Yes
HEY MIL-EMI 2026 (EN + FR)	Yes (99)	Yes (45: 26 EN, 19 FR; plus per-module evaluations, n = 21–31)	Yes (17)	Yes (the version received contains some unfilled placeholders in its results section)
HEY ICD-IFD 2025	List of 110 does not match the 46 selected	Yes (26)	Yes (17)	Yes
HEY ICD-IFD 2026	Yes (74)	Yes (41)	Pending (expected late 2026)	Yes (received August 2026)
TOTC ICD-IFD 2023	Yes (30)	Yes (20)	Yes (3; too few to use)	Yes
TOTC ICD-IFD 2024	Yes (45)	Yes (22)	Yes (2; too few to use)	Yes
TOTC Media 2023	Yes (30)	Yes (21)	Yes (11)	Yes

Activity	Participant list	Immediate evaluation (n)	Three-month follow-up (n)	Final report
TOTC Media 2024	List of 67 does not match the 32 selected	Yes (20)	Yes (7)	Yes
LTTC GE and ESD (blended)	Yes (25 selected, 21 attended)	Yes (20)	Follow-up phase documented in the final report	Yes
UYD 2023	Yes (240 listed, 166 participants)	Results stored as chart images (n≈108)	No	Yes
UYD 2024	Yes (154 curated, 101 participants)	Yes (49, of which 13 rows corrupted)	Yes (38)	Not received
UYD 2025	Yes (228 listed, 159 participants)	Yes (78)	Yes (38)	Yes
AUYD 2024	Yes (90 listed, 51 participants)	Yes (21)	No	Yes
GEN annual meetings and seminars 2023-24	Lists in PDFs (36, 37)	Results stored as images (n = 24, 24, 21)	n/a	Yes
GEW 2023	No	No raw data	No	Yes (narrative)
GEW 2024	Coordinator contacts only	Coordinator survey (30) and qualitative round (13)	No	Slides only
GEW 2025	No	Coordinator survey (36) and analysis workbook	No	Yes

Online courses: recomputed indicators

The table below presents the indicators recomputed from the immediate evaluation forms, participant lists and course reports for the tutored online course editions delivered or evaluated by the cut-off. Satisfaction cells give the mean on the six-point scale, the number of respondents and the share scoring 5 or 6; follow-up cells give the number of respondents, the response rate against selected participants, and the shares reporting having shared and applied the learning.

Course	Applications	Selected	Certified (share of selected)	Women	Residence: Council of Europe / Africa / Middle East / other	Satisfaction mean (n; % scoring 5-6)	Three-month follow-up
HEY Environment 2025	563	46	26 (57%)	50%	11 / 32 / 3 / 0	5.11 (28; 71%)	n = 12 (26%): shared 92%; applied 67%
HEY MIL-EMI 2025 (EN + FR)	806	76	42 (55%)	54%*	17 / 56 / 3 / 0*	EN 5.38 (21; 100%) FR 5.33 (18; 83%)	n = 20 (26%): shared 70%; applied 70%
HEY MIL-EMI 2026 (EN + FR)	4,584	99	64 (65%)	53%	52 / 44 / 3 / 0*	5.29 (45; 82%)	n = 17 (17%): shared 82%; applied 76%
HEY ICD-IFD 2025	294	46	32 (70%)	50%	13 / 28 / 5 / 0	5.42 (26; 88%)	n = 17 (37%): shared 94%; applied 76%
HEY ICD-IFD 2026	590	74	46 (62%)	58%	27 / 41 / 4 / 2	5.27 (41; 85%)	Not fielded by the cut-off
TOTC ICD-IFD 2023	162	30	25 (83%)	67%	17 / 9 / 1 / 3	4.60 (20; 55%)	n = 3 (10%): too few to report

Course	Applications	Selected	Certified (share of selected)	Women	Residence: Council of Europe / Africa / Middle East / other	Satisfaction mean (n; % scoring 5-6)	Three-month follow-up
TOTC ICD-IFD 2024	685	45	n/a†	n/a†	n/a†	4.86 (22; 64%)	n = 2 (4%): too few to report
TOTC Media 2023	61	30	23 (77%)	80%	21 / 8 / 1 / 0	5.38 (21; 95%)	n = 11 (37%): applied 82% (no sharing question)
TOTC Media 2024	78	32	23 (72%)	72%	27 / 5 / 0 / 0	5.55 (20; 85%)	n = 7 (22%): shared 71%; applied 29%

