

EDULAB

European Youth in Transition to
Education and Labour



www.edu-lab-project.eu

Design, Implementation and Outcomes of the First Two Rounds of OSES-Delphi and Design of Research Tools for the Concluding Rounds

Volume One (the Executive Summary) – this document

Volumes Two, Three and Four will be published on EDU-LAB web separately

Deliverable number and title: D3.5 “Outcomes of the first two rounds of OSES, design of research tools for the concluding rounds”

Preceding Deliverable: D3.4 “Terms of Reference for the Design and Implementation of the Online Survey of Experts and Stakeholders (OSES)”
(produced on 30 June 2025)

Due date: 30 June 2026

Nature: R = Report

Dissemination Level: PU = Public

Work Package: WP3B (OSES - International expert survey)

Lead Beneficiary: Economica

Contributing Beneficiaries: All project partners

Authors: Alexander Chvorostov, Daniel Helmenstein, Claudia Palt, et al.



Funded by
the European Union

A Horizon Europe collaborative project (2025-2027)
Grant Agreement #101177428

Disclaimer:

The content of this deliverable reflects only the authors' view. The European Commission and its Research Executive Agency are not responsible for any use that may be made of the information it contains.

Purpose and scope of the deliverable

This deliverable (D3.5) documents the design, conduct and first results of the Online Survey of Experts and Stakeholders (OSES), the policy-Delphi strand of Work Package 3B of the EDU-LAB project.

This report is a follow-up of the first deliverable in WP3B (D3.4), produced by 30 June 2025, which was initially marked as “CO” – confidential. Now, as the main phase of the OSES-Delphi fieldwork have been accomplished, and D3.5 has gone public, the dissemination level of D3.4 was also set to “PU” – public.

Document history

Version	Date	Description
0.1	06 June 2026	Initial draft (Daniel Helmenstein) (circulated within the Economica team members)
0.2	08 June 2026	Revised draft (Alexander), returned to Economica team
0.3	10 June 2026	Final Wave 2 results integrated after field close (N = 207; cut-off 18:00); convergence typology, per-country, profile and matched-sample analyses added (Daniel Helmenstein)
0.4	10 June 2026	PMB review feedback integrated (UNI BA, EVALAG): Policy Priority Matrix reframed; EEI-to-statement methodology detailed; panel-design note (re-contact retention, longitudinal matching) added; dimension bar charts moved to the annex; source-note convention (Daniel Helmenstein)
0.5	22 June 2026	Reworked and finalized by Economica team (the final text has been curated with the help of AI tools)
1.0 (final)	22 June 2026	Approved by PMB and the Coordinator

List of project-specific terms and abbreviations

CEDEFOP	The European Centre for the Development of Professional Training (an EU Agency)
DPA	Data-Processing Agreement
ECEC	Early Childhood Education and Care
EDU-LAB	The EU-funded project “European Youth in Transition to Education and Labour”
EEA	European Education Area
EEI	Exploratory Expert Interview
ETF	European Training Foundation
GE	General education
GDPR	General Data Protection Regulation
IQR	Inter-Quartile Range (Q3 – Q1)
ISCED	International Standard Classification of Education
Md	Median
OSES	International Online Survey of Experts and Stakeholders (the WP3B policy Delphi)
PMB	Project Management Board
PVET	Technical and Vocational Education and Training
SD	Standard deviation
TVET	Professional and vocational education and training
WP	Work package
W1 / W2	Wave 1 / Wave 2 of the Delphi; S01–S30 = statement codes

Codes for EDU-LAB target countries

FI Finland	IT Italy	UK the United Kingdom
AT Austria	GR Greece	PL Poland
XK Kosovo	PT Portugal	DE Germany

Table of Contents

1	Introduction and summary	4
1.1	Introducing OSES-Delphi as one of backbone empirical components of EDU-LAB project	4
	Preamble	4
	Background and objectives of EDU-LAB project and OSES-Delphi study	4
1.2	Structure of the reporting package	5
	Reader's guide to Deliverable D3.5	5
	A note on the character of this report (First Volume)	5
	Conventions used throughout D3.5	6
	Major groups of stakeholders targeted by OSES-Delphi research program	6
1.3	Summary of the main points of the methodology of OSES-Delphi	6
2	Methodology of OSES-Delphi	8
2.1	A sequential mixed-methods Delphi design	8
2.2	The Delphi method and its rationale	8
2.3	Exploratory Expert Interviews (EElis)	9
2.4	From interviews to statements: the road to the Delphi	9
2.5	The questionnaire and the assessment dimensions	11
2.6	Stakeholder groups and recruitment	13
2.7	Fieldwork, data processing and data quality	13
2.8	Analysis approach	14
3	Results of Wave 1	16
3.1	The expert panel	17
3.2	Overall patterns across the five dimensions	17
3.3	Results by trend statements	20
	Theme A — Education and Training	21
	Theme B — Labour Market and Employment	24
	Theme C — Transitions and Systemic Interfaces	27
	Theme D — Inclusion and Disadvantage	30
3.4	Consensus and controversy	33
3.5	Quantitative profile and qualitative engagement, side by side	35
	Illustrative qualitative themes (most-commented statements)	36
4	Wave 2 — reassessment rationale, design and final results	37
4.1	Why a second wave	37
4.2	Panel design: re-contact retention and longitudinal matching	37
4.3	The 21 carried-forward statements and the four blocks	38
4.4	Block-specific design and additional items	39
4.5	Final results	42
	Results by block	43
4.6	What the second round changed — a convergence typology	47
4.7	Per-country convergence	47
4.8	Driver and policy profiles by theme	48
4.9	Feasibility, agreement and certainty: what Blocks 3 and 4 add	48
4.10	How individual experts moved: the matched sample	48
5	Connections to other research streams in EDU-LAB	50
6	Outlook	52
	References	53
	ANNEX	54

1 Introduction and summary

1.1 Introducing OSES-Delphi as one of backbone empirical components of EDU-LAB project

Preamble

The Online Survey of Experts and Stakeholders (OSES) is the policy-Delphi strand of Work Package 3B (WP3B) of the EDU-LAB project, led by the Economica Institute of Economic Research in Vienna, Austria. As a transnational governance instrument, it gathers structured, cross-national expert judgements on the future of young people's transitions from education and training into the labour market and translates these judgements into input for EU- and country-level policy formulation (e.g. in a form of Policy Briefs, integration reports, toolkits, dashboard of empirical data and evidences, etc.) The instrument combines quantitative consensus measurement with qualitative interpretation in a sequential, multi-wave design.

Background and objectives of EDU-LAB project and OSES-Delphi study

EDU-LAB is a Horizon Europe collaborative research project (Grant Agreement No. 101177428) running from 2025 to 2027. It addresses systemic challenges in the transition of young people from education into employment, examining how policies, practices and institutional arrangements across different education and training systems (particularly in general education and in professional/vocational education and training) shape outcomes for young people. Among its core concerns are early school-leaving, work-based learning, regional disparities, the policy and investment landscape in the European Education Area, and the identification and promotion of effective and inclusive practices at local, national and EU levels. The project brings together a multidisciplinary consortium of universities, research institutions and policy-oriented organisations from across Europe and aims not only to generate knowledge but to facilitate its uptake through collaborative stakeholder engagement.

Within this frame, WP3B “Online Survey of Experts and Stakeholders” is one of the project's two core empirical components. Its purpose is to capture expert and stakeholder foresight on youth education-to-work transitions amid digital and demographic change; to measure where expert consensus and controversy lie; and to feed those findings into the project's synthesis (WP5) and into policy briefs, recommendations and publications (WP6), while also informing the work of WP1 and WP2. The instrument is deliberately built for impact: cross-national expert input is structured so that both stable agreement and persistent dissent can be read off and carried into policy formulation.

Four thematic areas structure the entire OSES effort. These frame the interview guide, the statement clusters and the final questionnaire alike, closing the methodological circle across stages:

- Education and Training: The content, quality and direction of general and vocational education for young people.
- Labour Market and Employment: Entry pathways, employability, skills demand and the position of young people in the labour market.
- Transitions and Systemic Interfaces: The passage from education into work and the institutional interfaces that govern it.
- Inclusion and Disadvantage: Educational inequality, social background, and the inclusion of disadvantaged and minority young people.

The OSES also strengthens the project methodologically. Where many other foresight exercises (for example single-country expert panels or one-off scenario workshops) rely on a single national panel or a small workshop, it assembles a large multi-country panel and measures, for every trend, not only the central expectation but the degree of agreement around it. That makes it possible to distinguish developments on which European experts broadly concur from those on which they remain divided, which is a distinction that matters for how confidently a finding can be carried into policy recommendation, and one that recurs throughout the results below.

1.2 Structure of the reporting package

Reader's guide to Deliverable D3.5

Deliverable D3.5 documents the OSES-Delphi in four complementary components. The Executive Summary (this text report) is the narrative anchor; the other three provide the underlying detail. The table below shows how they fit together and how this report relates to each other. The other three volumes can be downloaded from the project's website www.edu-lab-project.eu.

- The first volume** (Executive Summary) is the narrative anchor of Deliverable D3.5. It documents the methodology and the “road to the Delphi” — the path from exploratory expert interviews, through a structured statement-reduction process, to the final set of trend statements — and then summarises the descriptive results of the first wave and sets out the rationale, design and final results of the second wave, including the Wave 1 → Wave 2 convergence and matched-sample analyses. It is written to be read on its own while pointing to the companion components (three further volumes) for full detail.
- The second volume** (Quantitative Reference Tables) presents the complete statistical tables of survey responses received.
- The third volume** (Qualitative Reference Tables) contains a full collection of responses to open-ended questions.
- The fourth volume** (OSes-Delphi Tools and Documentation) includes a full collection of survey instruments (concise questionnaires in all project languages), full list of Delphi statements, GDPR materials, hidden online survey items, as well as the detailed ex-ante methodology in the Terms of Reference (D3.4).

Table 1. Matrix of the main components of D3.5 composite report

D3.5 Components (volumes)	Contents	Relation to this report
1. Executive Summary (this document)	Methodology, road to the Delphi, descriptive Wave 1 results, Wave 2 rationale and final results, links to WP1/WP2	The overview that ties the others together
2. Quantitative Reference Tables	Full Wave 1 and final Wave 2 tables (overall and per country), the Wave 1 ↔ Wave 2 comparison and the matched-sample (within-respondent) analyses; sample description and factor/policy frequencies	Source for every figure cited here; consult it for complete distributions
3. Qualitative Reference Tables	All open-ended responses, per country and per statement, with systematic coding	Source for the qualitative engagement and themes in Section 4.5
4. OSes-Delphi Tools and Documentation	Questionnaires and GDPR materials (in all languages), hidden items, the statement list	Reference for instruments and wording

A note on the character of this report (First Volume).

At the present stage the OSES is reported descriptively. The aim here is to set out transparently how the data were produced and what they show (distributions, levels of agreement, and notable patterns) rather than to advance causal explanation or policy recommendations. In-depth analytical interpretation and the integration of the quantitative and qualitative strands are reserved for the project's later synthesis (WP5) and dissemination (WP6) stages, once the concluding qualitative round is complete.

The report is structured as follows. Section 2 sets out the project background and objectives. Section 3 documents the methodology such as the design, the exploratory interviews, the road from interviews to the 30 Delphi statements, the questionnaire, the sample and recruitment, the fieldwork and data quality, and the analysis approach. Section 4 presents the descriptive Wave 1 results: the panel, the overall patterns, a statement-by-statement walk-through, and a consensus and controversy synthesis. Section 5 sets out the

Wave 2 reassessment with its rationale, its four-block design and its final results, including the Wave 1 → Wave 2 convergence map, the per-country comparison, the driver and policy profiles and the matched-sample analyses. Section 6 looks ahead to the concluding qualitative round.

Conventions used throughout D3.5:

- Terms “*Statement(s)*” and “*Trend Statement(s)*” are used interchangeably and stand for key thematic tokens suggested to Delphi respondents for assessment along various scales.
- Statements carry the codes S01–S30, identical across all components.
They are grouped into four thematic areas:
 - A: Education and Training,
 - B: Labour Market and Employment,
 - C: Transitions and Systemic Interfaces,
 - D: Inclusion and Disadvantage.
- Statement assessment ratings use the following measurement units:
 - a 1–5-point Likert-scale plus a non-substantive “cannot judge” option;
 - central tendency is reported as the median (Md)
 - dispersion as the interquartile range (IQR), with $IQR \leq 1$ read as consensus and $IQR \geq 2$ as divergence.
- Source (all tables and figures, unless stated otherwise): EDU-LAB Delphi Survey — Wave 1 (N = 433) and Wave 2 (N = 207), Horizon Europe Project EDU-LAB (2025–2027), Grant Agreement No. 101177428. Calculations and data processing by Economica Institute of Economic Research, Vienna.
- Statement texts are quoted verbatim from the fielded questionnaire (including their original US-English spellings); the surrounding text follows British English.

Major groups of stakeholders targeted by OSES-Delphi research program

Eligible respondents were recruited from six principal clusters of stakeholders, with multiple attribution across clusters explicitly permitted (a respondent may, for example, be both a researcher and a former teacher):

- A. Governance and administration: authorities funding and running general and vocational education and training; labour-market services; certification and accreditation bodies.
- B. Business actors: sector and production chambers; employers and their associations; employee, student and teacher unions; investors and sponsors.
- C. Teaching and training staff: teachers and trainers in general and vocational education and training, kept distinct from administrative and technical staff.
- D. Youth and civil society: parent and youth organisations; NGOs and grass-roots groups; media; other stakeholders involved in young people's lives.
- E. Research and academia: individual researchers and networks; private head-hunting / HR firms and consultancies.
- F. International organisations: EU bodies (Commission, CEDEFOP, ETF) and global institutions (UNESCO, OECD, World Bank, EBRD).

1.3 Summary of the main points of the methodology of OSES-Delphi

- **Design.** The OSES is a sequential mixed-methods policy Delphi: 47 exploratory expert interviews (EIs) mapped the issue space; a first quantitative wave measured where experts stand; a second wave feeds those results back for structured reassessment; and a concluding qualitative round will deepen the interpretation.
- **Languages.** Questionnaires for Exploratory Interviews and the Delphi survey tools have been initially designed in English and then localized for ten project languages. Interviews have been conducted in the language preferred by the respondents, while Delphi online survey was offered in English and the further languages of participating countries.

- **Instrument.** EEI findings and partner input produced a pool of 106 candidate statements, which a transparent, partner-rated reduction process condensed into 30 trend statements across four thematic areas: Education and Training (8 Statements), Labour Market and Employment (7 Statements), Transitions and Systemic Interfaces (8 Statements), and Inclusion and Disadvantage (7 Statements). Each statement was rated on five dimensions: desirability, importance for policymakers, conflict potential for society (5-point Likert scale), and likelihood of realisation within 5 and within 15 years.
- **Wave 1 sample.** Fielded 4 March – 10 April 2026, the cleaned dataset comprises N = 433 experts, predominantly highly qualified people, spanning all six stakeholder clusters at international, national and local levels. Respondents have been invited to take part in OSES-Delphi survey in the course intensive cross-country recruitment campaigns organized by eleven project partners in all nine EDU-LAB countries. The final sample includes 407 completed expert responses from project countries (though, these may account also for experts from international organizations and third countries); respectively, 26 respondents indicated their residence in countries other than the formal EDU-LAB geographic area. The most responses are available for Finland (71), Austria (58), Kosovo (55) and Italy (47).
- **Headline findings of Wave 1.** Experts agree most strongly on the likelihood of the trends and are most divided on their importance and desirability. The clearest descriptive signal is a normative gap in the inclusion domain: trends such as widening educational inequality, the persistence of social-class effects and intensified discrimination are judged highly undesirable yet among the most likely to materialise and carry the highest perceived conflict potential.
- **Wave 2.** 21 statements were carried forward, grouped into four blocks by their Wave 1 result pattern (controversial; diverging; uncertain; consensus). Fielded 4 May – 10 June 2026, Wave 2 closed with N = 207 completed reassessments (89.6% of started sessions; 205 linked to their Wave 1 records). The second round confirms convergence: on the 24 dimensions measured in both waves the spread of opinion narrowed in 9 cases and widened in none, medians held except for one upward shift on the five-year likelihood (S21), policy action on the consensus trends is rated feasible (median 4 throughout), and the matched sample shows a clear net pull of individual revisions toward the panel view.

2 Methodology of OSES-Delphi

This section describes the design and execution of the OSES-Delphi study in such detail that the results are interpretable. It remains a summary: the full methodological specification (instruments, ethics approvals, consent forms and the complete recruitment strategy) is laid out in the Terms of Reference (D3.4), to which the reader is referred for detail.

2.1 A sequential mixed-methods Delphi design

The OSES follows a sequential mixed-methods logic (Creswell & Plano Clark, 2018) that moves from qualitative exploration to quantitative consensus measurement and back to qualitative deepening. It comprises four stages, as outlined below.

First, the preceding thematic literature review and exploratory expert interviews (EElS) map the issue space and surface candidate trends.

Second, a first mostly quantitative Delphi wave establishes where experts stand on a closed set of trend statements.

Third, a subsequent wave returns the Wave 1 group results to respondents and invites a structured reassessment in the light of the collective view.

Fourth, a concluding round of qualitative interviews deepens the interpretation of the quantitative findings, focusing on the reasons behind persistent dissent and unexpected consensus.

This report covers stages one to three; the concluding qualitative round (an anticipated 30–40 online or face-to-face interviews, audio-recorded with informed consent under the same pseudonymised pipeline as the EElS) belongs to the later project timeline.

The rationale for the sequence is that each stage informs the next: the initial exploratory interviews ground the statements in real stakeholder concerns; the first Delphi wave quantifies agreement and disagreement; the second wave tests whether informed reflection moves the panel; and the final interviews explain the patterns that the numbers reveal. The design thereby combines the breadth of a large, standardised survey with the depth of qualitative inquiry.

The concluding qualitative round is designed to mirror the exploratory interviews, so that the study closes on the same methodological footing on which it opened. Thematically it concentrates on the causes of persistent dissent or unexpected consensus, the implicit assumptions behind divergent assessments, the practical implications for national and EU policy, and experts' own assessments of perceived policy effectiveness.

2.2 The Delphi method and its rationale

The Delphi method is a structured group-communication technique that elicits judgement from a panel of experts through successive rounds with controlled feedback between them. Rather than convening experts in a single meeting, where dominant voices, status and group dynamics can distort the outcome, the Delphi gathers views individually and anonymously, feeds back the aggregate, and lets participants reconsider. It is, as a structured group-judgement technique, particularly suited to forward-looking, contested questions for which no single data source provides an answer (Rowe & Wright, 1999; Okoli & Pawlowski, 2004).

The OSES applies the method specifically as a policy Delphi (Turoff, 1970). The goal of a policy Delphi is not to manufacture agreement but to map the structure of expert opinion honestly: to identify where there is stable consensus, where there is durable controversy, and where judgement is uncertain as each of these is informative for policy. Accordingly, the analysis reports central tendency with the median and dispersion with the interquartile range (IQR), the measures appropriate to ordinal rating data (Hsu & Sandford, 2007; von der Gracht, 2012) and treats a low IQR as a signal of consensus and a high IQR as a signal of genuine division rather than mere noise.

The method originated in technological forecasting in the mid-twentieth century (Dalkey & Helmer, 1963; Linstone & Turoff, 1975) and has since been widely adapted across policy, health, education and foresight

research (Landeta, 2006). Its enduring appeal rests on three features that suit the OSES well: anonymity, which frees participants from social pressure and allows minority views to be expressed; iteration with controlled feedback, which gives participants the chance to reconsider in the light of the group; and the systematic, quantified treatment of judgement, which makes the degree of agreement an explicit, reportable result rather than an impression. In a transnational setting these features also help to balance contributions across very different national contexts and stakeholder positions.

