Organisation for Economic Co-operation and Development

DCD(2026)18

Unclassified

English text only

25 September 2026

DEVELOPMENT CO-OPERATION DIRECTORATE

Fit for Purpose?

2026 Survey on the OECD DAC Guiding Principles on Managing for Sustainable Development Results

This report presents the results of the 2026 survey on the OECD DAC Guiding Principles on Managing for Sustainable Development Results. It covers 52 bilateral and multilateral development co-operation institutions and compares their results systems with the 2020 baseline. Providers report progress on all six principles, and results frameworks now reach more levels of their institutions. However, the organisational culture, skills, incentives and resources needed to use those frameworks have not kept pace. Results information is still used more for accountability than for learning. This report proposes a typology that places institutions in four positions, according to their capability and the balance between the formal and tacit conditions for results management. It also examines two developments since 2020. First, the indicator framework of the Sustainable Development Goals is losing ground as a common reference, with costs that fall first on partner countries. Second, artificial intelligence is being adopted early, mostly in institutions with stronger results systems.

This project is funded by the European Union (EU).

Alejandro Guerrero-Ruiz, alejandro.guerrero@oecd.org
Geovana Zoccal, geovana.zoccal@oecd.org
Ryan Tunç Kennedy, ryan.kennedy@oecd.org

JT03593387

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Fit for Purpose?

2026 Survey on the OECD DAC Guiding Principles on Managing for Sustainable Development Results



Abstract

This report presents the results of the 2026 survey on the OECD DAC Guiding Principles on Managing for Sustainable Development Results. It covers 52 bilateral and multilateral development co-operation institutions, comparing results with the 2020 baseline. Providers report progress on all six principles, and results frameworks now reach more levels of their institutions. However, the organisational culture, skills, incentives and resources needed to use those frameworks have not kept pace. Results information is still used more for accountability than learning. This report proposes a typology that places institutions in four positions, according to their capability and the balance between formal and tacit conditions for results management. It also examines two developments since 2020. First, the indicator framework of the Sustainable Development Goals is losing ground as a common reference, with costs falling first on partner countries. Second, artificial intelligence is being adopted early, mostly in institutions with stronger results systems.

Acknowledgements

The OECD would like to thank all the respondents of the 2026 survey on Managing for Sustainable Development Results, as well as all participants of the Peer Learning on Effective Results Frameworks. We hope these findings will help the institutions navigate the current context by strengthening their systems to manage for impact.

The authors are also sincerely grateful to the European Union for funding this work and analysis.

This report was prepared by Geovana Zoccal, Ryan Tunç Kennedy and Alejandro Guerrero (OECD Development Co-operation Directorate). The authors also thank Michaela Ditrych Lenc and Luke White for their support.

Funded by the European Union

Table of contents

Abstract 2
Acknowledgements 3
Executive summary 6

1 Key findings 8
Strengthening a results-focused system 9
Institutional architecture has strengthened, but practice still lags behind 11
Before the dashboard: accountability is routine, learning is not 17
The SDG indicator framework was never used in depth, and is now losing ground 18
AI has arrived in results systems as pilots, mostly where capability was already high 21

2 Findings by Guiding Principle 25
Guiding Principle 1: Support sustainable development goals and desired change 26
Guiding Principle 2: Adapt to context 28
Guiding Principle 3: Enhance country ownership, mutual accountability and transparency 30
Guiding Principle 4: Maximise the use of results information for learning and decision-making 32
Guiding Principle 5: Foster a culture of results and learning 33
Guiding Principle 6: Develop a results system that is manageable and reliable 36

References 38

Annex A. Survey methodology 39
Process 39
Participation sample 39
Methodological considerations 40
Questions 42

Annex B. Survey respondents 44

FIGURES

Figure 1. Collective progress per Guiding Principle, 2020-2026 10
Figure 2. Development of results frameworks at the corporate level, 2020-2026 10
Figure 3. Coverage of results framework architecture, 2020-2026 11
Figure 4. Linkages between formal and tacit conditions that support managing for development impact 12
Figure 5. Formal and tacit conditions for managing for development impact, 2020 and 2026 14
Figure 6. Four positions of results management capability, 2026 15
Figure 7. Practice score by the size of the ODA programme, bilateral providers, 2026 16
Figure 8. Use of results information by purpose, headquarters and country level, 2026 17