** Taken from course report tables (raw export missing or not the final cohort). † Not available.*

Three results stand out. First, demand vastly exceeds supply: applications ranged from 61 to 4,584 for 30 to 99 places, and the 2026 Media and Information Literacy course accepted approximately two percent of applicants. Second, delivery quality is consistent: completion among selected participants ranged from 55 to 83 percent, and satisfaction means ranged from 4.60 to 5.55 across the ten course-cohort evaluation datasets (nine editions, with the 2025 media course evaluated separately in English and French), with 55 to 100 percent of respondents scoring 5 or 6. Third, Africa is the largest region of residence in four of the five HEY-era editions with cohort data (55 to 74 percent of the cohort), confirming that the online modality reaches the programme's Southern target groups.

The 2024 Intercultural/Interfaith Dialogue edition scores 4.86 (n = 22). The two tutored editions on this theme (4.60 in 2023 and 4.86 in 2024) are the two lowest satisfaction scores in the record, while the HEY editions of the same theme score 5.42 (2025) and 5.27 (2026), a progression consistent with the iterative improvement of course design discussed under E2.

The three-month figures must be read against their denominators. Response rates ranged from 4 to 37 percent of selected participants (17 to 37 percent across the six usable returns), respondents are self-selected, and the two Intercultural/Interfaith Dialogue tutored-course returns (n = 3 and n = 2) are too small to report; the 2024 return moreover consists of two submissions made under the same name and e-mail address. In the 2025 Intercultural/Interfaith Dialogue follow-up, two submissions received about forty seconds apart are identical on every substantive field; treating them as a single respondent would not materially change the results (n = 16: shared 94 percent, applied 75 percent). The follow-up percentages are therefore indicative of potential rather than representative of each cohort.

Supplementary indicators from the immediate evaluation forms

Beyond the headline satisfaction item, the immediate evaluation forms score the achievement of the course objectives, the contribution to personal development and learning, and the trainer and tutor team (five items). The means recomputed across all editions corroborate the satisfaction record; the trainer and tutor teams are rated at or above the course-objective items in ten of the eleven datasets (range 5.25 to 5.74).

Course (immediate evaluation, n)	Objectives achievement (mean, 1-6)	Personal development (mean, 1-6)	Trainers and tutors (mean, 1-6)
TOTC ICD-IFD 2023 (20)	4.90	4.45	5.38
TOTC Media 2023 (21)	5.33	5.14	5.49
TOTC ICD-IFD 2024 (22)	4.97	4.73	5.25
TOTC Media 2024 (20)	5.40	5.35	5.43
LTTC GE and ESD (20)	5.11	5.15	5.55

Course (immediate evaluation, n)	Objectives achievement (mean, 1-6)	Personal development (mean, 1-6)	Trainers and tutors (mean, 1-6)
HEY ICD-IFD 2025 (26)	5.35	5.35	5.33
HEY MIL-EMI 2025 EN (21)	5.46	5.43	5.74
HEY MIL-EMI 2025 FR (18)	5.20	5.17	5.41
HEY Environment 2025 (28)	5.32	5.14	5.32
HEY MIL-EMI 2026 (45)	5.27	5.13	5.30
HEY ICD-IFD 2026 (41)	5.18	4.98	5.37

Residential activities

The residential datasets were received, except for the LTTC final evaluation. The indicators below are recomputed from the participant lists and, where the raw forms are usable, from the evaluation files. Satisfaction at the universities follows each course's own instrument (share of satisfied respondents plus an engagement score), which is why it is expressed differently from the course means above.