2.3 Exploratory Expert Interviews (EIs)

The basis of the OSES-Delphi questionnaire are 47 exploratory expert interviews. These expert interviews were conducted as semi-structured interviews. All interviews were conducted by EDU-LAB researchers with pre-identified respondents representing the principal stakeholder groups, primarily in the project countries and, where appropriate, EU-wide and from around the world. These interviews were taken by qualified interviewers from all partnering teams in a language preferred by respondents (mainly the local language).

The interviews were conducted online (video) or face-to-face, lasted roughly 30–60 minutes, and were audio-recorded with informed consent. Transcription, translation and summarisation ran through a GDPR-compliant, pseudonymised local pipeline, and the resulting summaries were made available on the freely accessible [EDU-LAB dashboard](#).

The interview guide was organised around the same four thematic areas that structure the whole OSES: the education and training of young people; their labour-market situation and employment; the transitions within and between education, training and the labour market; and the challenges, difficulties and disadvantages young people experience. Within each theme, open questions invited interviewees to identify the most important current and future developments, tensions and uncertainties. Within each of the four themes, the interview guide combined a small set of fixed start-up questions with open, probing questions. Under education and training, interviewees were asked about the most consequential changes in the content, quality and organisation of general and vocational education; under labour and employment, about young people's entry into work, employability and the skills employers seek; under transitions, about the interfaces between education, training and the labour market and where they function poorly; and under challenges and disadvantages, about the obstacles facing disadvantaged and minority young people. Concluding questions invited interviewees to name the developments they considered most important, most desirable and most uncertain. The principal function of the EIs for the Delphi strand was thus generative: they produced the raw material of trends, claims and points of contention from which the Delphi statements were later built.

2.4 From interviews to statements: the road to the Delphi

Turning the interview material into a workable Delphi questionnaire required a structured reduction process, designed to be transparent and defensible at each step. The exploratory findings and partner contributions were first developed into a large pool of candidate trend statements, organised into thematic buckets and then consolidated into 106 distinct statements.

From transcripts to a candidate pool: extraction and qualitative content analysis

The starting point was a consolidated interview corpus. All 47 exploratory expert interviews were fully transcribed, translated into English, anonymised and merged into a single master document, which served as the primary analytical dataset. The analytical phase ran from November 2025 to January 2026 and followed a structured, traceable workflow: all anonymised summaries were read in full; relevant passages were marked (colour-coded) in their original context at the level of individual sentences and idea units; marked passages were extracted per interview, merged into a single corpus, split by the four guiding interview themes through a scripted procedure, and transferred into a structured spreadsheet. This yielded 418 extracted idea units (candidate raw statements, each traceable back to its interview of origin).

The transformation of this material into Delphi statements followed structuring qualitative content analysis (QCA), which combines systematic reduction with conceptual abstraction while preserving the contextual depth of expert reasoning. Coding proceeded in an iterative inductive–deductive loop: categories were first

developed bottom-up from the material and then stabilised top-down against the project’s conceptual framework. Two criteria guided selection throughout: Relevance to the research objectives and the novelty or uniqueness of an insight, so that dominant themes and less frequent but analytically important perspectives were both retained, while ambiguous passages were flagged and revisited iteratively.

Extracted units were grouped into thematic categories using a combined keyword-based and interpretative approach. For example, labour market and skills; technology and digitalisation; education system and institutions; family and social influence; governance and policy; equity and inclusion; mental health and motivation; COVID-19 impact. Categories were merged, split and balanced iteratively, targeting a manageable system of eight to twelve domains. Through linguistic cleaning, de-duplication and consolidation of overlapping ideas, the 418 units were condensed to roughly 327 thematically structured statements, then to just over 200, and finally, through abstraction into analytically precise, cross-nationally applicable formulations, to the candidate pool of 106 statements that entered the partner rating documented below (Table 2). The degree of abstraction is best seen in an example: the raw unit “mental health issues are on the rise” became the final statement “Mental health challenges among youth ... will continue to affect life skills” (S16).

Because Delphi statements are deliberately condensed, consistency between source material and final statements was checked at the conceptual rather than the literal level, combining exact and fuzzy text matching with heuristic semantic matching as the primary method. In a Delphi context, QCA is extended by one further step that distinguishes it from conventional content analysis: the translation of categories into forward-looking, ratable survey statements as the bridge from qualitative narratives to measurable constructs on which both survey waves rest.

In a first, quantitative reduction step, all 11 EDU-LAB partner teams rated the 106 statements on a five-point importance scale with each team’s position being summarized by its median rating. Statements with a median of at least 4, that is, rated “important” or “very important” by the typical team, were retained, yielding 54 statements, initially grouped into six thematic clusters for discussion. The median was chosen deliberately: with a small number of rating teams, it represents the typical expert view and is robust against outliers, and it gives a clear, easily communicated selection criterion. A consensus indicator based on the standard deviation of the ratings (high / moderate / low agreement) was reported alongside, but for information only — low consensus flagged a statement as worth discussing further, not as a candidate for removal.

Following a joint partner meeting on 9 January 2026, the working set was further developed and then condensed by Economica. Overlapping and duplicate statements were merged or removed, with each removal documented by a short reason: Most often that a statement overlapped thematically with another, was too narrowly specific, or had attracted low consensus. In the same step the surviving statements were re-categorised from the six discussion clusters into the four EEI thematic areas, closing the methodological circle by using the same clusters throughout the study. This produced a near-final set of about 35–36 statements, which partners reviewed once more for reformulation, deletion or reintroduction.

The removals were documented individually. The examples below illustrate the kinds of reasons applied during the consolidation step.

Table 2. Examples of statements removed during consolidation, with the reason recorded

Removed statement (abridged)	Reason recorded
Migration background will become a more decisive factor	Overlaps with the broader inequality theme
Primary-school quality improvements	Overlaps with the inequality theme
AI-generated content will make cheating harder to detect	Low consensus; topic too specific
Digital addiction and isolation among young people	Low consensus
Soft-skills training (CVs, interviews)	Narrow focus on a specific employability tool
Prejudices against non-binary young people	Topic too specific for a general trend statement

A final, practical constraint then applied. A pilot test of the programmed questionnaire showed that it was still too long for respondents to complete comfortably. Economica therefore trimmed the set from 36 to 30 statements, balancing thematic coverage against respondent burden. The (final) 30 statements are distributed across the four themes as Education and Training (8), Labour Market and Employment (7), Transitions and Systemic Interfaces (8) and Inclusion and Disadvantage (7). The questionnaire was then finalised and sent to partners for translation into the national languages, with the English and German versions prepared in-house. The path from pool to final instrument is summarised below.

Table 3. The road to the Delphi — stepwise reduction of the statement pool from 106 to 30

Step	Statements	Basis / criterion
Candidate pool	106	EEl findings and partner contributions, consolidated into distinct statements
First reduction	54	Median importance rating ≥ 4 across 11 partner teams; six discussion clusters
Consolidation	~35–36	Joint partner meeting (9 Jan 2026); duplicates/overlaps removed with documented reasons; re-mapped to the four EEl themes
Pilot trimming	36	Pilot test showed the questionnaire too long; trimmed for respondent burden
Final instrument	30	Distributed 8 / 7 / 8 / 7 across Themes A–D; translated into all national languages

2.5 The questionnaire and the assessment dimensions

In Wave 1, respondents first provided basic information (Sociodemographics like country of residence, age, gender), their highest level and field of education (ISCED-2011 and ISCED-F; UNESCO Institute for Statistics, 2012, 2014), and their stakeholder affiliation(s), and could indicate further professional or social affiliations. For each thematic area they could also self-assess their level of expertise and the geographical scope of that expertise.

Each of the 30 statements was then assessed on five dimensions using a 5-point Likert scale. Desirability was rated from “very undesirable” (1) to “very desirable” (5). Importance for policymakers, conflict potential for society, and the chances of realisation within the next 5 years and within the next 15 years were each rated from “very low” (1) to “very high” (5). Every dimension additionally offered a non-substantive “prefer not to say / don't know / cannot judge” option, which is reported separately and excluded from the central-tendency measures. Each statement also carried an open-comment field. Together, the five dimensions allow a statement to be read on several axes at once: Whether a development is wanted, whether it matters, whether it would be socially divisive, and how soon it is expected, which is what makes the desirability-versus-likelihood and importance-versus-conflict contrasts in Section 4 possible.

Figures 1 and 2 reproduce the English-language questionnaire for a single statement (here S01), showing the five rating dimensions, the “cannot judge” option and the open comment field as respondents saw them.

Figure 1. Wave 1 questionnaire (English version): a trend statement (here S01) as presented to respondents, rated for importance, conflict potential and likelihood over five and fifteen years on a five-point scale, with a “prefer not to say / cannot judge” option.



European Youth in Transition to
Education and Labour
A Horizon Europe collaborative project (2025–2027)



Grant Agreement
101177428
Funded by
the European Union



Project
Coordinator

Theme A: Education and Training

Please assess the following statement: **"Double competencies (disciplinary plus IT) will become a universal employment requirement"** with regard to:

	Very low	Low	Moderate	High	Very high	Prefer not to say / don't know / cannot judge
The importance for policymakers	☐	☐	☐	☐	☐	☐
The conflict potential for society	☐	☐	☐	☐	☐	☐
The chances of realization within the next 5 years	☐	☐	☐	☐	☐	☐
The chances of realization within the next 15 years	☐	☐	☐	☐	☐	☐

Figure 2. Wave 1 questionnaire (English version): the desirability rating and the open comment field shown for the same statement

Would you please indicate how desirable or undesirable you consider the proposed trend **"Double competencies (disciplinary plus IT) will become a universal employment requirement"**?

Very undesirable

☐

Rather undesirable

☐

Moderate

☐

Rather desirable

☐

Very desirable

☐

Prefer not to say / don't know / cannot judge

☐

Comments















The questionnaire opened with information about the survey, data-protection assurances and the privacy policy, and with the respondent's informed consent, in line with the project's GDPR-compliant procedures. As such only respondents who confirmed consent were able to proceed. It then collected the basic and stakeholder information described above before presenting the four thematic blocks of statements in turn, each introduced by a short framing and a self-assessment of expertise. Several technical and routing items were included for administration and are documented, with the full instruments in all languages being present in the Documentation volume as they carry no substantive content and do not enter the analysis.

2.6 Stakeholder groups and recruitment

Eligible respondents were recruited from six principal clusters of stakeholders, with multiple attribution across clusters explicitly permitted (a respondent may, for example, be both a researcher and a former teacher):

- A. Governance and administration: authorities funding and running general and vocational education and training; labour-market services; certification and accreditation bodies.
- B. Business actors: sector and production chambers; employers and their associations; employee, student and teacher unions; investors and sponsors.
- C. Teaching and training staff: teachers and trainers in general and vocational education and training, kept distinct from administrative and technical staff.
- D. Youth and civil society: parent and youth organisations; NGOs and grass-roots groups; media; other stakeholders involved in young people's lives.
- E. Research and academia: individual researchers and networks; private head-hunting / HR firms and consultancies.
- F. International organisations: EU bodies (Commission, CEDEFOP, ETF) and global institutions (UNESCO, OECD, World Bank, EBRD).

Recruitment aimed at a geographically and professionally diverse panel, anchored in the nine EDU-LAB project countries and extended, where appropriate, to other EEA countries and worldwide. National teams ran the recruitment as a low-threshold campaign, and the questionnaire was offered in ten language versions to keep participation accessible. The detailed recruitment strategy, timeline and responsibilities, together with the formal stakeholder definitions and selection criteria, are set out in the Terms of Reference and the recruitment guide.

The questionnaire was programmed once and translated into the national languages by the partner teams, with the English and German versions prepared in-house, so that respondents in each country could participate in their own language. The realised sample (Section 4.1) reflects this design, with the largest national contingents in the countries that ran the most intensive campaigns.

2.7 Fieldwork, data processing and data quality

Administrative frameworks

Wave 1 was administered in LimeSurvey © on a survey instance hosted by the provider, LimeSurvey GmbH, under a data-processing agreement consistent with the project's GDPR framework, in group-by-group mode and ran from 4 March 2026 (first session) to the official field close on 10 April 2026, 13:00. Data were exported to SPSS/PSPP format and cleaned through a single processing chain with an identical structure across all countries (case selection, recoding, missing-value handling). Invalid responses were removed (one completion under five minutes and one identified fake response), and questionnaires submitted after the field close were set aside even where fully completed. Of the completed questionnaires recorded within the regular field period, 435 remained before final case selection and 2 were removed, giving the substantive Wave 1 analysis dataset of N = 433. The same dataset underlies the Quantitative Tables Volume and the Qualitative Volume. A separate metadata analysis documents the process and paradata¹ of the data

¹ Typical survey paradata include the survey start time, completion time, response duration, language version, device type, browser type, page completion times, and break-off points.

collection in full; its main indicators are summarised here because they speak to the quality of the dataset. The Wave 2 feedback materials, prepared while the panel still numbered 435, therefore refer to 435 respondents; the final Wave 1 analysis dataset used throughout this report is N = 433.

The responses to the open-ended questions were exported to Excel together with the key sociodemographic characteristics of the participants. All comments are to be systematically analysed, and the findings of this analysis to be presented in the final deliverable.

Response volume and timing

Responses accumulated fairly steadily across the roughly five-week field period, with pronounced daily peaks consistent with reminder mailings and a clear final push before the cut-off (the strongest single day, 8 April, saw 32 completions). Completions concentrated on weekdays as 91% fell Monday to Friday, led by Wednesday, and followed an office-hours pattern over the day, with the afternoon (12:00–18:00) the busiest window (times in server time, an approximation across ten time zones).

Effort and completion

Completing the questionnaire took a median of 28.6 minutes (lower/upper quartiles 18.7 / 47.6 minutes); 46% of respondents spent 30 minutes or more, and only 1.4% under ten minutes, which is a profile consistent with serious engagement rather than rushed clicking (on questionnaire length and response quality, see Galesic & Bošnjak, 2009). Within the regular field period, 433 completed questionnaires contrasted with 449 abandoned or incomplete sessions, a completion rate of 49.1%. Drop-off was heaviest very early (at the information and data-protection and consent form survey sections and before the thematic sections) and at the entry to the rating section; once respondents began rating, completion was high. The first rating statement took the longest per-page time (median 67 seconds), reflecting an initial learning curve with the rating scheme rather than item difficulty.

Language versions.

The questionnaire was completed in all ten language versions. The table below shows started and completed sessions and completion rates by version; the English version was the most used overall, while completion rates were highest for the Italian, Serbian and Bosnian versions and lowest for Polish.

Table 4. Wave 1 starts and completions by language version

Language version	Started	Completed	Completion rate
English	232	114	49.1%
Finnish	137	72	52.6%
German	136	64	47.1%
Albanian	88	51	58.0%
Polish	81	24	29.6%
Portuguese	72	34	47.2%
Greek	70	30	42.9%
Italian	56	38	67.9%
Bosnian	7	4	57.1%
Serbian	3	2	66.7%

2.8 Analysis approach

Overall remarks

All results of OSES-Delphi are reported mainly descriptively at the stage of issuing this report. For every statement and every dimension, the analysis reports the number of valid substantive ratings, the full frequency distribution across the five scale points, the number of “cannot judge” responses, and the median, interquartile range, mean and standard deviation (overall and for each of the nine project countries). “Cannot judge” answers and item non-response are excluded from the central-tendency and dispersion measures, and percentages are computed on the valid base.

As a consistent reading aid, an IQR of 1.0 or below is described as expert consensus and an IQR of 2.0 or above as marked divergence, consistent with common reporting criteria for Delphi studies (von der Gracht, 2012; Diamond et al., 2014). The complete set of tables is provided in the Quantitative Tables Volume; this report presents the headline figures and patterns.

Scope, limitations and reading guidance

A few points should be borne in mind when reading the results.

First, the OSES panel is a purposively recruited, highly qualified expert panel, not a representative sample of any population; the findings describe expert judgement and expectation, not measured outcomes, and are not weighted.

Second, at this stage the analysis is purely descriptive: it reports distributions, central tendencies and levels of agreement, without inferential testing, modelling or causal interpretation, which are reserved for later project stages.

Third, the per-country figures rest on varying and, for some countries, small numbers of respondents, so country comparisons are indicative rather than precise; the nine project countries are reported individually because they are the study's focus, while smaller national contingents are retained only in the overall figures.

Fourth, the Wave 2 results reported here are final (N = 207; field phase closed 10 June 2026, 18:00).

Finally, "cannot judge" responses are excluded from the medians, IQRs and means and reported separately, and the consensus reading (IQR ≤ 1 as agreement, IQR ≥ 2 as divergence) is a convention adopted consistently for ease of reading rather than a formal statistical test.

Throughout, figures are reported as medians and interquartile ranges on the 1–5 scale unless otherwise stated, and the reader is referred to the Quantitative Tables Volume for the complete distributions, means, standard deviations and per-country detail behind every figure cited in this report.

Robustness checks

Three checks support treating the Wave 1 results as robust. First, on the cut-off: the 19 questionnaires completed shortly after the field close (and therefore excluded from the analysis) were compared with the 433 retained cases. They are broadly similar in substance as statement medians differ only within the range expected from their very small number and with only minor demographic differences (a somewhat lower share of master's-level respondents and a similar age and gender profile). Excluding them, as the field-period rule requires, does not change the descriptive picture.

Second, on the treatment of "cannot judge": the headline figures exclude these non-substantive answers. Re-computing the importance ratings with "cannot judge" instead recoded to the scale midpoint changes the median for only 2 of the 30 statements and the IQR for only 3, so the consensus picture does not hinge on how this option is handled.

Third, on the statistics themselves: medians and interquartile ranges are computed by linear interpolation on the 1–5 scale. Because the scale is discrete and the samples large, the medians are stable and the consensus reading (IQR ≤ 1 / ≥ 2) does not depend on the choice of quartile convention. Full distributions are provided in the Quantitative Tables Volume so that any alternative summary can be checked directly.

3 Results of Wave 1

The chapter opens with the complete set of trend statements. The table below lists all 30 statements with their codes (S01–S30), theme and headline rating on each of the five dimensions; it is the key to the statement codes used in every figure and table that follows. The panel is described next, followed by the cross-dimensional patterns and the detailed, statement-by-statement results.