Figure 9. Extent to which results frameworks incorporate SDG indicators by level, 2020 and 2026 19
Figure 10. Role of the SDG framework in development providers' results management systems, 2026 20
Figure 11. Development providers' institutional approaches to using AI in management processes, 2026 21
Figure 12. Use of AI across management functions, 2026 22
Figure 13. AI adoption and institutional capability, 2026 23
Figure 14. Mean score by provider type and Guiding Principle, institutions answering both rounds, 2020-2026 25
Figure 15. GP1 trajectory table by provider type, 2020-2026 26
Figure 16. The extent to which established methodologies are used to define the results chain, by level, 2020-2026 27
Figure 17. The extent to which cross-sector approaches are promoted and included into measurement, learning and reporting systems, 2020-2026 27
Figure 18. GP2 trajectory table by provider type, 2020-2026 28
Figure 19. Flexibility for managers at country level at design and implementation phase, 2020-2026 29
Figure 20. GP3 trajectory table by provider type, 2020-2026 30
Figure 21. GP4 trajectory table by provider type, 2020-2026 32
Figure 22. GP5 trajectory table by provider type, 2020-2026 33
Figure 23. Leadership support for a culture of results and learning, 2020-2026 34
Figure 24. GP6 trajectory table by provider type, 2020-2026 36
Figure 25. Capability against reported practice, 51 institutions, 2026 41

TABLES

Table 1. Guiding Principles on Managing for Sustainable Development Results 8
Table 2. GP1: mean score by provider type, institutions answering both rounds 28
Table 3. GP2: mean score by provider type, institutions answering both rounds 30
Table 4. GP3: mean score by provider type, institutions answering both rounds 31
Table 5. GP4: mean score by provider type, institutions answering both rounds 33
Table 6. GP5: mean score by provider type, institutions answering both rounds 35
Table 7. GP6: mean score by provider type, institutions answering both rounds 37
Table 8. Participation by provider type 40
Table 9. List of survey questions 42
Table 10. List of survey respondents, 2020 and 2026 44

Executive summary

Fit for purpose? In part. Development co-operation providers manage for results better than they did in 2020, on all six OECD Development Assistance Committee (DAC) Guiding Principles on Managing for Sustainable Development Results. Results frameworks now reach further into institutions. But the culture, skills, incentives and resources that turn those frameworks into practice have not kept pace, and results information still serves accountability far more than learning. Meanwhile, the indicator framework of the Sustainable Development Goals (SDGs) is losing its place as a common reference, and artificial intelligence (AI) is entering results systems, mostly where capability was already high.

The report covers 52 bilateral and multilateral providers, 44 of which also took part in the 2020 baseline survey. It rests mostly on institutions' own assessments.

Progress is broad but uneven. Data show that practice improved on all six principles. The largest gain was in “fostering a culture of results and learning”, and the smallest in “country ownership, mutual accountability and transparency”. “Adapting to context” remains the strongest area, and a “manageable and reliable results system” is the weakest. Two-thirds of institutions now have a corporate results framework. The share with frameworks at every level, from corporate to project, has risen from about one in five to more than one in three.

The architecture has moved ahead of the conditions for using it. Providers rate the formal side of results management, i.e. what exists on paper and in systems, higher than the tacit side, i.e. what lives in people and practice: culture, skills, incentives and the actual use of results information. The gap continues as wide as it was in 2020. Dedicated resources and incentives are the weakest of the 12 building blocks used to rate the conditions needed for fulfilling the six principles, and dedicated resources have barely moved. The budget size of these institutions’ development co-operation is not what separates strong from weak practice for managing for impact. Overall, long-standing DAC members report stronger practice than newer and smaller providers, but within each group larger programmes do not report better practice.

Four positions describe where institutions stand, correlating overall capability and the balance between formal and tacit conditions:

- developing and in step,
- developing with architecture ahead,
- compliant and formal, or
- functional and learning.