Activity	Participants	Women	Residence: Europe-Council of Europe / N. Africa-Middle East / Sub-Saharan Africa / other	Satisfaction	Three-month outcomes
UYD 2023	240 listed, 166 participants	≈56%	116 / 30 / 14 / 6	n ≈ 108; results stored as chart images	None
UYD 2024	154 curated, 101 participants	57%	70 / 22 / 7 / 2	94% satisfied (36 valid of 49); engagement 4.96/6	n = 38: shared 97%; follow-up initiative 61%
UYD 2025	228 listed, 159 participants (210 in report)	≈53%	95 / 29 / 24 / 11	92% satisfied (78); engagement 4.5/6	n = 38: shared 92%; applied 92%
AUYD 2024	90 listed, 51 participants (79 in report)	51%	2 / 18 / 31 / —	90% satisfied (21); engagement 5.14/6	None
LTTC GE and ESD	25 selected, 21 attended	67% of attendees	12 / 7 / 2 / —	5.40/6 (20); 90% scoring 5 or 6	Phase 3 documented in the final report: 20 of 21 active; 9 initiatives; 4 taken into GEW
GEN annual meetings and seminars 2023, 2024	36 and 37 listed	Not stated	All Europe-Council of Europe	24, 24 and 21 forms; results stored as images	n/a

Residential events are Europe-majority (around 60 to 70 percent Europe/Council of Europe at the universities), with Sub-Saharan Africa at around 7 to 15 percent, except at the African edition, where Sub-Saharan participants are the largest group, confirming its role in bringing the programme to the region. Satisfaction is uniformly high (90 to 94 percent satisfied). The two usable university follow-ups (n = 38 in both years) show strong self-reported multiplication: 92 to 97 percent shared their learning and 61 percent of the 2024 respondents had initiated a follow-up activity. The LTTC lost two of its four Sub-Saharan participants to visa refusal and drop-out before the residential phase, a concrete instance of the access constraint discussed under E2.

Global Education Week: verifiability of the reach figures

The published reach figures rest entirely on annual self-report surveys of the national coordinators (20 to 36 respondents, mostly banded estimates); the NSC's own 2025 campaign evaluation states that no data are collected at activity level beyond estimated totals. The assessment below substantiates Table 5 under E3.

Year	Published claim	What the underlying data support	Verdict
2023	312,667 participants; 7,000 registered activities; 26 countries	No raw dataset exists for 2023; the planning document contains pre-event entries only. The campaign report itself presents banded results for 20 countries which cannot arithmetically yield the headline figure unless a single “more than 10,000” country contributed around 280,000. The activities figure appears in no source document.	Not verifiable; internally contradictory
2024	Over 300,000 participants; about 6,000 activities; 28 countries	A coordinator survey exists (n = 30), mostly banded. Band minima support roughly 240,000, of which one country's self-reported “more than 200,000” accounts for over 80 percent. The activities figure is not derivable from the data (band floor around 3,700).	Order of magnitude only
2025	430,717 participants; 5,698 activities; 45 countries	The best-documented year: per-country point estimates (n = 36). The workbook's own country sheet totals 397,380 participants and 5,182 activities; the published headline is hard-coded without reconciliation. The largest single entry, initially 700,000, was reduced to 160,000 in the compilation; documentation obtained during the evaluation shows the coordinator's actual collection (February 2026) at 88,976 direct participants plus approximately 112,194 reached through the national network, matching neither figure. Two countries account for 72 percent of the total, and about 12 of the 45 countries have no country-level data (they are counted as reached through another coordinator's national network).	Order of magnitude only; totals likely overstated

The fair reading, reflected under E3, is that the campaign mobilises hundreds of thousands of people across dozens of countries; the five- and six-digit precision of the published totals is not supported by the underlying data, and the activity counts are the weakest element. The 2025 revision of the largest entry is documented in correspondence obtained during the evaluation, which shows the compilation process to be traceable but the version control weak; this underpins recommendation R5.

Limitations of the underlying data

- Self-selection and small denominators: three-month surveys reached 4 to 37 percent of selected participants (17 to 37 percent for the usable returns), all self-selected completers; the resulting percentages indicate potential rather than cohort-wide change.
- Anonymity: the forms are anonymous or were anonymised, so immediate and follow-up responses cannot be linked and results cannot be disaggregated by gender, age or region (the same constraint that defeated the logframe's disaggregation clauses, Annex 1).
- Data completeness and quality: some records were missing, incomplete, image-based or inconsistent with final cohorts; a small number of duplicate submissions were also identified, and some raw data remained unavailable at the evaluation cut-off. Management indicated that several of these items could be provided or clarified but they do not alter the analysis presented in this report.

Taken together, the data picture is itself a finding: monitoring is strong at course level (clean cohorts, consistent instruments, six-point scales, three-month follow-ups fielded), weak at campaign level and absent at outcome level, matching the weakest causal links in the theory of change as discussed under E1 and in the contribution analysis.