Table 5. All 30 trend statements at a glance — codes, themes and headline ratings on every dimension

Code	Theme	Statement	Desirability	Importance for policy makers (Md/IQR)	Conflict potential for society	Likelihood in 5 years	Likelihood in 15 years
S01	A	Double competencies (disciplinary plus IT) will become a universal employment requirement	4	4/2.0	3	4	4
S02	A	Illiteracy after eight or nine years of schooling will be the norm	1	4/2.0	4	3	3
S03	A	The COVID-19 crisis will have long-term impacts on learning behaviour and abilities for young people	2	4/1.0	4	4	3
S04	A	Schools will implement effective programs to address behavioral problems (e.g.: violence, bullying, disengagement)	5	4/1.0	4	3	4
S05	A	Public investment in early childhood education will become the most critical part of education	4	4/2.0	3	3	3
S06	A	Education will diversify through vocational pathways that promote specific skills over standardized education	4	4/1.0	3	3	4
S07	A	Focus on university education will neglect TVET (Technical and Vocational Education and Training)	2	3/1.0	3	3	3
S08	A	Society's status preference for universities over vocational training will persist	3	4/1.0	3	4	3
S09	B	New career entry pathways will emerge as AI transforms traditional employment routes	4	4/1.0	4	4	5
S10	B	Work experience requirements for entry-level positions will become more prevalent	3	4/1.0	4	4	4
S11	B	Overqualification will increase due to the growing number of academic qualifications	2	4/1.0	4	4	4
S12	B	Lack of affordable childcare will prevent young parents' labor participation	1	4/2.0	4	4	4
S13	B	Young people will be first to be fired and last to be hired during downturns	2	4/1.0	4	4	3
S14	B	The gap between employer expectations for skilled workers and employer participation in training provision will widen	2	4/1.0	4	3	4
S15	B	Ageing workforce will create massive talent demand in certain sectors	3	4/1.0	4	4	4
S16	C	Mental health challenges among youth stemming from the COVID-19 crisis will continue to affect life skills	1	4/1.5	4	4	3
S17	C	Media and Influencers will significantly influence student program choices	2	4/1.0	4	4	4
S18	C	The transition from education to employment will become more unstable and fragmented for many young people	2	4/2.0	4	4	4
S19	C	Extended juvenile phase will hinder workplace engagement	2	3/1.0	3	3	4
S20	C	Schools will increasingly adapt curricula to emerging occupations and career paths	4	4/2.0	3	3	4

Code	Theme	Statement	Desirability	Importance for policy makers (Md/IQR)	Conflict potential for society	Likelihood in 5 years	Likelihood in 15 years
S21	C	Social partnerships between education, the private sector and academia will be widely formed	5	4/2.0	3	3	4
S22	C	Education and labour market policies will struggle to effectively govern young people's transition	2	4/1.0	4	3	4
S23	C	Integrated guidance centers for youth employment will be widely established	4	4/1.0	3	3	3
S24	D	Educational inequalities will exacerbate social polarization and segregation	1	4/1.0	5	4	4
S25	D	Resource inequities between public and private schools will widen	1	4/2.0	4	4	4
S26	D	Social class (socio-economic background) will continue to determine educational opportunities	1	4/2.0	4	4	4
S27	D	Initiatives breaking gender stereotypes will open previously male-dominated career pathways to women	5	4/2.0	3	3	4
S28	D	School system reliance on parental help will disadvantage non-native and low-education families	1	4/2.0	4	4	4
S29	D	Digital infrastructure and remote learning will reduce regional disparities between urban and rural education	4	4/1.0	3	3	4
S30	D	Minority ethnic and racial groups of young people will face intensified discrimination	1	4/2.0	4	4	4

3.1 The expert panel

The cleaned Wave 1 dataset comprises N = 433 experts. The panel is highly qualified: 376 respondents hold a master's, doctoral or equivalent degree (ISCED 7–8), 48 a bachelor's degree, and 9 an upper-secondary or post-secondary non-tertiary qualification. By gender, 242 respondents are women, 179 men, and 12 preferred not to say. Respondents' ages span 19–71 years (mean 49.4, median 50; 42 did not give an age), so the panel skews toward experienced professionals rather than early-career respondents.

By primary stakeholder affiliation, the largest groups are research and academia (115) and teaching and training staff (111), followed by public governance and administration (84) and business actors (62); civil society (26) and international organisations (10) are also represented, with multiple affiliations permitted. This gives the panel a strong research and practice base across the education and labour-market fields. Respondents also self-assessed their expertise and its geographical scope for each theme; the detailed distributions, with field of study and additional affiliations, are in the Quantitative Tables Volume.

Geographically, the nine EDU-LAB project countries account for 407 of the 433 cases and are reported individually throughout: Finland (71), Austria (58), Kosovo (55), Italy (47), Greece (38), Portugal (38), the UK (36), Poland (34) and Germany (30). A further 26 respondents come from other countries and are retained in all overall figures but not reported as separate country columns. Finally, 287 of the 433 respondents indicated willingness to take part in the second wave (against 146 who declined), defining the potential Wave 2 panel.

3.2 Overall patterns across the five dimensions

Overview of observed response patterns

When considered across the full set of 30 statements, the five dimensions display markedly different distributions and response dynamics.

Expert agreement is greatest on the likelihood of the trends: 24 of the 30 statements reach consensus (IQR ≤ 1) on realisation within five years and 23 within fifteen years. Agreement is weakest on importance and

desirability, where only 17 and 19 statements respectively reach consensus, and where the largest numbers of statements diverge markedly (importance: 12 with IQR ≥ 2 ; desirability: 11). The panel thus agrees more readily on what is likely than on how much it matters or how welcome it would be. The two tables below give the average level of each dimension by theme and the distribution of agreement across dimensions.

Table 6. Dimension profile by theme — mean of statement medians across the five dimensions

Theme	Desirability	Importance	Conflict potential	Likely 5y	Likely 15y
A · Education and Training	3.1	3.9	3.4	3.4	3.4
B · Labour Market and Employment	2.4	4.0	4.0	3.9	4.0
C · Transitions and Systemic Interfaces	2.8	3.9	3.5	3.4	3.8
D · Inclusion and Disadvantage	2.0	4.0	3.9	3.7	4.0

Table 7. Spread of expert opinion by dimension — statements at consensus, moderate spread and divergence

Dimension	Consensus (IQR ≤ 1)	Moderate spread	Divergent (IQR ≥ 2)
Desirability	19	0	11
Importance	17	1	12
Conflict potential	19	1	10
Likelihood 5 yrs	24	1	5
Likelihood 15 yrs	23	0	7

Desirability of discussed trends

Desirability is the dimension on which the four themes differ most. Statements describing positive developments are welcomed with the most desirable trends being S04: Schools will implement effective programs to address behavioural problems (e.g.: violence, bullying, disengagement) and S27: Initiatives breaking gender stereotypes will open previously male-dominated career pathways to women, while the inclusion and labour-market themes contain the survey's least desirable items. The least desirable trends of all, each rated 'very undesirable' on average, are S30: Minority ethnic and racial groups of young people will face intensified discrimination and S24: Educational inequalities will exacerbate social polarization and segregation: developments experts clearly do not wish to see. Desirability is also comparatively contested (11 of 30 statements with IQR ≥ 2), reflecting that experts genuinely differ on whether some developments are to be welcomed.

Importance for policymakers²

Importance is rated moderate-to-high throughout and is the dimension with the widest spread of opinion (12 statements with IQR ≥ 2). The trends seen as most important for policymakers include S04: Schools will implement effective programs to address behavioural problems (e.g.: violence, bullying, disengagement) and S09: New career entry pathways will emerge as AI transforms traditional employment routes; the lowest importance is attached to S19: Extended juvenile phase will hinder workplace engagement. Notably, even where the median importance is high, opinion is often split, so that several high-priority trends are simultaneously among the most divisive to rate.

Conflict potential for society³

Conflict potential is concentrated in the inclusion theme. The statements expected to be most socially divisive are S24: Educational inequalities will exacerbate social polarization and segregation, S30: Minority ethnic and racial groups of young people will face intensified discrimination and S26: Social class (socio-economic background) will continue to determine educational opportunities. These are all concerning inequality, social

² A ranked overview of all 30 statements on this dimension is provided in the annex (Figure A1).

³ A ranked overview of all 30 statements on this dimension is provided in the annex (Figure A2).

background or discrimination while more technical or labour-market trends attract lower conflict ratings. This pattern marks inequality and disadvantage as the politically most sensitive part of the agenda.

Likelihood of realisation (5 and 15 years)

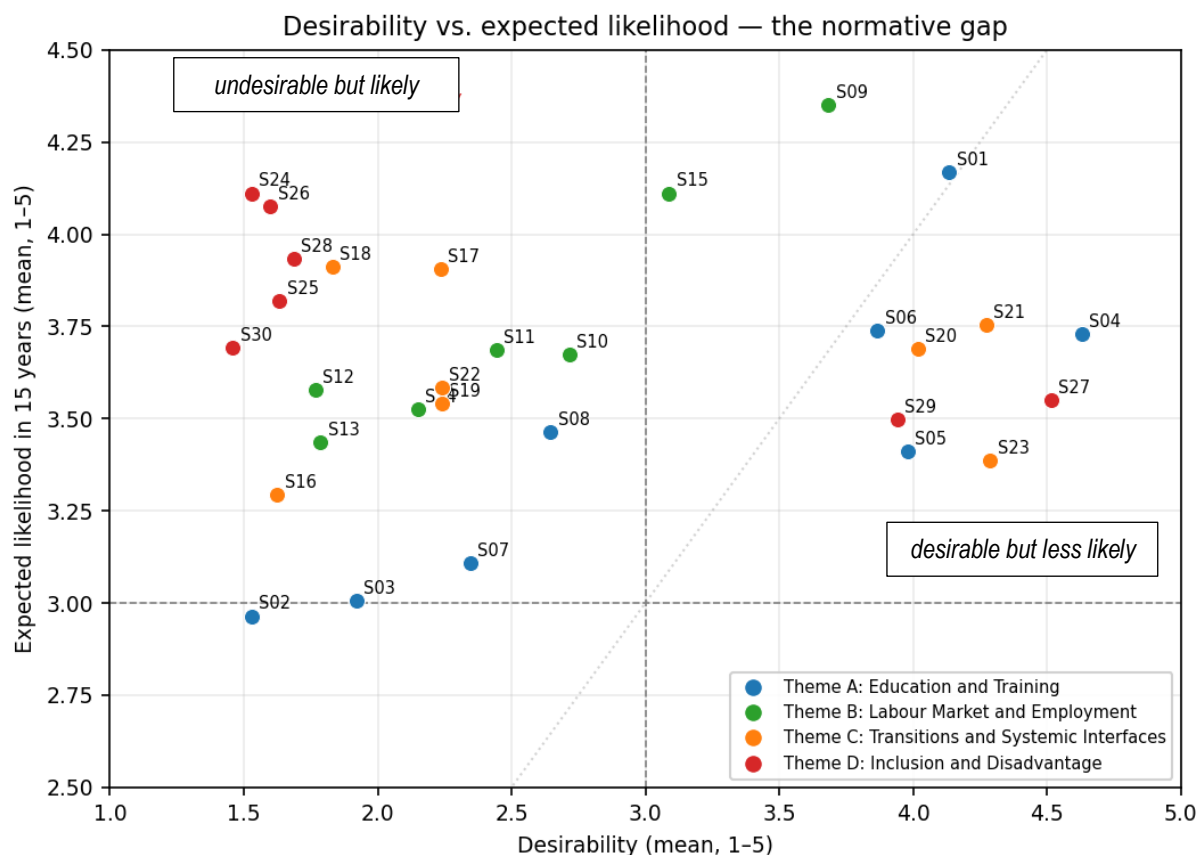
The two likelihood horizons are the dimensions of greatest agreement (24 and 23 of 30 statements at consensus). For most statements the fifteen-year likelihood is equal to or slightly higher than the five-year figure, indicating that experts expect these trends to strengthen over time rather than fade. The combination of high likelihood with low desirability in the inclusion theme is the single most consequential pattern in the data and is taken up statement by statement below.

Two cross-cutting contrasts therefore organise the detailed results: desirability versus likelihood (developments that are wanted but not expected, and others that are unwelcome but expected), and importance versus conflict potential (the most divisive issues being those of inequality and disadvantage).

Uncertainty

A further cross-cutting observation concerns uncertainty. The "cannot judge" option was used sparingly overall, which indicates that respondents generally felt able to assess the trends, but its use was uneven: it was somewhat more frequent on the desirability and the longer-horizon likelihood judgements of a handful of statements, where experts were evidently less comfortable committing to a rating. Where this is the case for an individual statement it is noted in the walk-through below, since a high share of "cannot judge" points to genuine uncertainty rather than to disagreement among those who did rate.

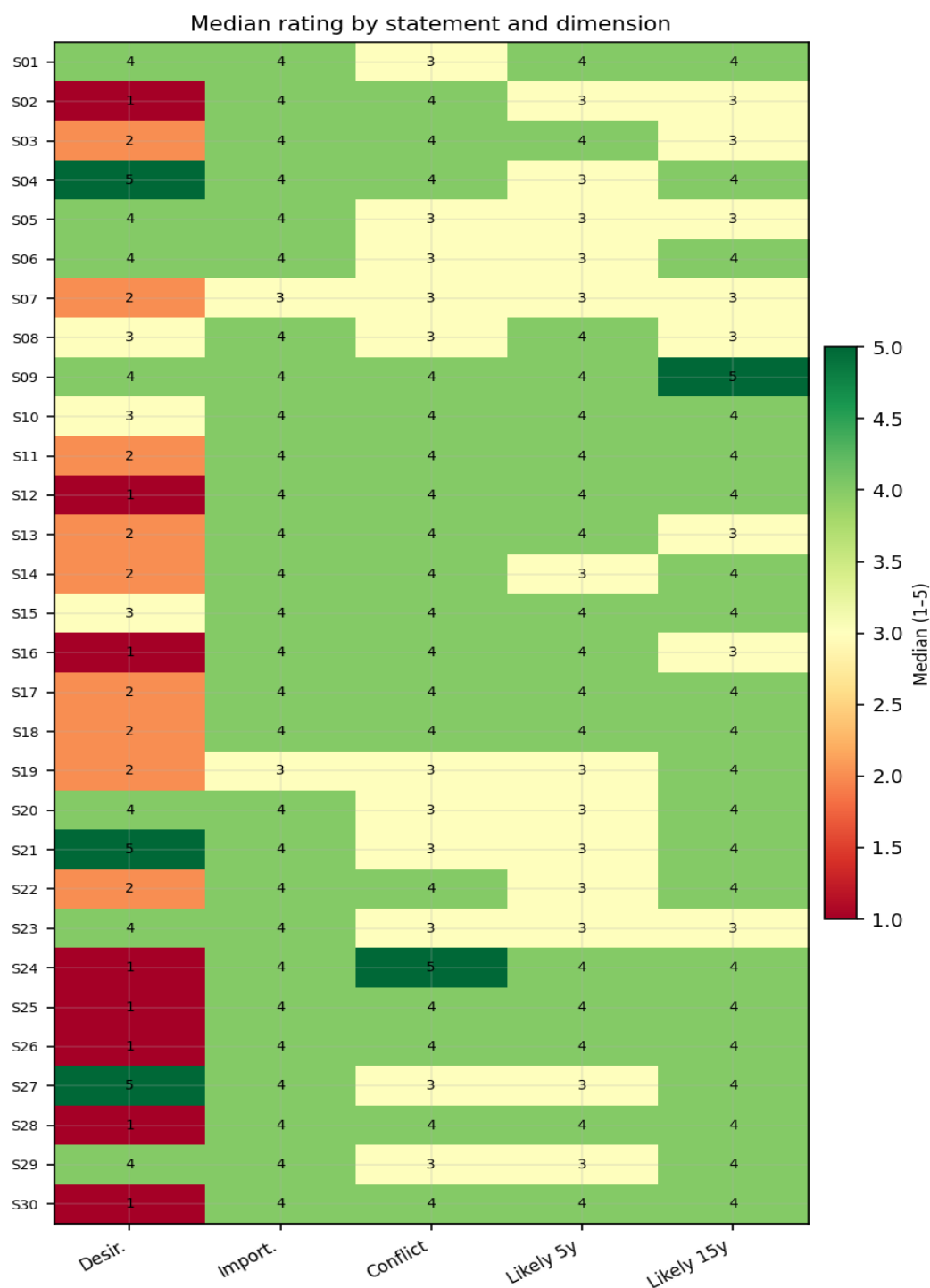
Figure 3. Each statement positioned by mean desirability (x) and mean expected likelihood in 15 years (y). The upper-left region (low desirability, high likelihood) collects the inclusion statements (Theme D). (Statement codes S01–S30 and their full wording are



3.3 Results by trend statements

This section walks through all 30 statements, grouped by theme. For each statement it reports, descriptively, how the panel rated it on the five dimensions, how much agreement there was, and any notable variation across the project countries with a compact profile table accompanying each. The 'results at a glance' table below first lists all 30 statements with their headline medians. All figures are overall medians and interquartile ranges (IQR) on the 1–5 scale; full distributions, means, standard deviations and per-country tables are in the Quantitative Tables Volume.

Figure 4. Median rating (1–5) for every statement and dimension. Greener = higher; note the low-desirability / high-importance pattern of the inclusion statements (S24–S30). (Statements are listed in Table 4; read each row across the five dimensions for one statement's full profile)



Theme A — Education and Training

The thematic section A “Education and Training” contains several of the panel's clearest points of agreement alongside some of its most divided judgements: statements on the COVID-19 learning legacy, behavioural problems, vocational diversification and the status preference for universities reach consensus on importance, while double competencies and basic-skills deficits are far more contested. Desirability is moderate, since the theme mixes welcome developments with risks.

S01 Double competencies (disciplinary plus IT) will become a universal employment requirement

Diverging: the statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with moderate conflict potential (median 3), and is high in likelihood over both horizons (median 4). Experts consider it desirable (median 4).

Table 8. S01 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	4	1.0	4.14	0.86	Very desirable
Importance	4	2.0	3.97	0.86	High
Conflict potential	3	1.0	3.44	0.96	Moderate
Likelihood 5 yrs	4	1.0	3.60	0.96	High
Likelihood 15 yrs	4	1.0	4.17	0.93	Very high

Table 9. S01 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (2.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)

S02 Illiteracy after eight or nine years of schooling will be the norm

Controversial: the statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with high conflict potential (median 4), and is moderate in likelihood over both horizons (median 3), though its timing is contested. Experts consider it undesirable (median 1) — an unwelcome yet expected development. Importance varies across the project countries (highest in Finland, lowest in Kosovo).

Table 10. S02 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	1	1.0	1.53	1.08	Very undesirable
Importance	4	2.0	3.72	1.24	Very high
Conflict potential	4	2.0	3.87	1.16	Very high
Likelihood 5 yrs	3	1.5	2.72	1.13	Moderate
Likelihood 15 yrs	3	2.0	2.96	1.22	Moderate

Table 11. S02 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.0)	3 (3.0)	4 (2.0)	4 (1.8)	4 (2.0)	4 (2.0)	3 (2.0)	4 (3.0)

S03 The COVID-19 crisis will have long-term impacts on learning behaviour and abilities for young people

Consensus: the statement is broadly rated important (median 4), with high conflict potential (median 4), and is more likely in the near term (median 4 at five years), though its timing is contested. Experts consider it undesirable (median 2) — an unwelcome yet expected development. Importance varies across the project countries (highest in Finland, lowest in Austria).

Table 12. S03 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	2	2.0	1.92	1.15	Very undesirable
Importance	4	1.0	3.51	1.08	High
Conflict potential	4	1.0	3.52	1.05	High
Likelihood 5 yrs	4	1.0	3.42	1.10	High
Likelihood 15 yrs	3	2.0	3.01	1.20	Moderate

Table 13. S03 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	3 (1.0)	4 (2.0)	3 (1.0)	4 (2.0)	3 (2.2)	4 (1.0)	4 (2.0)	4 (1.0)

S04 Schools will implement effective programs to address behavioral problems (e.g.: violence, bullying, disengagement)

Consensus: the statement is broadly rated important (median 4), with high conflict potential (median 4), and is more likely over the longer 15-year horizon (median 4, vs 3 at five years). Experts consider it desirable (median 5) — welcomed but less firmly expected. Importance varies across the project countries (highest in Greece, lowest in Finland).

Table 14. S04 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	5	0.0	4.63	0.76	Very desirable
Importance	4	1.0	4.16	0.88	Very high
Conflict potential	4	2.0	3.97	1.02	Very high
Likelihood 5 yrs	3	1.0	3.40	0.99	High
Likelihood 15 yrs	4	1.0	3.73	0.96	High

Table 15. S04 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.0)	4 (2.0)	4 (2.0)	5 (1.0)	4 (1.0)	4 (1.0)	4 (1.2)	4 (1.0)

S05 Public investment in early childhood education will become the most critical part of education

Diverging: the statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with moderate conflict potential (median 3), and is moderate in likelihood over both horizons (median 3), though its timing is contested. Experts consider it desirable (median 4) — welcomed but less firmly expected. Importance varies across the project countries (highest in Austria, lowest in Finland).

Table 16. S05 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	4	2.0	3.98	1.18	Very desirable
Importance	4	2.0	3.80	1.07	Very high
Conflict potential	3	1.0	3.36	1.11	Moderate
Likelihood 5 yrs	3	2.0	3.10	1.07	Moderate
Likelihood 15 yrs	3	1.0	3.41	1.08	Moderate

Table 17. S05 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	3 (1.0)	4 (2.0)	4 (1.0)	4 (2.0)	4 (2.0)	4 (2.0)	4 (1.0)	3 (2.0)	4 (1.0)

S06 Education will diversify through vocational pathways that promote specific skills over standardized education

The statement is broadly rated important (median 4), with moderate conflict potential (median 3), and is more likely over the longer 15-year horizon (median 4, vs 3 at five years). Experts consider it desirable (median 4).