More than half of analysed institutions are developing and in step. Positions have been largely stable since 2020. Most institutions that did move ended up with their formal architecture running ahead of the conditions needed to use it. A more elaborate architecture is not necessarily more effective: tacit conditions are lost when people leave and are often the first to weaken when budgets are cut.

Results information serves accountability more than learning. At headquarters, about four in five institutions use results information routinely for accountability, while fewer than two in five use it for learning. Learning and steering are more common at country level, as in 2020. Routine learning is most closely associated with organisational culture and structure.

The SDG indicator framework is losing ground. Commitment to the goals remains, but alignment with indicators is weaker than alignment with the goals. It is also weaker at project level than at corporate level. Only about a quarter of institutions still treat the framework as their single most important reference. In-depth use of its indicators rests increasingly on United Nations (UN) agencies. With no successor to the 2030 Agenda for Sustainable Development yet agreed, providers are re-anchoring their frameworks

independently, leaving partner countries to bear the immediate costs. Providers increasingly align with partner-country indicators, but their use of partner countries' own monitoring and statistical systems has not changed since 2020. The open question is which functions of the framework – a common language, agreed indicators and custodians, investment in national data systems, and a neutral basis for describing progress – can be protected until and beyond 2030.

Finally, AI is arriving as pilots, mostly where capability was already high. More than half of institutions are at least piloting AI in management processes, and integration is concentrated in multilateral development banks and UN agencies. Informal use runs further ahead: more than four in five respondents report AI in use in at least one management function. AI use is centred on evaluation, reporting and communication, where documents are abundant and providers already perform best. Institutions with stronger results systems are further ahead, so AI risks widening the distance between providers. The survey did not ask what AI has changed. This report points to two windows for a collective approach, on interoperability and on governance, both open while systems and norms are still forming.

Taken together, the findings suggest that further gains depend less on adding architecture than on the conditions that make it work. They also depend on using results information to learn and decide, not only to report.

1 Key findings

This report presents the results of the 2026 survey on the Guiding Principles on Managing for Sustainable Development Results [1].

Multiple global crises with direct effects on the development co-operation landscape have increased pressure to achieve and demonstrate impact. The foundations for effectively doing so are strong, results-focused development institutions.

In 2019, the OECD and dozens of development co-operation providers developed a shared benchmark for **what results-focused development institutions could look like: the OECD DAC Guiding Principles on Managing for Sustainable Development Results** (MfSDR). These guiding principles help development organisations navigate complex development co-operation and humanitarian challenges to reach their expected results and support sustainable development effectively. They mark an important step towards recognising the importance of the results agenda as a catalyst for more effective development co-operation.

Table 1. Guiding Principles on Managing for Sustainable Development Results

Guiding Principle (GP)	Brief description
Guiding Principle 1: Support sustainable development goals and desired change	MfSDR approaches should maximise the impact of efforts towards achieving social, economic and environmentally sustainable development that leaves no one behind.
Guiding Principle 2: Adapt to context	MfSDR systems, methods and approaches should be flexible to allow for tailoring to different operational contexts, modalities of engagement, and types of partnerships.
Guiding Principle 3: Enhance country ownership, mutual accountability and transparency	MfSDR approaches should foster development effectiveness by strengthening and using partner countries' systems for strategic planning, monitoring and statistics, and promoting participatory approaches to enhance ownership, mutual accountability and transparency.
Guiding Principle 4: Maximise the use of results information for learning and decision-making	While responding to communication and accountability needs, results information should systematically be used for learning and decision making in order to improve delivery and enhance impact.
Guiding Principle 5: Foster a culture of results and learning	A culture of results should be promoted and sustained through consistent leadership, and supported by appropriate guidance, tools and capacity building as well as proper incentives.
Guiding Principle 6: Develop a results system that is manageable and reliable	While providing credible results information that can be used by all partners, results frameworks, measurement and reporting systems need to be suitable to each organisation's needs and capacities.

[1] This report positions 52 institutions against the OECD DAC Guiding Principles on Managing for Sustainable Development Results: 24 are larger and longer-standing DAC members reporting official development assistance (ODA) statistics to the OECD before 2010 (bilateral providers – Group A), eight are smaller or newer DAC members and providers outside the DAC (bilateral providers – Group B), six are multilateral development banks, 13 are UN agencies and one is a regional financial institution. Of those 52, 46 institutions responded to the survey between April and July 2026 (the responses of one were calibrated), and six were expert-assessed based on previous responses, practice and documentation. Data are available for both rounds (2020 and 2026) for 44 institutions, and the comparisons over time in this report rest on those 44 institutions unless stated otherwise. Responses to the survey are institutions' own assessments of their results systems, unweighted and uncalibrated, unless stated otherwise. Further details on the survey's method, scope and process are described in Annex A, which also contains the questionnaire.

Source: OECD (2019[1]), Managing for Sustainable Development Results: Guiding Principles.

Back in 2019, the development co-operation community was lagging on all those principles, below the level the 2021 baseline report treated as good enough (Guerrero-Ruiz, Schnatz and Verger, 2021[2]). Comparing the 2026 responses with the baseline shows improvement on every principle. Despite this progress, the 2026 findings still show considerable room for improvement in managing for sustainable development results and in applying results approaches for learning, decision-making, accountability and communication.

This report discusses the main figures and key findings of the 2026 survey. Beyond the broad progress, it points to the gaps providers need to address to combine formal management practices for development impact with effective practice, and it proposes a typology that places providers in four positions according to their institutional capability and the balance between the formal and the tacit conditions of results management: developing and in step, developing with architecture ahead, compliant and formal, and functional and learning.

The report also sheds light on two developments of the past five years. **First, the 2030 Agenda is no longer the primary reference for most providers' results systems.** Previous OECD research had shown that most providers aligned to the goals rather than to the targets and indicators. The 2026 responses show that the SDG indicator framework is rarely the single most important reference for a development organisation and is often treated alongside other priorities of equal or greater weight, with no successor framework discussed and with consequences for partner countries' own systems.

Second, institutions are adopting artificial intelligence in their management processes. Most have staff using AI, but institutional integration is uneven. The survey identifies a risk of divergence: leading institutions build their own systems while institutions with fewer resources are left behind, and no shared conventions have yet been agreed.

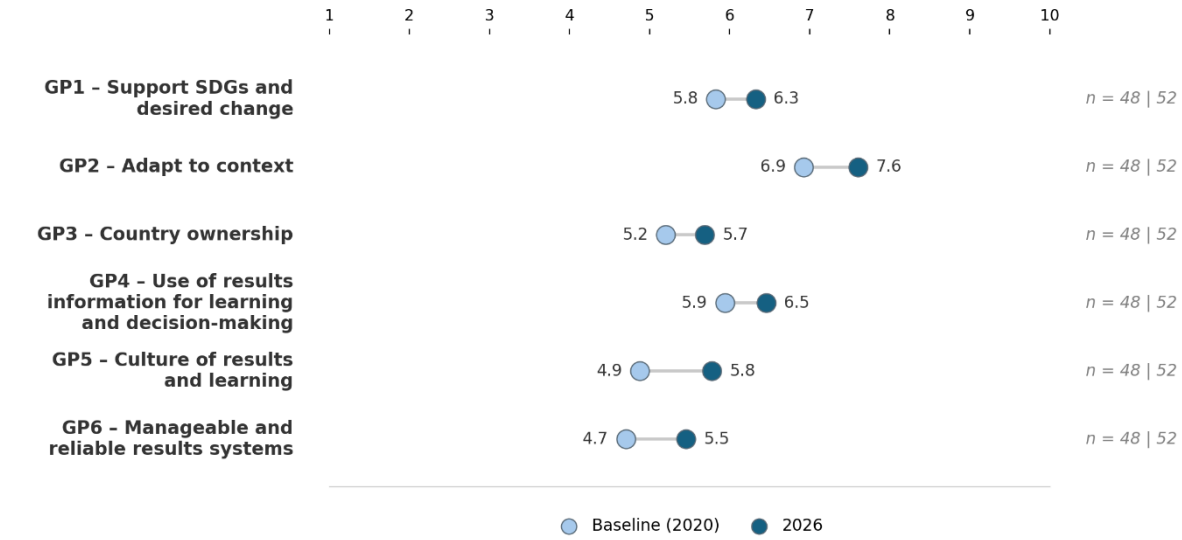
Finally, the report distils the detailed results of the survey per Guiding Principle, followed by annexes on the survey methodology and questionnaire (Annex A) and the list of respondents (Annex B).