Table 18. S06 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	4	2.0	3.87	1.00	Rather desirable
Importance	4	1.0	3.80	0.90	High
Conflict potential	3	1.0	3.24	0.98	Moderate
Likelihood 5 yrs	3	1.0	3.33	0.94	Moderate
Likelihood 15 yrs	4	1.0	3.74	0.91	High

Table 19. S06 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.0)	4 (2.0)	4 (2.0)	4 (1.0)	4 (0.0)	4 (1.0)	4 (1.0)	4 (1.5)

S07 Focus on university education will neglect TVET (Technical and Vocational Education and Training)

Controversial: the statement is of moderate importance (median 3), with moderate conflict potential (median 3), and is moderate in likelihood over both horizons (median 3), though its timing is contested. Experts consider it undesirable (median 2). Importance varies across the project countries (highest in Finland, lowest in Austria).

Table 20. S07 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	2	1.0	2.35	1.08	Rather undesirable
Importance	3	1.0	3.38	0.99	Moderate
Conflict potential	3	1.0	3.28	0.93	Moderate
Likelihood 5 yrs	3	2.0	3.06	0.95	Moderate
Likelihood 15 yrs	3	2.0	3.11	0.99	Moderate

Table 21. S07 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	3 (1.0)	3 (1.0)	4 (1.0)	4 (2.0)	3 (1.0)	3 (1.0)	3 (2.0)	3 (2.0)

S08 Society's status preference for universities over vocational training will persist

The statement is broadly rated important (median 4), with moderate conflict potential (median 3), and is more likely in the near term (median 4 at five years). Desirability is moderate (median 3). Importance varies across the project countries (highest in Finland, lowest in Poland).

Table 22. S08 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	3	1.0	2.65	1.09	Moderate
Importance	4	1.0	3.65	0.90	High
Conflict potential	3	1.0	3.41	0.90	Moderate
Likelihood 5 yrs	4	1.0	3.55	0.94	High
Likelihood 15 yrs	3	1.0	3.46	0.94	Moderate

Table 23. S08 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	3 (1.0)	4 (1.0)

Theme A in summary. Within this theme, the statement rated most important for policymakers is S04 (Schools will implement effective programs to address behavioural problems (e.g.: violence, bullying, disengagement)), while opinion on importance is most divided for S01. The least desirable development in the theme is S02, and the highest conflict potential attaches to S04. These within-theme contrasts feed directly into the selection of statements for the second wave (Section 5).

Theme B — Labour Market and Employment

Labour Market and Employment statements are rated consistently important and comparatively likely, combining firmly agreed trends (AI-driven entry pathways, an ageing workforce creating talent demand) with one of the survey's most cross-nationally variable items, on affordable childcare and young parents' labour participation. Average desirability is low, reflecting that many trends describe pressures rather than opportunities.

S09 New career entry pathways will emerge as AI transforms traditional employment routes

Consensus: the statement is broadly rated important (median 4), with high conflict potential (median 4), and is more likely over the longer 15-year horizon (median 5, vs 4 at five years), though its timing is contested. Experts consider it desirable (median 4). Importance varies across the project countries (highest in Greece, lowest in Finland).

Table 24. S09 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	4	1.0	3.68	0.94	Rather desirable
Importance	4	1.0	4.16	0.82	High
Conflict potential	4	2.0	3.97	0.86	High
Likelihood 5 yrs	4	2.0	3.98	0.86	High
Likelihood 15 yrs	5	1.0	4.35	0.77	Very high

Table 25. S09 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.0)	4 (2.0)	4 (1.0)	5 (1.0)	4 (1.0)	5 (1.0)	4 (1.0)	4 (1.0)

S10 Work experience requirements for entry-level positions will become more prevalent

The statement is broadly rated important (median 4), with high conflict potential (median 4), and is high in likelihood over both horizons (median 4). Desirability is moderate (median 3). Importance varies across the project countries (highest in Finland, lowest in Italy).

Table 26. S10 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	3	2.0	2.72	1.15	Rather undesirable
Importance	4	1.0	3.57	0.89	High
Conflict potential	4	1.0	3.60	0.95	High
Likelihood 5 yrs	4	1.0	3.62	0.90	High
Likelihood 15 yrs	4	1.0	3.67	1.00	High

Table 27. S10 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.0)	4 (1.0)	3 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	3 (1.5)

S11 · Overqualification will increase due to the growing number of academic qualifications

The statement is broadly rated important (median 4), with high conflict potential (median 4), and is high in likelihood over both horizons (median 4). Experts consider it undesirable (median 2) — an unwelcome yet expected development. Importance varies across the project countries (highest in Kosovo, lowest in Finland).

Table 28. S11 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	2	1.0	2.44	1.05	Rather undesirable
Importance	4	1.0	3.58	1.04	High
Conflict potential	4	1.0	3.51	1.00	High
Likelihood 5 yrs	4	1.0	3.54	0.98	High
Likelihood 15 yrs	4	1.0	3.68	1.04	High

Table 29. S11 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	3 (2.0)	4 (1.0)	4 (1.5)	4 (1.0)	4 (1.0)	4 (2.0)	4 (2.0)	4 (1.0)	4 (1.2)

S12 · Lack of affordable childcare will prevent young parents' labor participation

Consensus: the statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with high conflict potential (median 4), and is high in likelihood over both horizons (median 4). Experts consider it undesirable (median 1) — an unwelcome yet expected development. Importance varies across the project countries (highest in Austria, lowest in Finland).

Table 30. S12 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	1	1.0	1.77	1.19	Very undesirable
Importance	4	2.0	3.95	1.04	Very high
Conflict potential	4	2.0	3.85	1.04	High
Likelihood 5 yrs	4	1.0	3.52	1.05	High
Likelihood 15 yrs	4	1.0	3.58	1.08	High

Table 31. S12 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	3 (1.0)	5 (1.0)	4 (2.0)	4 (1.5)	5 (1.0)	4 (2.0)	4 (1.0)	4 (2.0)	4 (1.0)

S13 · Young people will be first to be fired and last to be hired during downturns

Diverging: the statement is broadly rated important (median 4), with high conflict potential (median 4), and is more likely in the near term (median 4 at five years). Experts consider it undesirable (median 2) — an unwelcome yet expected development. Importance varies across the project countries (highest in Austria, lowest in Finland).

Table 32. S13 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	2	1.0	1.79	0.97	Very undesirable
Importance	4	1.0	3.63	1.02	High
Conflict potential	4	2.0	3.86	0.93	High
Likelihood 5 yrs	4	1.0	3.45	1.00	High
Likelihood 15 yrs	3	1.0	3.43	1.07	High

Table 33. S13 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	3 (1.0)	4 (1.0)	4 (2.0)	4 (2.0)	4 (2.0)	4 (1.2)	4 (1.0)	4 (1.0)	3 (2.0)

S14 · The gap between employer expectations for skilled workers and employer participation in training provision will widen

Consensus: the statement is broadly rated important (median 4), with high conflict potential (median 4), and is more likely over the longer 15-year horizon (median 4, vs 3 at five years). Experts consider it undesirable (median 2) — an unwelcome yet expected development. Importance varies across the project countries (highest in Austria, lowest in Finland).

Table 34. S14 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	2	2.0	2.15	1.18	Very undesirable
Importance	4	1.0	3.61	0.98	High
Conflict potential	4	1.0	3.46	0.95	High
Likelihood 5 yrs	3	1.0	3.44	0.87	Moderate
Likelihood 15 yrs	4	1.0	3.52	0.96	Moderate

Table 35. S14 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	3 (1.0)	4 (1.0)	4 (1.0)	3 (1.0)	4 (2.0)	3 (1.0)	4 (1.0)	4 (2.0)	4 (1.0)

S15 · Ageing workforce will create massive talent demand in certain sectors

The statement is broadly rated important (median 4), with high conflict potential (median 4), and is high in likelihood over both horizons (median 4). Desirability is moderate (median 3). Importance varies across the project countries (highest in Greece, lowest in Finland).

Table 36. S15 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	3	2.0	3.09	1.24	Moderate
Importance	4	1.0	4.03	0.88	High
Conflict potential	4	1.0	3.72	0.97	High
Likelihood 5 yrs	4	1.0	3.77	0.85	High
Likelihood 15 yrs	4	1.0	4.11	0.90	Very high

Table 37. S15 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.5)	5 (1.0)	4 (1.0)	4 (1.0)	4 (0.0)	5 (1.0)

Theme B in summary. Within this theme, the statement rated most important for policymakers is S09 (New career entry pathways will emerge as AI transforms traditional employment routes), while opinion on importance is most divided for S12. The least desirable development in the theme is S12, and the highest conflict potential attaches to S09. These within-theme contrasts feed directly into the selection of statements for the second wave (Section 5).

Theme C — Transitions and Systemic Interfaces

Transitions and Systemic Interfaces is where the desirability-likelihood gap is most visible: developments experts would welcome (integrated guidance centres, stronger social partnerships) are expected only moderately, while more negative dynamics (unstable, fragmented transitions; lingering youth mental-health effects) attract broad agreement on their importance.

S16 · Mental health challenges among youth stemming from the COVID-19 crisis will continue to affect life skills

Diverging: the statement is broadly rated important (median 4), with high conflict potential (median 4), and is more likely in the near term (median 4 at five years). Experts consider it undesirable (median 1) — an unwelcome yet expected development. Importance varies across the project countries (highest in Finland, lowest in Kosovo).

Table 38. S16 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	1	1.0	1.63	0.99	Very undesirable
Importance	4	1.5	3.65	1.10	High
Conflict potential	4	1.0	3.68	1.01	High
Likelihood 5 yrs	4	1.0	3.62	1.03	High
Likelihood 15 yrs	3	1.0	3.29	1.07	Moderate

Table 39. S16 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.8)	4 (2.0)	3 (1.0)	4 (1.0)	4 (2.0)	3 (1.0)	4 (1.0)	4 (2.0)	4 (0.5)

S17 · Media and Influencers will significantly influence student program choices

Diverging: the statement is broadly rated important (median 4), with high conflict potential (median 4), and is high in likelihood over both horizons (median 4), though its timing is contested. Experts consider it undesirable (median 2) — an unwelcome yet expected development. Importance varies across the project countries (highest in Austria, lowest in Finland).

Table 40. S17 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	2	2.0	2.24	1.06	Rather undesirable
Importance	4	1.0	3.65	1.01	High
Conflict potential	4	1.0	3.62	1.01	High
Likelihood 5 yrs	4	1.0	3.84	0.91	High
Likelihood 15 yrs	4	2.0	3.91	0.98	High

Table 41. S17 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
------------------------------------	----	----	----	----	----	----	----	----	----

median (IQR)	3 (1.0)	4 (1.0)	4 (1.0)	4 (2.0)	4 (2.0)	4 (1.0)	4 (1.5)	4 (2.0)	4 (1.0)
--------------	---------	---------	---------	---------	---------	---------	---------	---------	---------

S18 · The transition from education to employment will become more unstable and fragmented for many young people

Uncertain: the statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with high conflict potential (median 4), and is high in likelihood over both horizons (median 4), though its timing is contested. Experts consider it undesirable (median 2) — an unwelcome yet expected development. Importance varies across the project countries (highest in Greece, lowest in Finland).

Table 42. S18 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	2	1.0	1.83	1.03	Very undesirable
Importance	4	2.0	3.87	0.99	High
Conflict potential	4	2.0	3.92	0.85	High
Likelihood 5 yrs	4	1.0	3.82	0.89	High
Likelihood 15 yrs	4	2.0	3.91	0.93	High

Table 43. S18 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.8)	4 (2.0)	4 (1.0)	4 (2.0)	5 (1.0)	4 (2.0)	4 (2.0)	4 (3.0)	4 (2.0)

S19 · Extended juvenile phase will hinder workplace engagement

The statement is of moderate importance (median 3), with moderate conflict potential (median 3), and is more likely over the longer 15-year horizon (median 4, vs 3 at five years). Experts consider it undesirable (median 2) — an unwelcome yet expected development. Importance varies across the project countries (highest in Greece, lowest in Finland).

Table 44. S19 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	2	2.0	2.24	1.01	Rather undesirable
Importance	3	1.0	3.35	1.03	Moderate
Conflict potential	3	1.0	3.40	1.04	Moderate
Likelihood 5 yrs	3	1.0	3.43	0.94	Moderate
Likelihood 15 yrs	4	1.0	3.54	1.00	High

Table 45. S19 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	3 (1.0)	3 (1.0)	3 (1.0)	3 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	3 (1.0)	3 (1.0)

S20 · Schools will increasingly adapt curricula to emerging occupations and career paths

Consensus: the statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with moderate conflict potential (median 3), and is more likely over the longer 15-year horizon (median 4, vs 3 at five years). Experts consider it desirable (median 4).

Table 46. S20 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	4	2.0	4.02	1.07	Very desirable
Importance	4	2.0	3.89	0.93	High
Conflict potential	3	1.0	3.29	1.02	Moderate

Dimension	Median	IQR	Mean	SD	Most frequent rating
Likelihood 5 yrs	3	1.0	3.33	0.98	Moderate
Likelihood 15 yrs	4	1.0	3.69	0.95	High

Table 47. S20 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (2.0)	4 (2.0)	4 (0.2)	4 (1.2)	4 (1.0)	4 (1.0)	4 (1.8)	4 (1.0)

S21 · Social partnerships between education, the private sector and academia will be widely formed

Diverging: the statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with moderate conflict potential (median 3), and is more likely over the longer 15-year horizon (median 4, vs 3 at five years). Experts consider it desirable (median 5) — welcomed but less firmly expected.

Table 48. S21 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	5	1.0	4.27	0.89	Very desirable
Importance	4	2.0	3.91	0.92	High
Conflict potential	3	2.0	2.96	1.14	Moderate
Likelihood 5 yrs	3	1.0	3.38	0.93	Moderate
Likelihood 15 yrs	4	1.0	3.75	0.90	High

Table 49. S21 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (2.0)	4 (1.2)	4 (2.0)	4 (1.0)	4 (1.0)	4 (2.0)	4 (1.0)	4 (1.0)	4 (1.0)

S22 · Education and labour market policies will struggle to effectively govern young people's transition

The statement is broadly rated important (median 4), with high conflict potential (median 4), and is more likely over the longer 15-year horizon (median 4, vs 3 at five years). Experts consider it undesirable (median 2) — an unwelcome yet expected development.

Table 50. S22 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	2	2.0	2.24	1.18	Rather undesirable
Importance	4	1.0	3.86	0.90	High
Conflict potential	4	1.0	3.59	0.88	High
Likelihood 5 yrs	3	1.0	3.46	0.86	High
Likelihood 15 yrs	4	1.0	3.58	0.90	High

Table 51. S22 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)

S23 · Integrated guidance centers for youth employment will be widely established

Diverging: the statement is broadly rated important (median 4), with moderate conflict potential (median 3), and is moderate in likelihood over both horizons (median 3), though its timing is contested. Experts consider

it desirable (median 4) — welcomed but less firmly expected. Importance varies across the project countries (highest in Finland, lowest in Portugal).

Table 52. S23 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	4	1.0	4.29	0.83	Very desirable
Importance	4	1.0	3.68	1.02	High
Conflict potential	3	2.0	2.87	1.12	Moderate
Likelihood 5 yrs	3	2.0	3.02	0.96	Moderate
Likelihood 15 yrs	3	1.0	3.39	0.97	Moderate

Table 53. S23 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.0)	4 (2.0)	4 (1.8)	4 (1.2)	3 (1.0)	4 (1.2)	3 (1.0)	4 (1.0)

Theme C in summary. Within this theme, the statement rated most important for policymakers is S21 (Social partnerships between education, the private sector and academia will be widely formed), while opinion on importance is most divided for S18. The least desirable development in the theme is S16, and the highest conflict potential attaches to S18. These within-theme contrasts feed directly into the selection of statements for the second wave (Section 5).

Theme D — Inclusion and Disadvantage

Inclusion and Disadvantage produces the report's most striking pattern: the lowest desirability ratings in the survey combined with among the highest importance, conflict-potential and likelihood ratings — developments experts find unwelcome, divisive and yet likely. The statement on widening educational inequality records the highest conflict potential of all 30 statements.

S24 · Educational inequalities will exacerbate social polarization and segregation

Consensus: the statement is broadly rated important (median 4), with very high conflict potential (median 5), and is high in likelihood over both horizons (median 4), though its timing is contested. Experts consider it undesirable (median 1) — an unwelcome yet expected development. Importance varies across the project countries (highest in Greece, lowest in Finland).

Table 54. S24 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	1	1.0	1.53	1.04	Very undesirable
Importance	4	1.0	4.12	0.95	Very high
Conflict potential	5	1.0	4.31	0.83	Very high
Likelihood 5 yrs	4	2.0	3.95	0.90	High
Likelihood 15 yrs	4	1.0	4.11	0.88	Very high

Table 55. S24 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.0)	4 (2.0)	4 (2.0)	5 (1.0)	4 (2.0)	5 (1.0)	4 (1.0)	4 (1.0)

S25 · Resource inequities between public and private schools will widen

Uncertain: the statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with high conflict potential (median 4), and is high in likelihood over both horizons (median 4), though its timing is contested. Experts consider it undesirable (median 1) — an unwelcome yet expected development. Importance varies across the project countries (highest in Greece, lowest in Finland).

Table 56. S25 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	1	1.0	1.63	1.04	Very undesirable
Importance	4	2.0	3.68	1.15	Very high
Conflict potential	4	2.0	3.90	1.03	High
Likelihood 5 yrs	4	1.0	3.60	1.05	High
Likelihood 15 yrs	4	2.0	3.82	1.10	Very high

Table 57. S25 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	3 (2.0)	4 (2.0)	4 (2.0)	4 (1.0)	5 (1.0)	4 (2.0)	4 (2.0)	3 (1.0)	3 (1.0)

S26 · Social class (socio-economic background) will continue to determine educational opportunities

The statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with high conflict potential (median 4), and is high in likelihood over both horizons (median 4). Experts consider it undesirable (median 1) — an unwelcome yet expected development. Importance varies across the project countries (highest in Greece, lowest in Finland).

Table 58. S26 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	1	1.0	1.60	1.05	Very undesirable
Importance	4	2.0	3.93	1.05	Very high
Conflict potential	4	1.0	4.19	0.86	Very high
Likelihood 5 yrs	4	1.0	4.01	0.89	High
Likelihood 15 yrs	4	1.0	4.08	0.85	High

Table 59. S26 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.5)	4 (1.5)	4 (2.0)	4 (1.0)	5 (1.0)	4 (1.5)	4 (1.0)	4 (2.0)	4 (1.0)

S27 · Initiatives breaking gender stereotypes will open previously male-dominated career pathways to women

Consensus: the statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with moderate conflict potential (median 3), and is more likely over the longer 15-year horizon (median 4, vs 3 at five years). Experts consider it desirable (median 5) — welcomed but less firmly expected. Importance varies across the project countries (highest in Greece, lowest in Portugal).

Table 60. S27 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	5	1.0	4.52	0.85	Very desirable
Importance	4	2.0	3.76	1.05	High
Conflict potential	3	1.0	3.31	1.06	Moderate
Likelihood 5 yrs	3	1.0	3.21	0.96	Moderate
Likelihood 15 yrs	4	1.0	3.55	0.97	Moderate

Table 61. S27 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (1.0)	4 (1.5)	4 (2.0)	4 (2.0)	5 (2.0)	3 (1.0)	4 (1.0)	4 (1.0)	4 (1.2)

S28 · School system reliance on parental help will disadvantage non-native and low-education families

Uncertain: the statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with high conflict potential (median 4), and is high in likelihood over both horizons (median 4), though its timing is contested. Experts consider it undesirable (median 1) — an unwelcome yet expected development. Importance varies across the project countries (highest in Greece, lowest in Finland).