Strengthening a results-focused system

Reported practice improved on every Guiding Principle, and adapting to context remains the strongest area. Developing a manageable and reliable results system (Guiding Principle 6) remains the lowest, at 5.5. Among the 44 institutions that answered both rounds, the largest improvement is on the culture of results and learning (Guiding Principle 5, +0.79), followed by manageable and reliable systems (+0.66) and adapting to context (+0.58), while the smallest are on the use of results information (+0.45), supporting the SDGs and desired change (+0.44), and on country ownership (+0.37).

Providers are largely able to adapt to context, with quality context analyses enabling a sound understanding of local dynamics needed to inform country- and project-level results frameworks, the identification and monitoring of conditions for success and risk factors in achieving intended outcomes, and sufficient flexibility for managers at country level to tailor their project design and implementation approaches to local contexts. Guiding Principle 2 (Adapt to context) is the principle on which providers reported performing best, and this has remained the case over time.

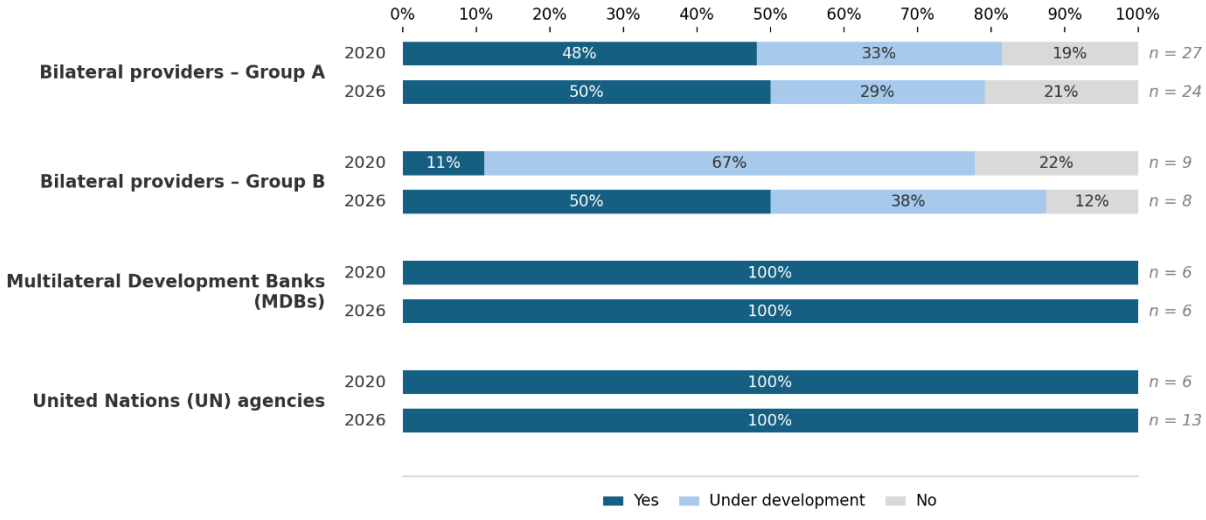
Figure 1. Collective progress per Guiding Principle, 2020-2026



Note: Score is the unweighted mean of the survey items corresponding to each Guiding Principle, measured on a four-point frequency scale and rescaled to 1-10. On the four-point scale, 1 is "no or marginally", 2 "in some cases", 3 "in a majority of cases" and 4 "systematically", so that 7 on the rescaled measure corresponds to "in a majority of cases". The corporate-level results-framework item that uses a different scale (yes, no, under development) is folded into Guiding Principle 1 as yes = 10, under development = 5.5 and no = 1 on the ten-point scale; the item on the role of the SDG framework is excluded. In 2020 guidance and incentives were one item under GP5, counted once. n is the number of responding institutions in 2020 and 2026, respectively.

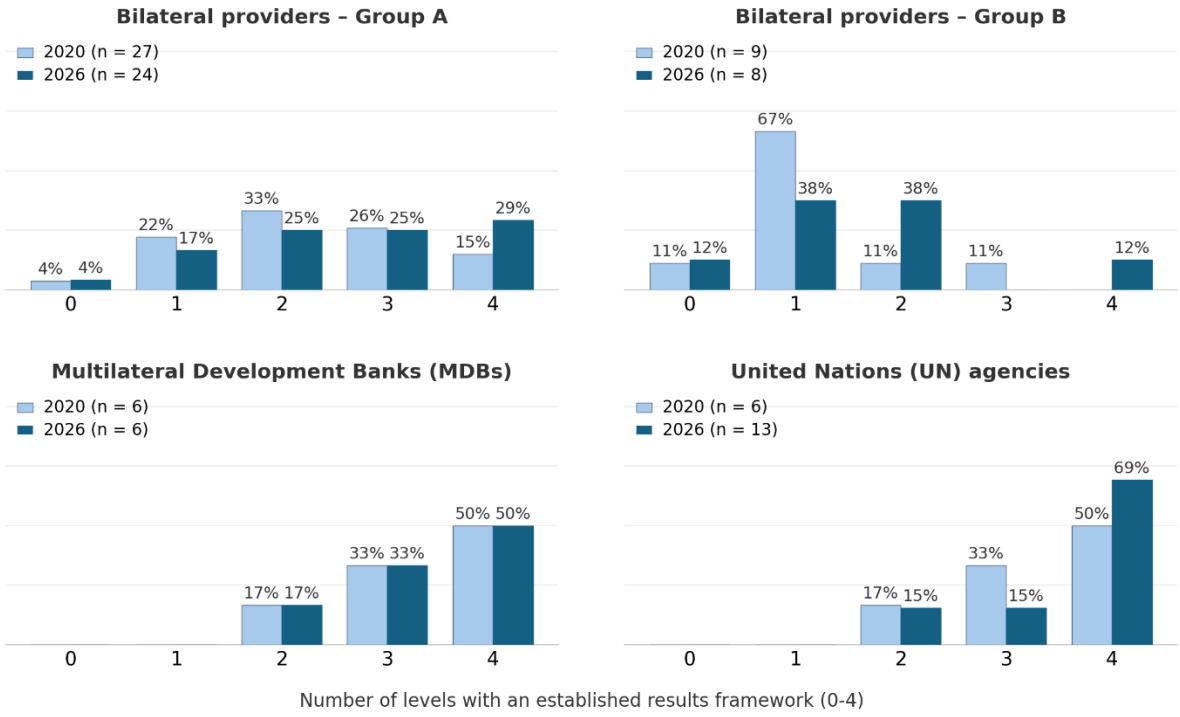
Providers made notable progress in developing and adopting results frameworks at various levels, a key component for impact-oriented management, enlarging their internal results framework architecture. The share of institutions with a corporate results framework in place rose from 54% in 2020 to 67% in 2026, with the largest gain among bilateral providers from Group B (Figure 2). The corporate results framework is close to being established as the standard instrument through which providers translate their overarching strategies into measurable objectives and link resources to intended outcomes (OECD, 2024[3]).

Figure 2. Development of results frameworks at the corporate level, 2020-2026



Over a third of institutions (38%) now have results frameworks established at all four levels (corporate, sector, partner country and project), compared with 21% in 2020; among the 44 institutions that answered both rounds the share rose from 18% to 34%. Nearly three in five institutions (58%) now have frameworks in place at three or more levels, up from 46% in 2020. This vertical coherence, connecting strategic vision at the corporate level to what is delivered on the ground, suggests that **providers are putting in place the structural conditions for collecting and aggregating results information across all levels**. For example, an MDB describes a corporate results framework from which indicators cascade into sectoral and thematic frameworks, regional and country strategies with their own frameworks related to the corporate one, and project frameworks that may incorporate corporate or sector indicators (Figure 3).