Table 62. S28 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	1	1.0	1.69	1.05	Very undesirable
Importance	4	2.0	3.84	1.08	High
Conflict potential	4	1.2	4.05	0.88	High
Likelihood 5 yrs	4	1.0	3.85	0.91	High
Likelihood 15 yrs	4	2.0	3.93	0.92	High

Table 63. S28 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (2.0)	4 (1.5)	4 (1.8)	4 (1.0)	5 (2.0)	4 (1.0)	4 (1.8)	4 (1.0)	4 (1.0)

S29 · Digital infrastructure and remote learning will reduce regional disparities between urban and rural education

Diverging: the statement is broadly rated important (median 4), with moderate conflict potential (median 3), and is more likely over the longer 15-year horizon (median 4, vs 3 at five years). Experts consider it desirable (median 4). Importance varies across the project countries (highest in Kosovo, lowest in Finland).

Table 64. S29 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	4	2.0	3.94	0.98	Rather desirable
Importance	4	1.0	3.54	1.01	Moderate
Conflict potential	3	2.0	3.08	1.06	Moderate
Likelihood 5 yrs	3	1.0	3.15	0.95	Moderate
Likelihood 15 yrs	4	1.0	3.50	1.00	High

Table 65. S29 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	3 (1.0)	3 (1.0)	4 (2.0)	4 (2.0)	4 (2.0)	3 (1.0)	3 (1.0)	3 (1.0)	3 (1.0)

S30 · Minority ethnic and racial groups of young people will face intensified discrimination

The statement is broadly rated important (median 4), though views on its importance vary widely (IQR 2.0), with high conflict potential (median 4), and is high in likelihood over both horizons (median 4). Experts consider it undesirable (median 1) — an unwelcome yet expected development.

Table 66. S30 — rating profile across the five dimensions (median, IQR, mean, SD, most frequent rating)

Dimension	Median	IQR	Mean	SD	Most frequent rating
Desirability	1	0.0	1.46	0.97	Very undesirable
Importance	4	2.0	3.81	1.07	High
Conflict potential	4	1.0	4.21	0.92	Very high
Likelihood 5 yrs	4	1.0	3.69	0.92	High
Likelihood 15 yrs	4	1.0	3.69	0.97	High

Table 67. S30 — importance rating by project country (median and IQR)

Importance by country (Md, IQR)	FI	AT	XK	IT	GR	PT	UK	PL	DE
median (IQR)	4 (2.0)	4 (1.0)	4 (2.0)	4 (1.0)	4 (2.0)	4 (1.0)	4 (1.0)	4 (1.0)	4 (1.0)

Theme D in summary. Within this theme, the statement rated most important for policymakers is S24 (Educational inequalities will exacerbate social polarization and segregation), while opinion on importance is most divided for S25. The least desirable development in the theme is S30, and the highest conflict potential attaches to S24. These within-theme contrasts feed directly into the selection of statements for the second wave (Section 5).

3.4 Consensus and controversy

Read together, the Wave 1 results are best described not by a simple agreement/disagreement split but by four patterns, assigned to each statement from its full response structure: the median, the interquartile range, the shape of the distribution and a comparative reading across statements. A consensus pattern shows a clear directional tendency held with high consistency; diverging views show a clear direction but with substantial variation in intensity; uncertainty shows dispersion with only a weak or unconsolidated shared tendency, often with a higher share of 'cannot judge' responses; and polarised shows no shared direction at all, with substantial shares at both ends of the scale, meaning fundamentally opposing positions rather than mere spread. The category of polarised responses is therefore deliberately narrow: a wide interquartile range alone does not make a statement controversial, and only S02 (illiteracy becoming the norm) clearly shows this opposing-positions pattern, with S07 (a university focus neglecting TVET) treated as a borderline case on the same logic. This four-pattern classification of the Wave 1 results is what defines the four blocks of the Wave 2 reassessment, whose full statement assignment is given in Section 5.2. Cutting across the classification, three descriptive patterns also stand out in their own right: the statements with the highest perceived conflict potential, those judged undesirable yet likely, and those judged desirable yet less likely:

Table 68. Three cross-cutting descriptive patterns (these are not the four-pattern classification; they highlight, respectively, the statements with the highest conflict potential, those judged undesirable yet likely, and those judged desirable yet less likely).

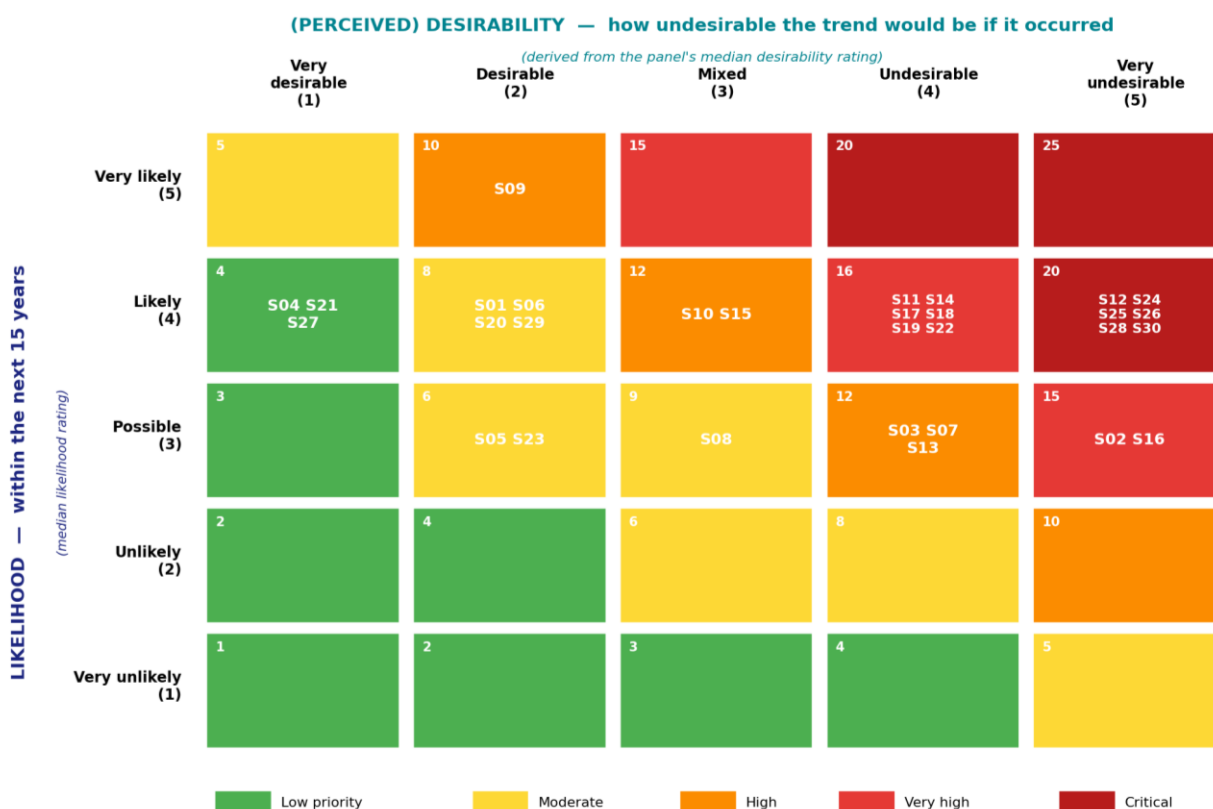
Pattern	Statement	Key figures
Highest conflict potential	S24: Educational inequalities will exacerbate social polarization and segregation	Md 5, mean 4.31
	S30: Minority ethnic and racial groups of young people will face intensified discrimination	Md 4, mean 4.21
	S26: Social class (socio-economic background) will continue to determine educational opportunities	Md 4, mean 4.19
	S28: School system reliance on parental help will disadvantage non-native and low-education families	Md 4, mean 4.05
Undesirable yet likely	S24: Educational inequalities will exacerbate social polarization and segregation	des. 1.53 vs L15y 4.11
	S26: Social class (socio-economic background) will continue to determine educational opportunities	des. 1.60 vs L15y 4.08
	S28: School system reliance on parental help will disadvantage non-native and low-education families	des. 1.69 vs L15y 3.93
Desirable yet less likely	S27: Initiatives breaking gender stereotypes will open previously male-dominated career pathways to women	des. 4.52 vs L15y 3.55
	S23: Integrated guidance centers for youth employment will be widely established	des. 4.29 vs L15y 3.39
	S04: Schools will implement effective programs to address behavioural problems (e.g.: violence, bullying, disengagement)	des. 4.63 vs L15y 3.73

Together, this three-pattern classification and these descriptive contrasts set the agenda for the second wave, which re-rates the controversial statements on all dimensions, revisits the diverging and uncertain statements where opinion is unsettled, and deepens the consensus statements with a forward-looking

policy-feasibility question (Section 5). In the Wave 2 feedback shown to respondents, the four patterns were presented as Shared view, Diverging in intensity, Uncertain and Polarised, respectively.

The same desirability and likelihood information can be condensed into a single display such as a Policy Priority Matrix (Figure 5). Neither risk nor impact was measured in the survey: respondents rated the desirability and the likelihood of the trends, and the matrix combines exactly these two ratings descriptively. Each statement is placed by the panel's median likelihood that it is realised within fifteen years (vertical) and by how undesirable it would be if it occurred, derived from the median desirability rating (horizontal); a priority score (undesirability × likelihood, 1–25) shades each cell from green (low) to dark red (critical). Read this way, the statements separate into two policy-relevant regions: a red, high-priority cluster of undesirable-and-likely developments, dominated by the inclusion theme, and a green-to-yellow region of desirable or less likely developments. The display is an illustrative visualisation of the panel's collective judgement, not a formal risk assessment, which would require dedicated probability and severity measures. Re-drawing the matrix with the five-year likelihood changes the picture only at the margins: the eleven-statement core of the high-priority region (S02, S11, S12, S16, S17, S18, S24, S25, S26, S28, S30) is identical under both horizons; near-term pressures (S03, the COVID-19 learning effects, and S13, young people's labour-market exposure) enter it only in the five-year view, while slower-moving institutional trends (S14, S19, S22) reach it only over fifteen years.

Figure 5. Policy Priority Matrix of the 30 trend statements.



Priority score = undesirability × likelihood (bands: 1–4 low · 5–9 moderate · 10–14 high · 15–19 very high · ≥ 20 critical).

Illustrative visualisation of the panel's median desirability and likelihood ratings — not a formal risk assessment. Likelihood rows 1–2 are empty: no median 15-year likelihood below 3.

Legend:

- Each statement is placed by the panel's median likelihood that the trend is realised within fifteen years (vertical, 1 = very unlikely to 5 = very likely) and by how undesirable the trend would be if it occurred, derived from the median desirability rating (horizontal, 1 = very desirable to 5 = very undesirable).
- The score shown in each cell is undesirability × likelihood (1–25), shaded from green (low priority) through yellow and orange to red and dark red (critical priority).
- The matrix is an illustrative policy-priority visualisation of the panel's desirability and likelihood ratings; it is not a formal risk assessment. Undesirable-and-likely developments fall in the red upper-right region, dominated by the inclusion statements.

3.5 Quantitative profile and qualitative engagement, side by side

The OSES is a mixed-methods instrument: every rated statement also invited open comments. This section places the two strands side by side. The table gives each statement's quantitative headline together with the number of open comments it attracted (1,533 in total); the paragraphs that follow distil illustrative themes from the open responses for the most-commented statements (drawn from the English-language responses; the full, coded qualitative material is in the Qualitative Volume).

Table 69. Quantitative profile and qualitative engagement side by side — headline ratings and volume of open comments per statement

Code	Imp Md/IQR	Des Md	Conf Md	L15 Md	Open comments	Statement
S01	4/2.0	4	3	4	84	Double competencies (disciplinary plus IT) will become a universal employment requirement
S02	4/2.0	1	4	3	82	Illiteracy after eight or nine years of schooling will be the norm
S03	4/1.0	2	4	3	63	The COVID-19 crisis will have long-term impacts on learning behaviour and abilities for young people
S04	4/1.0	5	4	4	64	Schools will implement effective programs to address behavioural problems (e.g.: violence, bullying, disengagement)
S05	4/2.0	4	3	3	67	Public investment in early childhood education will become the most critical part of education
S06	4/1.0	4	3	4	59	Education will diversify through vocational pathways that promote specific skills over standardized education
S07	3/1.0	2	3	3	54	Focus on university education will neglect TVET (Technical and Vocational Education and Training)
S08	4/1.0	3	3	3	45	Society's status preference for universities over vocational training will persist
S09	4/1.0	4	4	5	49	New career entry pathways will emerge as AI transforms traditional employment routes
S10	4/1.0	3	4	4	45	Work experience requirements for entry-level positions will become more prevalent
S11	4/1.0	2	4	4	63	Overqualification will increase due to the growing number of academic qualifications
S12	4/2.0	1	4	4	59	Lack of affordable childcare will prevent young parents' labour participation
S13	4/1.0	2	4	3	49	Young people will be first to be fired and last to be hired during downturns
S14	4/1.0	2	4	4	49	The gap between employer expectations for skilled workers and employer participation in training provision will widen
S15	4/1.0	3	4	4	47	Ageing workforce will create massive talent demand in certain sectors
S16	4/1.5	1	4	3	39	Mental health challenges among youth stemming from the COVID-19 crisis will continue to affect life skills
S17	4/1.0	2	4	4	50	Media and Influencers will significantly influence student program choices
S18	4/2.0	2	4	4	39	The transition from education to employment will become more unstable and fragmented for many young people
S19	3/1.0	2	3	4	38	Extended juvenile phase will hinder workplace engagement
S20	4/2.0	4	3	4	54	Schools will increasingly adapt curricula to emerging occupations and career paths
S21	4/2.0	5	3	4	45	Social partnerships between education, the private sector and academia will be widely formed
S22	4/1.0	2	4	4	44	Education and labour market policies will struggle to effectively govern young people's transition
S23	4/1.0	4	3	3	44	Integrated guidance centers for youth employment will be widely established
S24	4/1.0	1	5	4	41	Educational inequalities will exacerbate social polarization and segregation
S25	4/2.0	1	4	4	44	Resource inequities between public and private schools will widen
S26	4/2.0	1	4	4	44	Social class (socio-economic background) will continue to determine educational opportunities
S27	4/2.0	5	3	4	47	Initiatives breaking gender stereotypes will open previously male-dominated career pathways to women
S28	4/2.0	1	4	4	44	School system reliance on parental help will disadvantage non-native and low-education families

Code	Imp Md/IQR	Des Md	Conf Md	L15 Md	Open comments	Statement
S29	4/1.0	4	3	4	47	Digital infrastructure and remote learning will reduce regional disparities between urban and rural education
S30	4/2.0	1	4	4	34	Minority ethnic and racial groups of young people will face intensified discrimination

Illustrative qualitative themes (most-commented statements).

S01 — Double competencies (84 comments). Respondents stress that the requirement is occupation- and branch-specific rather than truly universal, that AI is rapidly redefining what 'IT competence' means, and that treating double competence as universal could exclude or disadvantage some groups.

S02 — Illiteracy as the norm (82 comments). Much of the commentary turns on the definition of 'illiteracy' (functional versus literal) and on unease with the framing; respondents treat rising functional-literacy problems as real and serious but resist the idea that illiteracy will become a 'norm'.

S04 — Schools addressing behavioural problems (64 comments). Respondents doubt that schools alone can or should solve these problems; effectiveness is seen as contingent on resources and trained staff, with root causes located in society and the media beyond the school gate.

S05 — Early-childhood investment (67 comments). There is broad agreement on its importance but frequent objection to the phrase 'the most critical' (one of several priorities), emphasis that quality and resources are decisive, and the point that early-childhood education is not solely a state responsibility.

S24 — Educational inequality (41 comments). Many note that inequality is already severe (so hard to read as a future 'trend'), link it to the unequal valuation of vocational versus academic pathways, and point to marked country differences in mobility.

S30 — Minority discrimination (34 comments). A strong, country-specific signal — most visibly from Finland — of rising racism and discrimination amplified by politics and media, combined with scepticism about the political will to act.

The remaining statements show comparable engagement (38–63 comments each); their coded themes are documented in the Qualitative Volume, and the Wave 2 open 'biggest risk' responses will extend this qualitative layer.

4 Wave 2 — reassessment rationale, design and final results

Wave 2 closed on 10 June 2026, 18:00, after five and a half weeks in the field (first invitations 4 May 2026). Of the 287 Wave 1 panellists invited (those who had provided an email address), 231 started and 207 completed the reassessment — a completion rate of 89.6% of started sessions, with a median completion time of 36.1 minutes. The figures in this section are based on the final Wave 2 analysis dataset ($N = 207$; no post-cut-off submissions, no cases removed in cleaning) and replace the preliminary figures of earlier drafts.

4.1 Why a second wave

Iteration with controlled feedback is the defining feature of a Delphi, and it is what distinguishes the OSES from a one-off survey. The first wave establishes a baseline: for each statement it shows where the panel's centre of gravity lies and how widely opinion is spread. The second wave then returns that information to the respondents. For every carried-forward statement, participants are shown a structured summary of the Wave 1 group result — the distribution of answers, the median, and the IQR, with the level of consensus indicated on a simple low/moderate/high (red/amber/green) scale — and are asked to reassess in the light of the collective view.

This serves two purposes. The first is convergence: Delphi theory expects that, confronted with the aggregate view, some respondents revise toward it, so that dispersion narrows where genuine common ground exists, while durable disagreement persists where it is real. Measuring whether and where this happens is itself a finding. The second purpose is explanation: Wave 2 goes beyond re-rating to ask why experts hold the views they do, which underlying factors they see as driving each trend and which policy responses they consider most needed, thereby beginning to convert the quantitative pattern into actionable substance.

In practice the feedback is presented per statement in a compact, comparable format: a short distribution table showing the percentage of respondents who chose each scale point, the median, and the IQR, with the level of consensus colour-coded from low (red) through moderate (amber) to high (green). Respondents thus see at a glance both where the panel landed and how tightly it agreed before deciding whether to revise their own assessment. This controlled-feedback step is precisely what distinguishes the second wave from a simple repeat of the first.

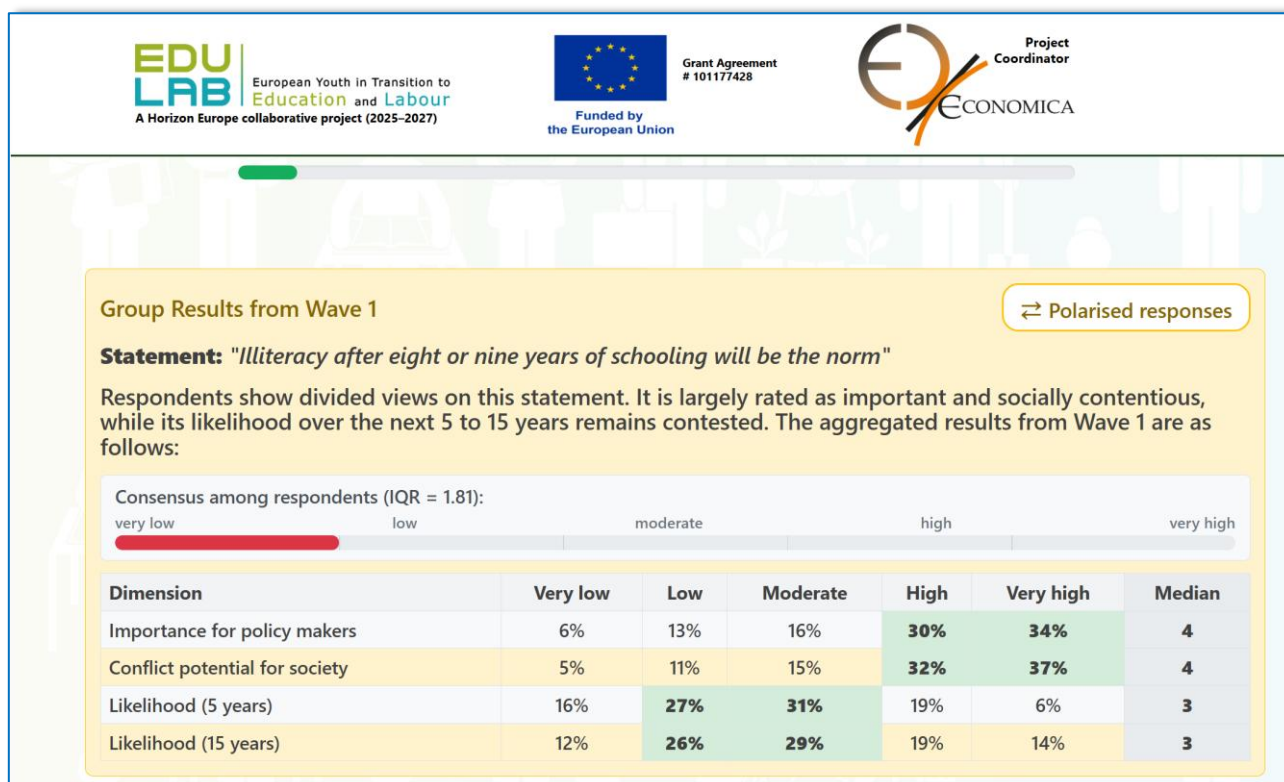
Figure 6 reproduces this per-statement feedback display as it was shown to respondents at the start of the second wave.