Figure 3. Coverage of results framework architecture, 2020-2026

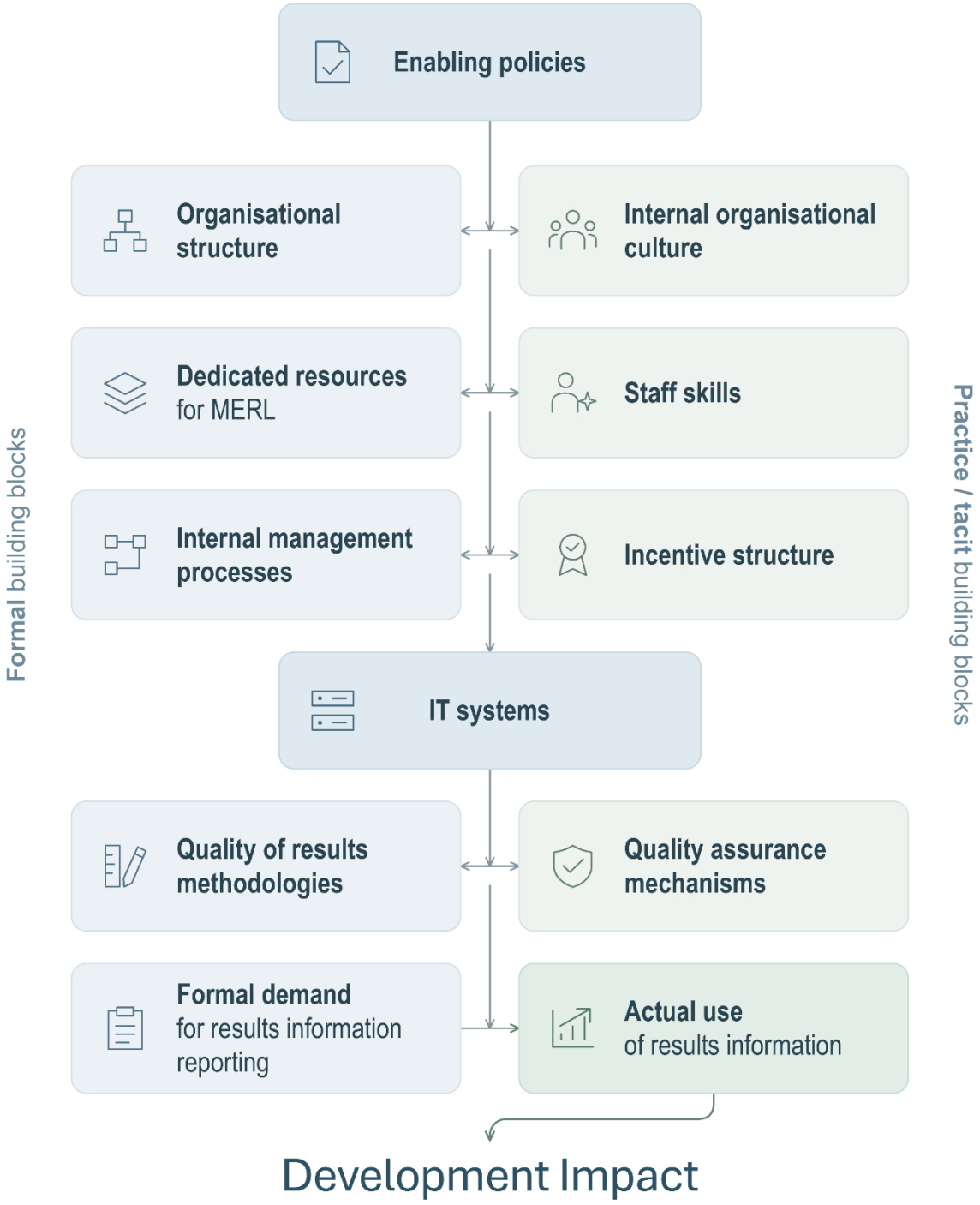


Note: The four levels are corporate, sector, partner country and project. Established results frameworks are counted at levels where a provider reports "systematically" or "in a majority of cases" (sector, partner country and project levels), or answers "yes" at the corporate level. Shares rest on all respondents in each round. The 2020 and 2026 respondents differ: all 2026 bilateral respondents also answered in 2020, while six of the 13 UN agencies and all six MDBs did.