4.2 Panel design: re-contact retention and longitudinal matching

Wave 2 differs from a conventional second Delphi round in one structural respect: under the project's GDPR framework, only Wave 1 respondents who explicitly consented to re-contact (by leaving an email address at the end of the Wave 1 questionnaire) could be invited again. The study therefore combines two designs: a cross-sectional Delphi sample in each wave ($N = 433$; $N = 207$), analysed at the aggregate level, and an embedded longitudinal re-contact panel of respondents who consented, received a personal access token and took part again, analysed at the level of individual experts (see the matched-sample analysis at the end of this section and the Quantitative Tables Volume, Section 6).

Three analytically distinct populations follow from this design. First, the Wave 1 sample ($N = 433$). Second, the re-contact consent panel: 281 of the 433 Wave 1 respondents (64.9%) provided an email address. Combining this with eight post-cut-off Wave 1 completers and one manually added participant, and after de-duplication, this yielded 288 personal access tokens, of which 287 panellists were ultimately invited (one duplicate identity was excluded before dispatch). Third, the matched longitudinal panel: the 205 Wave 2 respondents whose answers could be linked to their own Wave 1 record. Describing the step from 433 to 207 as conventional panel attrition would be misleading, because the full Wave 1 sample was never eligible for re-contact. Two indicators describe longitudinal participation more accurately: the re-contact retention rate (completed Wave 2 questionnaires per invited consenting panellist, $207/287 = 72.1\%$) and the longitudinal match rate (Wave 2 responses linkable across waves, $205/207 = 99.0\%$).

Figure 6. Wave 2 feedback display (English version): the Wave 1 group result returned to respondents before re-assessment — here for statement S02 — showing the percentage distribution across the response options, the median, and an inter-quartile-range consensus bar.



Explicit respondent-level matching of this kind remains uncommon in Delphi research, where participation and consensus are usually discussed at the aggregate level (Hasson, Keeney & McKenna, 2000; Keeney, Hasson & McKenna, 2011; von der Gracht, 2012; Diamond et al., 2014). Under increasingly stringent data-protection and research-ethics requirements, the distinction between the survey sample, the consented re-contact panel and the matched longitudinal panel offers a transferable framework for Delphi designs that need to balance respondent privacy with longitudinal analytical capacity; the within-respondent analyses reported at the end of this section are its direct payoff.

4.3 The 21 carried-forward statements and the four blocks

Wave 2 does not repeat all 30 statements. Carrying every statement forward would burden respondents and waste effort on items already settled in Wave 1. Instead, 21 statements are carried forward and grouped into four blocks according to their Wave 1 result pattern, so that the reassessment can be tailored to what each statement most needs. The statement codes are identical to Wave 1, preserving comparability.

Table 70. The four Wave 2 blocks — carried-forward statements and their reassessment focus

Block	n	Statements	Reassessment focus / design
Block 1	2	S02, S07	Controversial — re-rated on all four Wave 1 dimensions
Block 2	8	S01, S05, S13, S16, S17, S21, S23, S29	Diverging expert views — re-rated on importance and likelihood (5 yrs)
Block 3	3	S18, S25, S28	Expert uncertainty — two new items (agreement with the Wave 1 group result; certainty in one's own Wave 1 rating)
Block 4	8	S03, S04, S09, S12, S14, S20, S24, S27	Consensus, optional deepening — one new item (feasibility of effective policy implementation)

The logic is that controversial statements are put fully back to the panel on all four dimensions while statements with diverging views are revisited on the two dimensions where they split most (importance and the five-year likelihood), uncertain statements are probed not for a fresh rating but for the source of their uncertainty (how far respondents agree with the group result and how certain they were of their own Wave

1 judgement) and consensus statements, where re-rating would add little, are deepened instead with a single forward-looking question on the feasibility of effective policy implementation. The block design thus follows directly from the consensus/controversy structure identified in Section 4.4.

In plain terms, the four blocks simply sort the Wave 1 statements by how the panel answered. Where experts broadly agreed (a shared view), the statement sits in the consensus block and is carried forward only to probe how feasible it would be to act on it. Where they leaned the same way but with widely differing intensity, it is a diverging statement. Where answers were spread with no settled direction, it is uncertain. And where answers split into two opposing camps, it is controversial (polarised). Each block is then re-asked in the way that best fits its pattern (Section 5.3), and the exact Wave 1 figures behind each statement's classification are given in Section 4 and the Quantitative Tables Volume. The way these results were fed back to respondents at the start of the second wave (the per-statement feedback display) is reproduced in the Documentation component of this deliverable.

4.4 Block-specific design and additional items

All Wave 2 rating dimensions use the same five-point scale as Wave 1 (very low to very high) plus a “cannot judge” option, keeping the measures comparable. Beyond the block-specific rating, every statement in Wave 2 carries two structured multiple-choice items and one open question, applied uniformly across all 21 statements. The first multiple-choice item asks which types of factors might contribute to the trend, from a fixed taxonomy of seven categories: structural features of population and society; policy frameworks and regulations; market dynamics and economic conditions; technological change and digitalisation; social and cultural shifts; institutional practices and organisational capacities; and the individual behaviour and choices of young people and their families. The second asks which types of policy measures are most needed, again from a fixed taxonomy: regulation and legislation; funding and financial incentives; investment in education, training and skills development; support and counselling services; cross-sector partnerships and stakeholder engagement; research, monitoring and evidence-building; and awareness raising and information campaigns. A final free-text field asks for the single biggest risk if the issue is not addressed. The multiple-choice frequencies are tabulated in the Quantitative Tables Volume; the open “biggest risk” responses are documented in the Qualitative Volume.

The two taxonomies are deliberately generic and identical across all 21 statements. Using the same factor and policy categories everywhere means the profiles can be compared across statements and themes, showing, for example, whether a given trend is seen as driven mainly by structural and economic forces or by individual behaviour, and whether the preferred response leans toward regulation, investment, or partnership. The fixed categories also keep the item short enough to sit at the end of each statement without further lengthening the questionnaire, while the open “biggest risk” field preserves room for considerations the categories do not capture.

Figures 7, 8 and 9 show the reassessment question and the two additional multiple-choice items, asked for every statement: respondents could select up to three contributing factors and up to three policy measures.

Figure 7. Wave 2 questionnaire (English version): The reassessment question, asked for all four dimensions here (S02).

Please reassess the statement "**Illiteracy after eight or nine years of schooling will be the norm**" with regard to:

	Very low	Low	Moderate	High	Very high	Prefer not to say / don't know / cannot judge
Importance for policy makers	☐	☐	☐	☐	☐	☐
Conflict potential for society	☐	☐	☐	☐	☐	☐
Likelihood within 5 years	☐	☐	☐	☐	☐	☐
Likelihood within 15 years	☐	☐	☐	☐	☐	☐

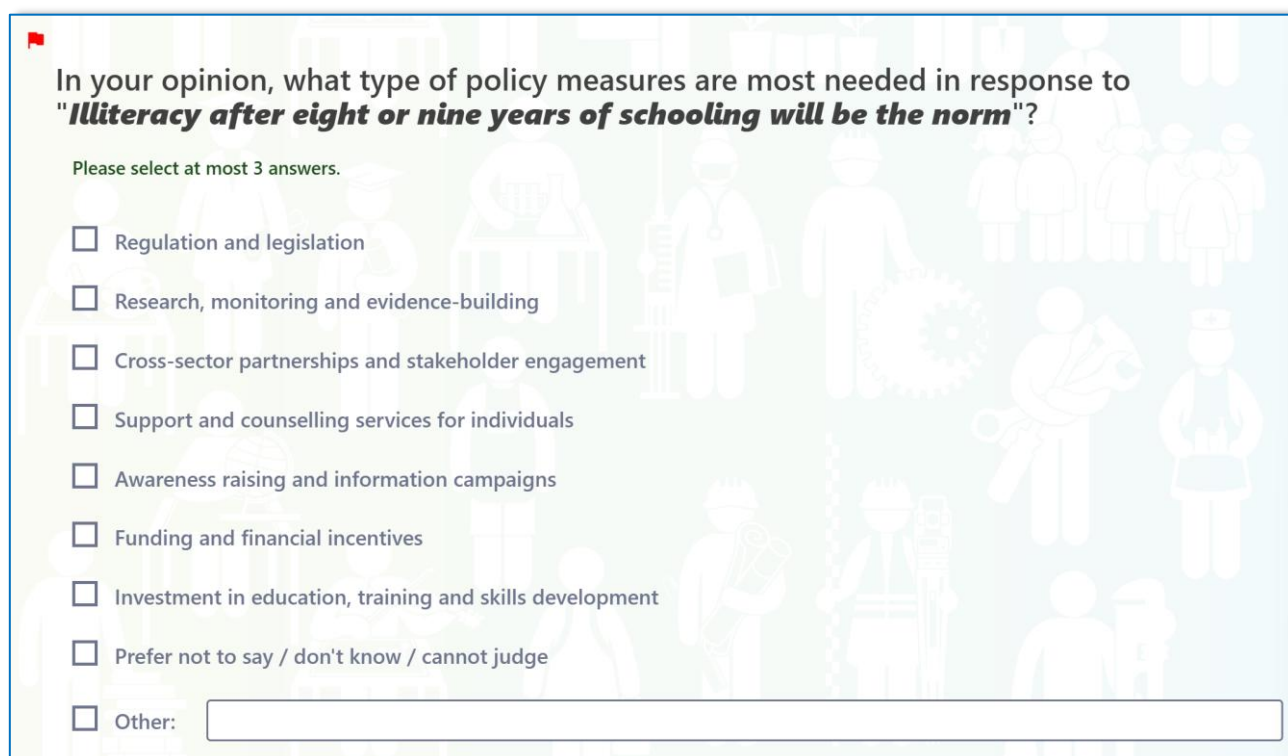
Figure 8. Wave 2 questionnaire (English version): the contributing-factor item (categories F1–F7), asked for every statement (here S02); respondents could select up to three.

In your opinion, what type of factors might contribute to "**Illiteracy after eight or nine years of schooling will be the norm**" becoming reality?

Please select at most 3 answers.

- ☐ Technological change and digitalisation
- ☐ Institutional practices and organisational capacities of education and training institutions
- ☐ Social and cultural shifts
- ☐ Policy frameworks and regulations
- ☐ Market dynamics and economic conditions
- ☐ Structural features of population and society
- ☐ Individual behaviour and choices of young people and their families
- ☐ Prefer not to say / don't know / cannot judge
- ☐ Other:

Figure 9. Wave 2 questionnaire (English version): the policy-measure item (categories P1–P7), asked for every statement (here S02); respondents could select up to three.



 In your opinion, what type of policy measures are most needed in response to **"Illiteracy after eight or nine years of schooling will be the norm"**?

Please select at most 3 answers.

- ☐ Regulation and legislation
- ☐ Research, monitoring and evidence-building
- ☐ Cross-sector partnerships and stakeholder engagement
- ☐ Support and counselling services for individuals
- ☐ Awareness raising and information campaigns
- ☐ Funding and financial incentives
- ☐ Investment in education, training and skills development
- ☐ Prefer not to say / don't know / cannot judge
- ☐ Other:

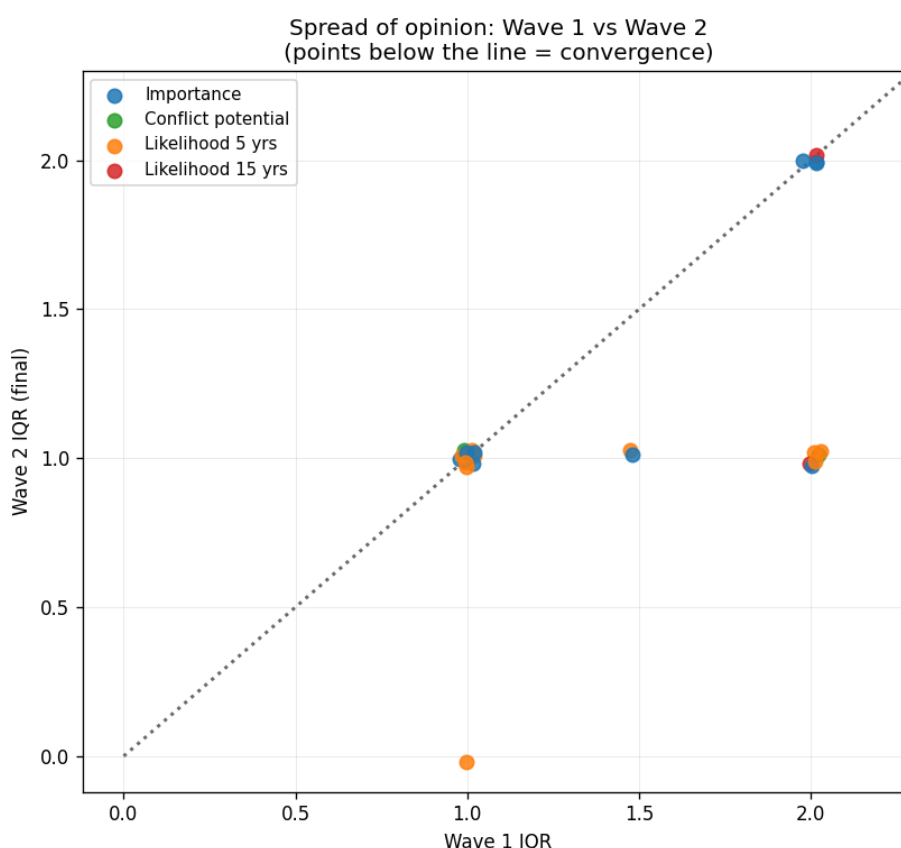
4.5 Final results

The second wave behaves as a Delphi should. Among the 24 statement-by-dimension comparisons that can be made on the dimensions measured in both waves (the four dimensions of the controversial block and the two of the diverging block), the interquartile range narrowed in 9 cases, was unchanged in 15, and widened in none — that is, wherever opinion moved at all, it moved toward consensus rather than away from it. Median ratings are stable in 23 of the 24 comparisons; the single exception — the five-year likelihood of S21 (3.0 → 4.0) — moved upward, re-anchoring rather than unsettling the substantive picture established in Wave 1. The per-statement comparison, including median and IQR side by side with the change between waves, is in the Quantitative Tables Volume (Section 5).

The structured factor and policy items give a first, descriptive sense of the “why” and the “what to do”, applied consistently across the 21 statements. The factors most frequently identified as driving the trends are market dynamics and economic conditions (selected on average by 43% of respondents per statement), social and cultural shifts (42%) and technological change and digitalisation (40%). On the policy side, the measures most often seen as needed are investment in education, training and skills development (65%), funding and financial incentives (37%) and regulation and legislation (36%). These are averages across statements; the per-statement profiles, which vary substantially by topic, are in the Quantitative Tables Volume (Section 4.5).

Blocks 3 and 4 use new Wave 2 judgement dimensions (agreement with the group result and certainty in one's own Wave 1 rating for the uncertain block, and the feasibility of effective policy implementation for the consensus block) which have no Wave 1 counterpart and are therefore reported in Wave 2 terms only; their final distributions are in the Quantitative Tables Volume (Section 4).

Figure 10. Spread of opinion (IQR) in Wave 1 versus final Wave 2 (N = 207), for the dimensions measured in both waves.



Legend: Points below the diagonal indicate convergence; none lie above it. How to read: each point is one statement on one dimension; points on the dotted line are unchanged between waves, while points below it have a narrower Wave 2 inter-quartile range — i.e. convergence. The message is that, where opinion shifted, it moved toward consensus and none diverged. (Statement codes: Table 4.)

Results by block.

Block 1 — controversial statements (S02, S07). Re-rated on all four Wave 1 dimensions. The table contrasts the Wave 1 and final Wave 2 medians and IQRs; the IQR is the key figure, since convergence shows as a narrowing range.

Table 71. Wave 2 — Block 1 (controversial statements): change in median and IQR from Wave 1 to Wave 2 on all four assessment dimensions

Code	Dimension	W1 Md/IQR	W2 Md/IQR	Δ IQR
S02	Importance	4 / 2.0	4 / 2.0	+0.0
S02	Conflict pot.	4 / 2.0	4 / 1.0	-1.0
S02	Likelihood 5y	3 / 1.5	3 / 1.0	-0.5
S02	Likelihood 15y	3 / 2.0	3 / 2.0	+0.0
S07	Importance	3 / 1.0	3 / 1.0	+0.0
S07	Conflict pot.	3 / 1.0	3 / 1.0	+0.0
S07	Likelihood 5y	3 / 2.0	3 / 1.0	-1.0
S07	Likelihood 15y	3 / 2.0	3 / 1.0	-1.0

Block 2 — diverging statements (8 items). Re-rated on importance and the five-year likelihood, the two dimensions on which Wave 1 opinion was most split. Across these items the medians are stable — except for an upward shift in the five-year likelihood of S21 (3.0 → 4.0) — while several IQRs narrow.

Table 72. Wave 2 — Block 2 (diverging statements): change in median and IQR from Wave 1 to Wave 2

Code	Dimension	W1 Md/IQR	W2 Md/IQR	Δ IQR
S01	Importance	4 / 2.0	4 / 1.0	-1.0
S01	Likelihood 5y	4 / 1.0	4 / 1.0	+0.0
S05	Importance	4 / 2.0	4 / 2.0	+0.0
S05	Likelihood 5y	3 / 2.0	3 / 1.0	-1.0
S13	Importance	4 / 1.0	4 / 1.0	+0.0
S13	Likelihood 5y	4 / 1.0	4 / 1.0	+0.0
S16	Importance	4 / 1.5	4 / 1.0	-0.5
S16	Likelihood 5y	4 / 1.0	4 / 1.0	+0.0
S17	Importance	4 / 1.0	4 / 1.0	+0.0
S17	Likelihood 5y	4 / 1.0	4 / 0.0	-1.0
S21	Importance	4 / 2.0	4 / 2.0	+0.0
S21	Likelihood 5y	3 / 1.0	4 / 1.0	+0.0
S23	Importance	4 / 1.0	4 / 1.0	+0.0
S23	Likelihood 5y	3 / 2.0	3 / 1.0	-1.0
S29	Importance	4 / 1.0	4 / 1.0	+0.0
S29	Likelihood 5y	3 / 1.0	3 / 1.0	+0.0

Block 3 — uncertain statements (S18, S25, S28). Rather than re-rating, Wave 2 asks two reflective questions: how far respondents agree with the Wave 1 group result, and how certain they were of their own Wave 1 judgement (both on the 1–5 scale). The distributions are summarised below.

Table 73. Wave 2 — Block 3 (uncertain statements): agreement with the Wave 1 group result and respondents' certainty

Code	Statement	Agreement Md/IQR	Certainty Md/IQR
S18	The transition from education to employment will become more unstable and fragmented for many young people	4 / 2.0	4 / 1.0
S25	Resource inequities between public and private schools will widen	4 / 1.0	4 / 1.0
S28	School system reliance on parental help will disadvantage non-native and low-education families	4 / 1.0	4 / 1.0

Block 4 — consensus statements (8 items). Not re-rated; deepened instead with a single question on the feasibility of effective policy implementation (1–5). The medians indicate how realistic experts consider effective policy action on each consensus trend.

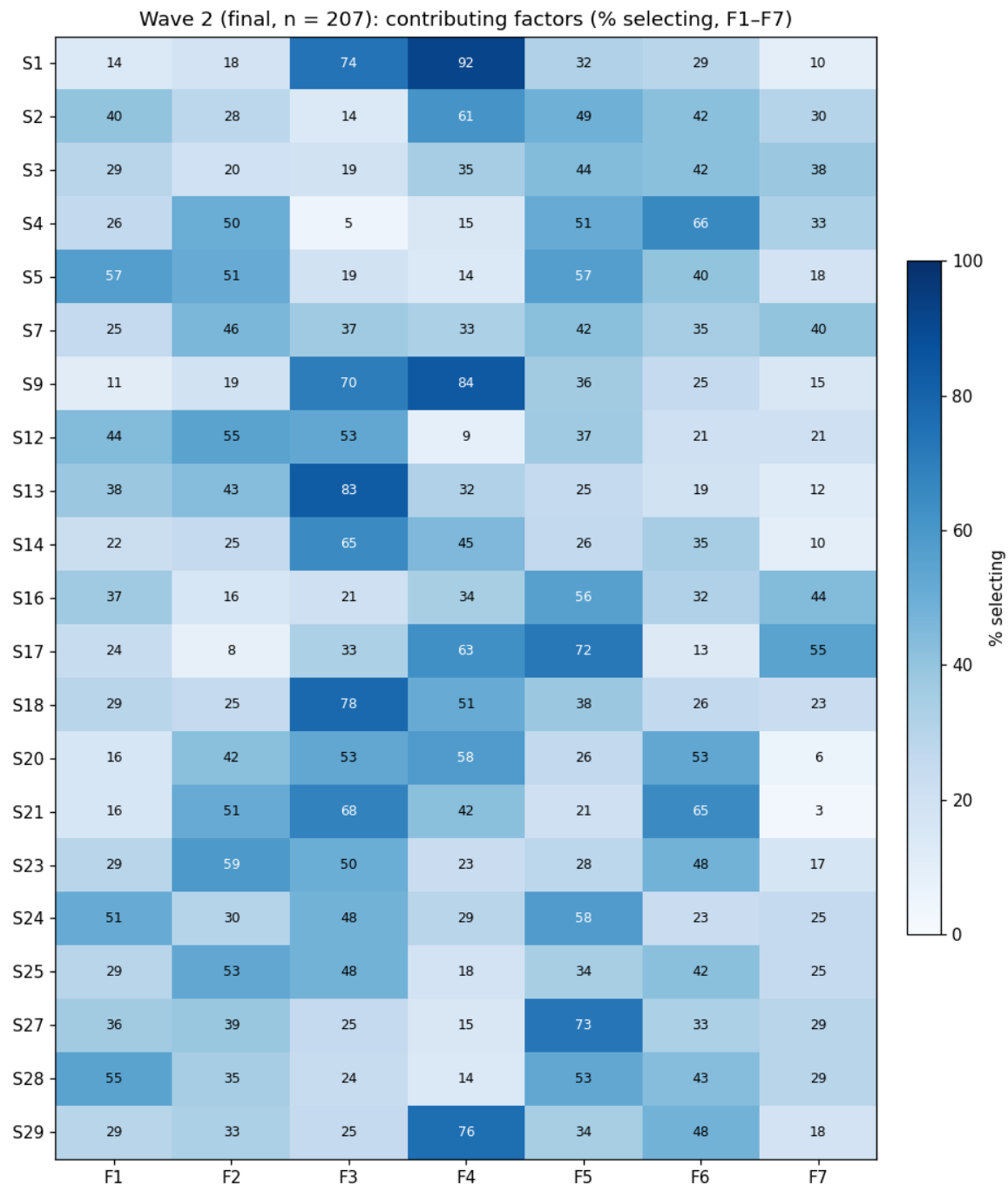
Table 74. Wave 2 — Block 4 (consensus statements): perceived feasibility of policy action

Code	Statement	Feasibility for policy actions (Md/IQR)
S03	The COVID-19 crisis will have long-term impacts on learning behaviour and abilities for young people	4 / 1.0
S04	Schools will implement effective programs to address behavioural problems (e.g.: violence, bullying, disengagement)	4 / 1.0
S09	New career entry pathways will emerge as AI transforms traditional employment routes	4 / 1.0
S12	Lack of affordable childcare will prevent young parents' labour participation	4 / 2.0
S14	The gap between employer expectations for skilled workers and employer participation in training provision will widen	4 / 1.0
S20	Schools will increasingly adapt curricula to emerging occupations and career paths	4 / 1.0
S24	Educational inequalities will exacerbate social polarization and segregation	4 / 2.0
S27	Initiatives breaking gender stereotypes will open previously male-dominated career pathways to women	4 / 1.0

A note on linkage

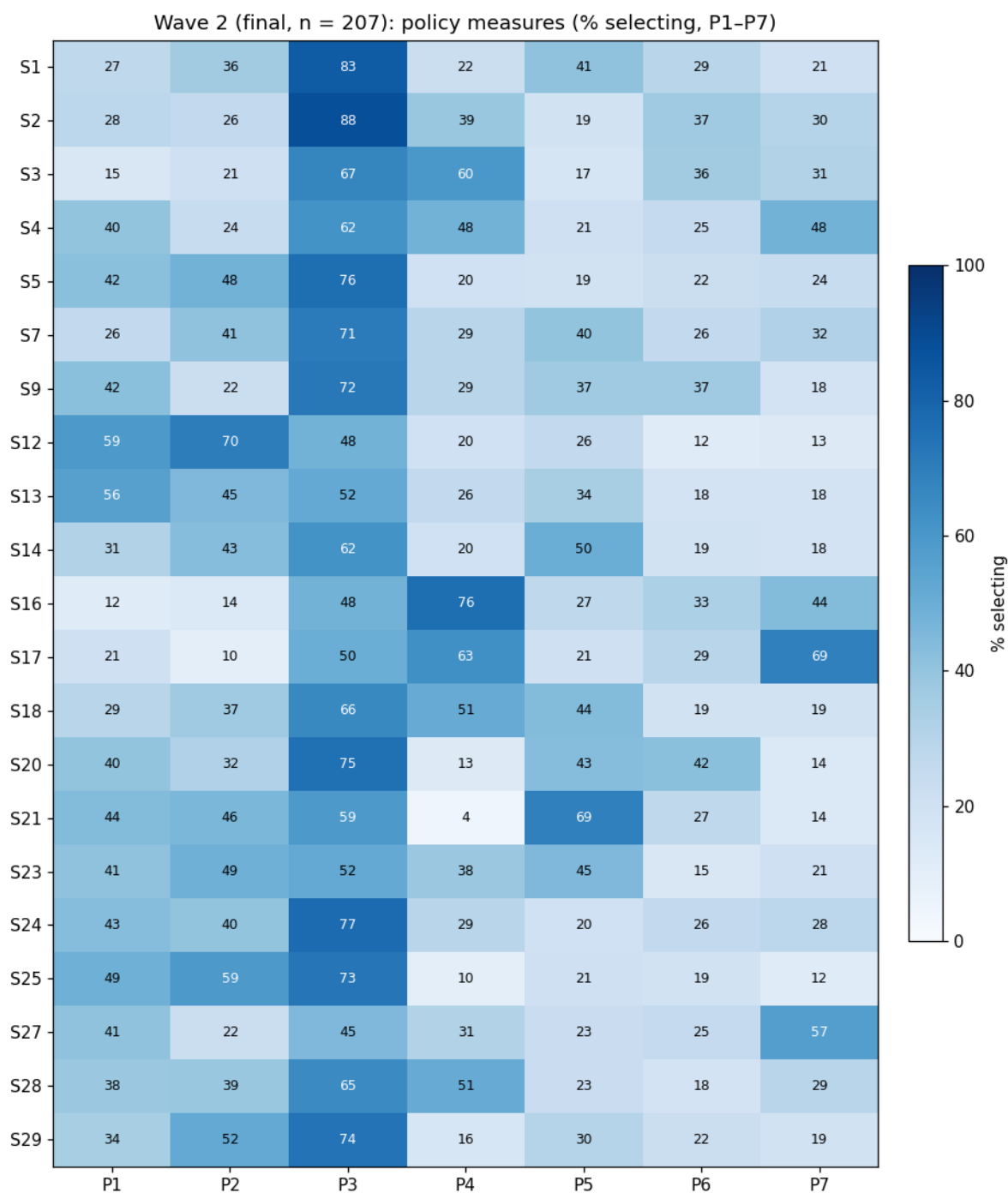
Wave 2 was a token-based reassessment that invited the Wave 1 respondents who had provided an email address, so country is taken from the matched Wave 1 record; 205 of the 207 completed questionnaires were linked to a Wave 1 record in the analysis dataset. The cross-wave special cases are documented in the Quantitative Tables Volume (Section 4.0): one respondent who could not be matched to a specific Wave 1 record (no identifying information was given in Wave 1), one post-cut-off Wave 1 completer who took part in Wave 2, and one email address attached to two duplicate Wave 1 records (relevant only for the within-person analyses of the matched sample). A further pair of Wave 1 records was identified post hoc as one respondent registered under two email addresses (only one identity was invited to Wave 2); a sensitivity check confirms that excluding the duplicate would leave all reported Wave 1 medians and consensus readings unchanged, so the Wave 1 dataset remains as published (N = 433).

Figure 11. Wave 2 (final, N = 207): share of respondents selecting each contributing-factor category (F1–F7), by statement.



How to read: the darker the cell, the larger the share of respondents who selected that contributing factor (F1–F7) for the statement in that row; reading across a row gives the statement's perceived 'driver' profile. The factor categories F1–F7 are defined in the legend above and the statements are listed in Table 4.

Figure 12. Wave 2 (final, N = 207): share of respondents selecting each policy-measure category (P1–P7), by statement.



How to read: the darker the cell, the larger the share of respondents who selected that policy measure (P1–P7) for the statement in that row; reading across a row gives the statement's preferred policy-response profile. The policy categories P1–P7 are defined in the legend above and the statements are listed in Table 4.

4.6 What the second round changed — a convergence typology

With the final dataset in place, the 24 comparable statement × dimension pairs can be condensed into a compact account of what the second round changed (Table 74; the full per-dimension figures are in the Quantitative Tables Volume, Sections 5.1–5.2). The headline is convergence without reversal: seven of the ten re-rated statements narrowed their interquartile range on at least one dimension, the remaining three were already at, and stayed at, their Wave 1 spread, and no statement widened on any dimension. The median moved in only one of the 24 pairs, the five-year likelihood of S21 (3.0 → 4.0), upward, so the substantive picture of Wave 1 holds while opinion consolidates around it.

Table 75. Wave 2 — what the second round changed: convergence typology of the ten re-rated statements (Blocks 1 and 2). “IQR narrowed” counts the re-rated dimensions with a smaller Wave 2 inter-quartile range; no dimension widened.

Code	Block	IQR narrowed	Median change	Pattern after Wave 2
S02	1 — controversial	2 of 4	—	Converged (conflict potential, 5-yr likelihood); importance and 15-yr likelihood remain at IQR 2.0
S07	1 — controversial	2 of 4	—	Converged; all four dimensions now at IQR ≤ 1
S01	2 — diverging	1 of 2	—	Converged; both dimensions at IQR ≤ 1
S05	2 — diverging	1 of 2	—	Converged on likelihood; importance remains at IQR 2.0
S13	2 — diverging	0 of 2	—	Stable; both dimensions at IQR ≤ 1
S16	2 — diverging	1 of 2	—	Converged; both dimensions at IQR ≤ 1
S17	2 — diverging	1 of 2	—	Converged; 5-yr likelihood now at IQR 0
S21	2 — diverging	0 of 2	Likelihood 5 yrs: 3.0 → 4.0	Spread stable but re-anchored upward; importance remains at IQR 2.0
S23	2 — diverging	1 of 2	—	Converged; both dimensions at IQR ≤ 1
S29	2 — diverging	0 of 2	—	Stable; both dimensions at IQR ≤ 1

After the second round, the residual dispersion is concentrated and informative. Seven of the ten re-rated statements now show consensus-level spread (IQR ≤ 1, the convention used throughout this report) on every re-rated dimension. Where dispersion persists, it sits almost entirely on a single dimension — importance for policymakers — for S02, S05 and S21 (IQR 2.0 in both waves), joined only by the fifteen-year likelihood of S02: experts increasingly agree on what they expect to happen, but continue to differ on how high these items should rank on the policy agenda (and, for S02, on how the picture plays out over the longer horizon). This is the kind of residual disagreement a policy Delphi is meant to surface rather than smooth over: it marks genuine prioritisation choices, not unsettled expectations.

4.7 Per-country convergence

Read per project country, the second round moves in the same direction everywhere (Table 75): the mean change in the country-level IQR across the 24 comparable pairs is negative in all nine project countries: strongest in Germany (−0.56), Greece (−0.55) and Portugal (−0.45), weakest in Poland (−0.10), the only country in which narrowing and widening pairs balance (8 vs 8). National panels are small in Wave 2 (n between 8 and 35), so these are tendencies rather than precise estimates, the per-country medians and IQRs behind them are in the Quantitative Tables Volume (Sections 4 and 5.3).

Table 76. Wave 2 — per-country convergence across the 24 comparable statement × dimension pairs (change in the country-level IQR from Wave 1 to Wave 2; pairs neither narrowed nor widened are unchanged; a negative mean Δ IQR indicates net convergence)

Country	Wave 2 n	IQR narrowed (of 24)	IQR widened (of 24)	Mean Δ IQR
Finland	33	6	2	-0.23
Austria	35	10	2	-0.27
Kosovo	26	6	3	-0.21
Italy	33	7	4	-0.21
Greece	8	16	3	-0.55

Country	Wave 2 n	IQR narrowed (of 24)	IQR widened (of 24)	Mean Δ IQR
Portugal	17	12	2	-0.45
the UK	8	12	5	-0.39
Poland	12	8	8	-0.10
Germany	17	14	2	-0.56

4.8 Driver and policy profiles by theme

Read by theme, the driver profiles are instructively different (Figure 11). The labour-market statements (Theme B) are attributed above all to market dynamics and economic conditions (F3; selected on average by 68% of respondents per statement). The education statements (Theme A) draw a more even profile led by social and cultural shifts (46%), institutional practices (42%) and technological change (42%); the transition statements (Theme C) combine market dynamics (51%) with technological change (45%); and the inclusion statements (Theme D) are read as socially and structurally driven (social and cultural shifts 51%, structural features of population and society 40%). The drivers' experts see thus track the substance of each theme rather than a generalised technology or policy narrative.

The policy-need profiles (Figure 12) are uniform on one point and differentiated on another. Investment in education, training and skills development (P3) is the most-selected category in every theme, from 58% (Theme B) to 75% (Theme A), confirming investment as the panel's default instrument. Around this core, the instrument mix shifts with the problem: regulation (P1, 47%) and funding (P2, 45%) weigh most in the labour-market theme, cross-sector partnerships (P5, 42%) and support and counselling services (P4, 41%) in the transition theme, and funding and regulation again in the inclusion theme (42% / 41%). The full per-statement profiles are tabulated in the Quantitative Tables Volume (Section 4.5).

4.9 Feasibility, agreement and certainty: what Blocks 3 and 4 add

Blocks 3 and 4 carry judgement dimensions that have no Wave 1 counterpart and are reported in Wave 2 terms only (Tables 72 and 73). For the eight consensus statements, the panel considers effective policy implementation feasible: the median feasibility rating is 4 ("high") for every statement, with consensus-level spread (IQR 1.0) for six of the eight and a wider spread (IQR 2.0) only for S12 (affordable childcare) and S24 (educational inequalities). These are the two consensus trends where the room for effective policy action is most contested.

For the three uncertain statements, respondents broadly endorse the Wave 1 group result (median agreement 4 throughout; IQR 1.0 for S25 and S28, 2.0 for S18) while reporting high certainty in their own Wave 1 judgement (median 4, IQR 1.0 in all three cases). The Wave 1 uncertainty pattern therefore reflected genuinely scattered expectations across experts rather than low individual confidence, which is useful to know for the concluding round, where these statements are natural candidates for qualitative probing.

4.10 How individual experts moved: the matched sample

Because Wave 2 was token-based, responses can be followed at the level of individual experts: 205 of the 207 Wave 2 completions are linked to their own Wave 1 record (one further case, matching two duplicate Wave 1 submissions, is excluded from the within-person figures, which rest on 204 linked respondents and 4,419 comparable rating pairs; details in the Quantitative Tables Volume, Section 6). This is the analysis the aggregate comparison cannot deliver: it shows who moves, not just whether the distribution tightens.

Three findings stand out (Table 76). First, individual revision is the norm, not the exception: every linked respondent changed at least one rating, and the median respondent changed 59% of their comparable ratings, with the aggregate stability of the medians being the net result of widespread, largely offsetting individual movement. Second, the movement is directed: 40.6% of rating pairs are unchanged, 30.6% moved toward the Wave 1 group median and 22.4% away from it (6.4% switched to an equally distant value on the opposite side), a clear net pull toward the panel view. Third, where movement away from the group median dominates, most visibly the five-year likelihood of S05 (67 vs 51 moves, mean change +0.39), it is consistently upward, mirroring the shift that did move the group median for S21 (3.0 → 4.0): the panel leans toward higher likelihoods rather than dispersing.

Table 77. Wave 2 — matched-sample movement by statement: linked respondents' rating pairs classified by distance to the Wave 1 group median. Unchanged, toward and away do not sum to 100%; the remainder switched to an equally distant value on the opposite side of the median.

Code	Rating pairs (n)	Unchanged (%)	Toward group median (%)	Away from median (%)
S02	727	40.0	33.1	20.2
S07	748	40.6	29.4	23.4
S01	382	41.9	32.5	17.0
S05	387	35.1	29.5	26.9
S13	369	41.7	32.2	22.2
S16	353	46.5	27.8	20.7
S17	355	47.6	27.9	20.3
S21	362	40.1	29.0	24.9
S23	355	38.6	28.2	27.0
S29	381	35.2	34.4	23.1

The matched sample thus completes the convergence story: the second round produced genuine individual reassessment that consolidated, rather than merely repeated, the first-round picture. The remaining Wave 2 material, the open “biggest risk” field and its integration with the quantitative shifts, is documented in the Qualitative Volume, and the distilled consensus-and-controversy dashboard for the WP6 policy briefs builds directly on the typology of Table 74.

5 Connections to other research streams in EDU-LAB

The OSES is one strand of a larger evidence-gathering effort. Two companion deliverables completed in December 2025 approach the same youth-transition agenda from complementary angles: D1.1 (WP1, “Modelling pathways and transitions in GE and PVET and to the labour market”) maps educational pathways, transitions and their intersectional determinants and the gaps in the data that cover them; and D2.1 (WP2, “Assessing the efficacy of policies and investment in GE and PVET”) reviews the effectiveness and efficiency of policies, programmes and investments, including their unintended effects. This section sketches, descriptively, how the OSES findings connect to them; the systematic integration belongs to the project's synthesis stage (WP5).

Connection to WP1 (D1.1, pathways and determinants): WP1 catalogues the pathways young people take through education and training and into work, together with a structured set of some 80 intersectional determinants and 32 transitions and assesses how well existing datasets cover them, documenting where the data are thin. The OSES addresses the same objects from the opposite direction: where WP1 records what the data do and do not capture, the OSES records what a large expert panel expects and how strongly it agrees. The correspondence is closest in two themes. Theme C (Transitions and Systemic Interfaces) concerns the same transitions WP1 catalogues: the school-to-work passage, its stability and the institutional interfaces that govern it. Theme D (Inclusion and Disadvantage) concerns the same intersectional determinants: social class (S26), migration and parental background (S28), gender (S27) and ethnicity and discrimination (S30). Strikingly, the OSES finding that these inequality determinants are judged highly likely yet highly undesirable aligns with WP1's concern that precisely these determinants are often under-covered in available data: the developments experts worry about most are also among the hardest to monitor. The OSES can therefore serve as an expert-judgement complement wherever WP1 identifies a data gap.

Connection to WP2 (D2.1, policy effectiveness and investment): WP2 reviews what is known about the effectiveness and efficiency of policies and investments supporting youth transitions, the structural gaps in that evidence, and unintended effects. The OSES connects to this most directly through its second wave, which asks, for every statement, which types of policy measures experts consider most needed using categories (regulation and legislation; funding and financial incentives; investment in education, training and skills; support and counselling services; cross-sector partnerships; research and monitoring; awareness raising) that mirror WP2's policy landscape and, for the consensus statements, how feasible effective policy implementation is judged to be. The importance and conflict-potential ratings flag where experts see the greatest policy priority and the greatest social sensitivity, and the desirability-likelihood gap flags where policy attention is most warranted. Concrete points of contact include WP2's evidence on the effectiveness of strong vocational pathways, which the OSES complements with expert judgements on vocational diversification (S06) and the persistent, and, to experts, undesirable, status preference for universities over vocational training (S07, S08) as well as WP2's interest in unintended effects, which the open 'biggest risk' question in Wave 2 directly elicits.

The table below summarises the most evident points of contact.

Table 78. Thematic connections between the OSES and WP1 (D1.1) and WP2 (D2.1)

OSES element	Connection to WP1 (D1.1, pathways & determinants)	Connection to WP2 (D2.1, policy & investment)
Theme C — Transitions	Same transitions WP1 catalogues (school-to-work passage, interfaces)	Effectiveness of transition-support policies and guidance
Theme D — Inclusion	Same intersectional determinants (class, migration, gender, ethnicity)	Equity-oriented policies; unintended effects on disadvantaged groups
Wave 2 policy-measure item	—	Maps directly onto WP2 policy categories and effectiveness focus
Wave 2 feasibility item	—	Expert read on whether effective policy action is realistic
Importance & conflict ratings	Which determinants/transitions matter most to experts	Policy priorities and politically sensitive areas

OSES element	Connection to WP1 (D1.1, pathways & determinants)	Connection to WP2 (D2.1, policy & investment)
Wave 2 'biggest risk' (open)	Risks linked to under-covered determinants	Evidence on unintended effects / non-achievement

The statement-level crosswalk below makes the connections concrete: for each statement it names the most directly related WP1 determinant or transition and the most directly related WP2 policy area. The mapping is indicative and intended as a starting point for the WP5 synthesis.

Table 79. Statement-level crosswalk — each statement linked to the relevant WP1 (D1.1) and WP2 (D2.1) evidence

Code	Statement	WP1 (D1.1) link	WP2 (D2.1) link
S01	Double competencies (disciplinary plus IT) will become a universal employment requirement	Skills/qualification determinant; education→labour	Curriculum and digital-skills policy
S02	Illiteracy after eight or nine years of schooling will be the norm	Basic-skills / prior-attainment determinant	Basic-skills and remedial-education policy
S03	The COVID-19 crisis will have long-term impacts on learning behaviour and abilities for young people	Participation determinant (learning disruption)	Learning-recovery and support policy
S04	Schools will implement effective programs to address behavioural problems (e.g.: violence, bullying, disengagement)	Participation / well-being determinant	School support services and resourcing
S05	Public investment in early childhood education will become the most critical part of education	Access determinant (early stage)	Investment in early-childhood education and care
S06	Education will diversify through vocational pathways that promote specific skills over standardized education	Pathway structure / selection	VET/TVET diversification policy
S07	Focus on university education will neglect TVET (Technical and Vocational Education and Training)	Selection (academic vs. vocational)	VET parity and attractiveness policy
S08	Society's status preference for universities over vocational training will persist	Selection determinant (social valuation)	VET attractiveness policy
S09	New career entry pathways will emerge as AI transforms traditional employment routes	Education→labour transition	Labour-market and skills policy (AI/automation)
S10	Work experience requirements for entry-level positions will become more prevalent	Education→labour transition (entry)	Work-based-learning policy
S11	Overqualification will increase due to the growing number of academic qualifications	Education→labour transition (matching)	Skills-matching / qualification policy
S12	Lack of affordable childcare will prevent young parents' labor participation	Family / socio-economic determinant; participation	Childcare and family policy
S13	Young people will be first to be fired and last to be hired during downturns	Education→labour transition (cyclical)	Youth labour-market activation
S14	The gap between employer expectations for skilled workers and employer participation in training provision will widen	Education→labour (employer side)	Employer engagement / WBL policy
S15	Ageing workforce will create massive talent demand in certain sectors	Labour-market / demographic context	Skills, migration and labour policy
S16	Mental health challenges among youth stemming from the COVID-19 crisis will continue to affect life skills	Well-being / participation determinant	Youth mental-health support policy
S17	Media and Influencers will significantly influence student program choices	Selection determinant (information/aspiration)	Guidance and information policy
S18	The transition from education to employment will become more unstable and fragmented for many young people	Education→labour transition (core)	Transition-support policy
S19	Extended juvenile phase will hinder workplace engagement	Transition timing / engagement	Activation and engagement policy
S20	Schools will increasingly adapt curricula to emerging occupations and career paths	Within-education transition; selection	Curriculum–labour alignment policy
S21	Social partnerships between education, the private sector and academia will be widely formed	Systemic interface	Partnership and governance policy

Code	Statement	WP1 (D1.1) link	WP2 (D2.1) link
S22	Education and labour market policies will struggle to effectively govern young people's transition	Systemic interface / governance	Governance and coordination (core WP2)
S23	Integrated guidance centers for youth employment will be widely established	Transition-support structure	Guidance and counselling policy
S24	Educational inequalities will exacerbate social polarization and segregation	Socio-economic determinant (core)	Equity policy
S25	Resource inequities between public and private schools will widen	Access / resource determinant	Funding-equity policy
S26	Social class (socio-economic background) will continue to determine educational opportunities	Socio-economic determinant (core)	Equity and redistribution policy
S27	Initiatives breaking gender stereotypes will open previously male-dominated career pathways to women	Gender determinant	Gender-equality policy
S28	School system reliance on parental help will disadvantage non-native and low-education families	Family / migration determinant	Inclusion and family-support policy
S29	Digital infrastructure and remote learning will reduce regional disparities between urban and rural education	Regional / urban-rural determinant; access	Digital-infrastructure policy
S30	Minority ethnic and racial groups of young people will face intensified discrimination	Ethnicity / migration determinant	Anti-discrimination and inclusion policy

Taken together, the three work packages triangulate the same agenda: WP1 on what the data can and cannot show, WP2 on what the evidence says about policy effectiveness, and WP3B (OSES) on what experts expect, prioritise and consider feasible. Joining these strands systematically is a task for WP5 as this report simply flags the clearest points of contact so that they are not lost.

6 Outlook

Three steps follow from the results reported here. First, with the final Wave 2 dataset now integrated, the remaining analytical step on the second round is interpretive: integrating the quantitative shifts with the open “biggest risk” material documented in the Qualitative Volume, and distilling the headline consensus-and-controversy dashboard for the policy briefs of WP6. Second, a concluding round of 30–40 qualitative interviews will deepen the interpretation of the quantitative findings: it will probe the causes of persistent dissent and unexpected consensus, the implicit assumptions behind divergent assessments, the practical implications for national and EU policy contexts, and experts’ own assessments of perceived policy effectiveness. Third, the consolidated OSES findings feed the project’s synthesis (WP5) and its dissemination through policy briefs, recommendations and publications (WP6), and inform the work of WP1 and WP2.

Even at this descriptive stage, the OSES yields a clear and policy-relevant message: experts across the consortium agree that the developments they find least desirable (those concentrated in the inclusion and disadvantage domain) are among the most likely to materialise and the most socially divisive. Translating that descriptive signal into explanation and recommendation is the task of the project's later stages, for which this report and its companion volumes provide the empirical foundation.

References

- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
- Dalkey, N. C., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science*, 9(3), 458–467. <https://doi.org/10.1287/mnsc.9.3.458>
- Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., & Wales, P. W. (2014). Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology*, 67(4), 401–409. <https://doi.org/10.1016/j.jclinepi.2013.12.002>
- Galesic, M., & Bošnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, 73(2), 349–360. <https://doi.org/10.1093/poq/nfp031>
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4), 1008–1015. <https://doi.org/10.1046/j.1365-2648.2000.01567.x>
- Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation*, 12(10), 1–8. <https://doi.org/10.7275/pdz9-th90>
- Keeney, S., Hasson, F., & McKenna, H. (2011). *The Delphi technique in nursing and health research*. Wiley-Blackwell.
- Kuckartz, U., & Rädiker, S. (2023). *Qualitative content analysis: Methods, practice and software* (2nd ed.). SAGE Publications.
- Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological Forecasting and Social Change*, 73(5), 467–482. <https://doi.org/10.1016/j.techfore.2005.09.002>
- Linstone, H. A., & Turoff, M. (Eds.). (1975). *The Delphi method: Techniques and applications*. Addison-Wesley.
- Mayring, P. (2022). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (13th ed.). Beltz.
- Meskel, P., Murphy, K., Shaw, D., & Casey, D. (2023). The Delphi method in nursing and health research. *International Journal of Nursing Studies Advances*, 5, 100123. <https://doi.org/10.1016/j.ijnsa.2023.100123>
- Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. *Information & Management*, 42(1), 15–29. <https://doi.org/10.1016/j.im.2003.11.002>
- Powell, C. (2003). The Delphi technique: Myths and realities. *Journal of Advanced Nursing*, 41(4), 376–382. <https://doi.org/10.1046/j.1365-2648.2003.02537.x>
- Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: Issues and analysis. *International Journal of Forecasting*, 15(4), 353–375. [https://doi.org/10.1016/S0169-2070\(99\)00018-7](https://doi.org/10.1016/S0169-2070(99)00018-7)
- Schreier, M. (2012). *Qualitative content analysis in practice*. SAGE Publications.
- Schoemaker, P. J. H., Day, G. S., & Snyder, S. A. (2013). Integrating organizational networks, weak signals, strategic radars and scenario planning. *Technological Forecasting and Social Change*, 80(4), 815–824. <https://doi.org/10.1016/j.techfore.2012.10.020>
- Turoff, M. (1970). The design of a policy Delphi. *Technological Forecasting and Social Change*, 2(2), 149–171. [https://doi.org/10.1016/0040-1625\(70\)90061-9](https://doi.org/10.1016/0040-1625(70)90061-9)
- UNESCO Institute for Statistics. (2012). *International Standard Classification of Education: ISCED 2011*. UNESCO Institute for Statistics.
- UNESCO Institute for Statistics. (2014). *ISCED fields of education and training 2013 (ISCED-F 2013): Manual to accompany the International Standard Classification of Education 2011*. UNESCO Institute for Statistics.
- van der Steen, M., & van Twist, M. (2013). Foresight and long-term policy-making: An analysis of anticipatory governance in practice. *Futures*, 54, 33–42. <https://doi.org/10.1016/j.futures.2013.09.005>
- von der Gracht, H. A. (2012). Consensus measurement in Delphi studies: Review and implications for future quality assurance. *Technological Forecasting and Social Change*, 79(8), 1525–1536. <https://doi.org/10.1016/j.techfore.2012.04.013>
- Weaver, W. T. (1971). The Delphi forecasting method. *Phi Delta Kappan*, 52(5), 267–271.
- Weimer, D. L., Vining, A. R., & Thomas, R. K. (2022). *Policy analysis: Concepts and practice* (7th ed.). Routledge.

ANNEX

Annex figures: distribution overviews by assessment dimension

The following charts rank the 30 statements by their mean rating on a single dimension. They complement the medians used in the main text (Section 4) and are provided here for reference.

Figure A1. Statements ranked by mean importance for policymakers, coloured by theme. (Statement codes S01–S30 and wording are given in Table 4; bar colours denote the four themes A–D.)

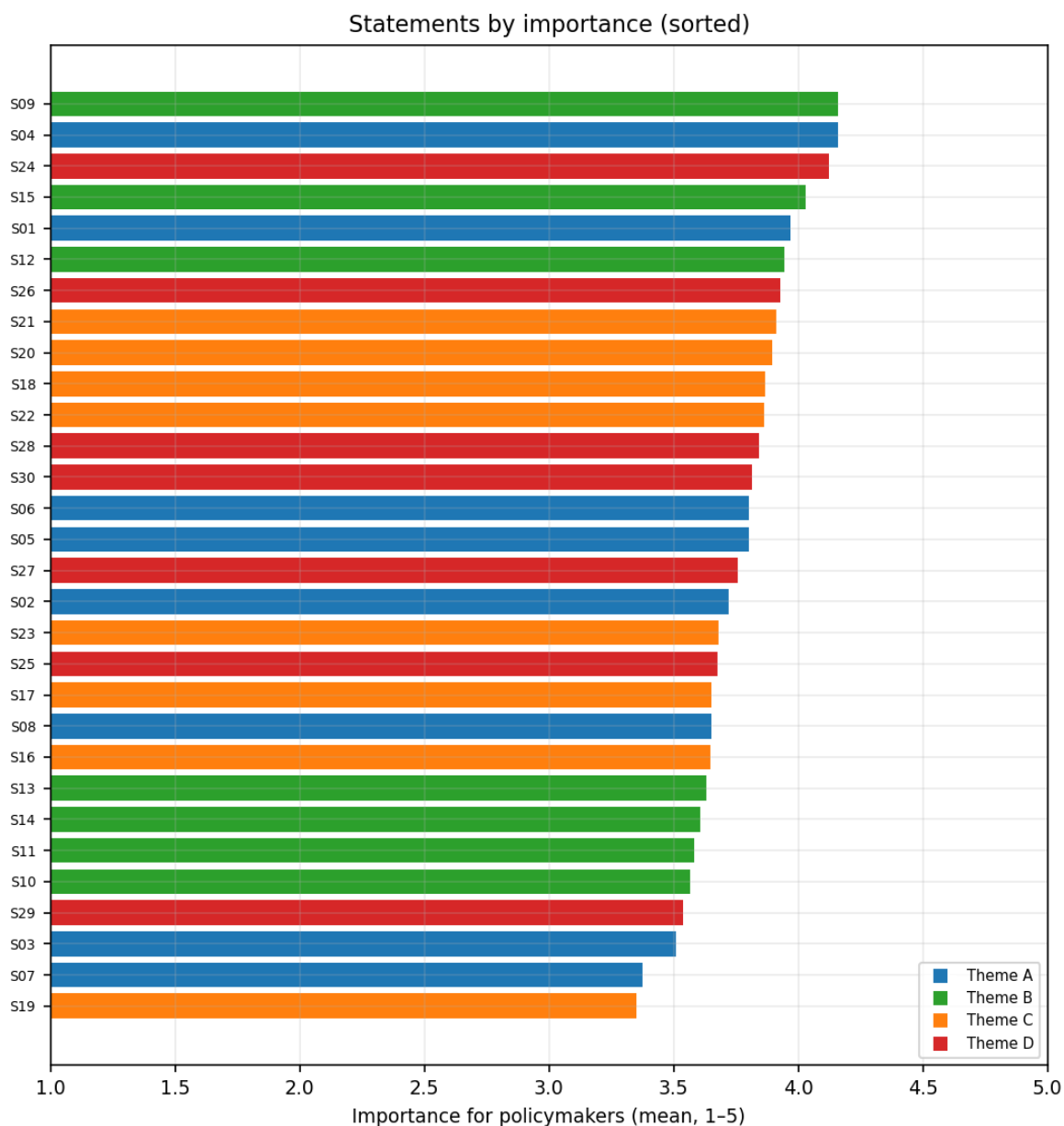


Figure A2. Statements ranked by mean conflict potential for society. The inclusion statements (red, Theme D) dominate the top. (Statement codes S01–S30 and wording are given in Table 4; bar colours denote the four themes A–D.)

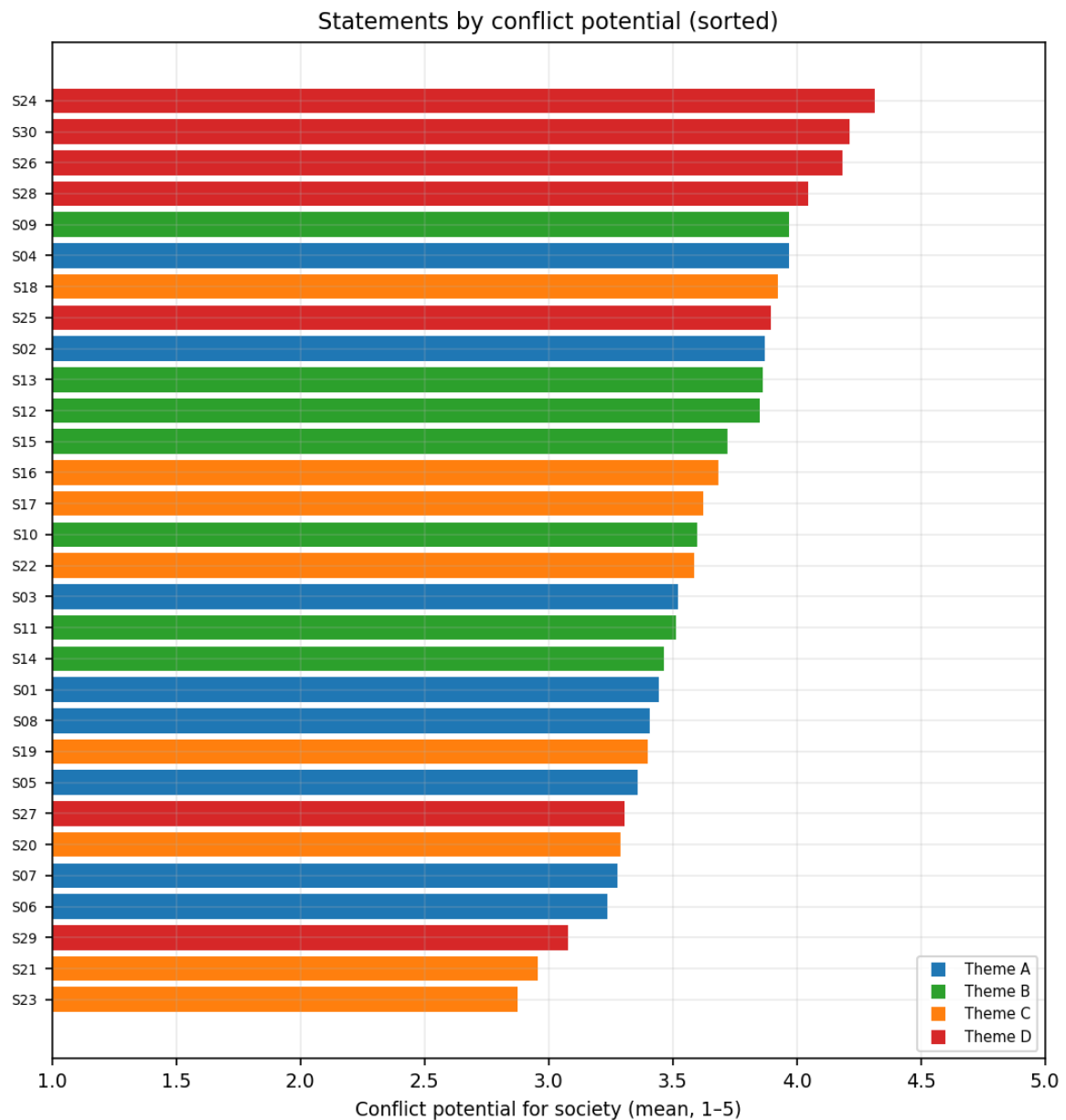


Figure A2. Statements ranked by mean conflict potential for society. The inclusion statements (red, Theme D) dominate the top. (Statement codes S01–S30 and wording are given in Table 4; bar colours denote the four themes A–D.)