Institutional architecture has strengthened, but practice still lags behind

Two kinds of condition underpin results management. Formal conditions exist on paper and in systems and correspond to the visible architecture of results-based management: policies, organisational structures, dedicated resources, internal processes, methodologies, information technology systems and a formal demand for results reporting. Tacit conditions exist in people and in practice, and translate formal conditions into results-focused operations: organisational culture, skills and training, incentives, quality assurance as it is carried out, and the actual use of results information.

Figure 4. Linkages between formal and tacit conditions that support managing for development impact



Source: Authors' elaboration. Pairings reflect conceptual matches between formal and tacit building blocks.

Institutions can build formal structures for managing sustainable development results, but without meeting tacit conditions, the translation of those structures into practice is limited. For example, an adequate organisational structure and formal processes leave little uptake or collaboration across an organisation where the culture and the skills needed for a results-focused practice are missing. Equally, a

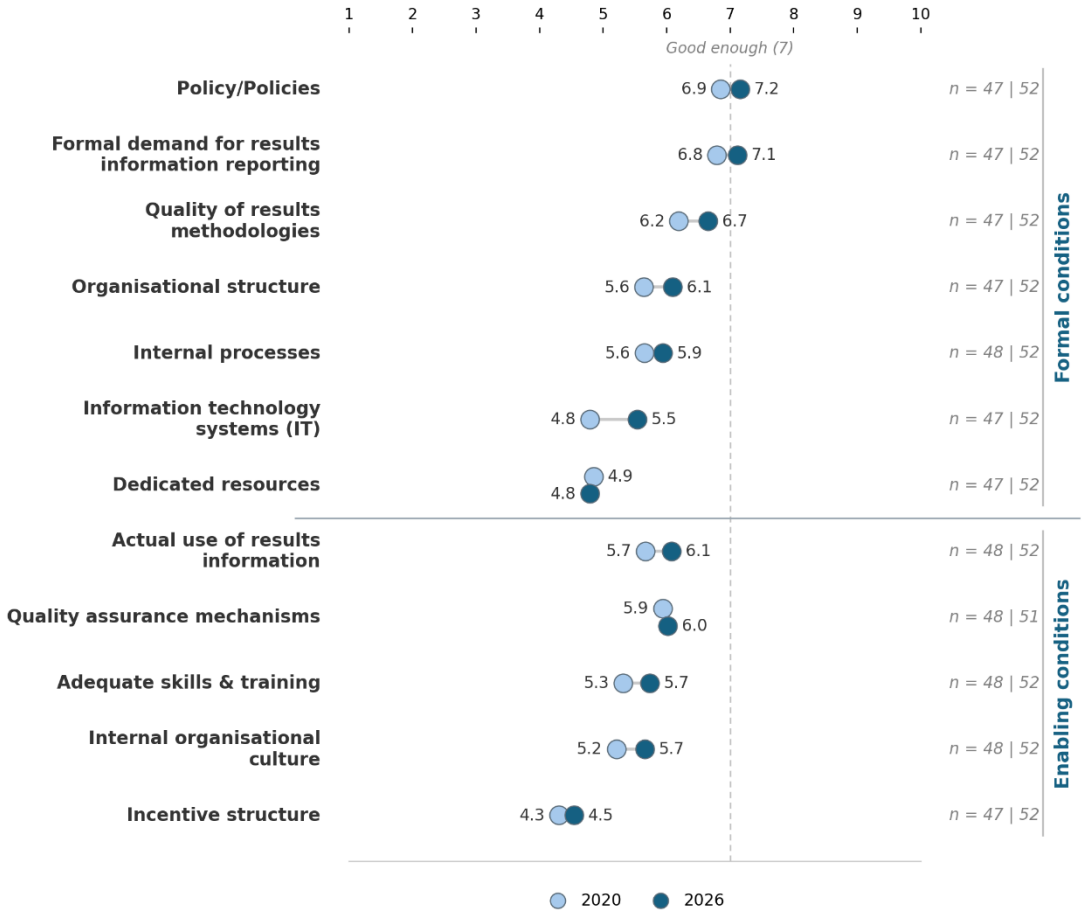
formal demand for results information reporting tends to generate documentation rather than performance management if the demand is not geared towards learning and decision-making.

The survey data show further associations between formal and tacit conditions beyond the pairs drawn in Figure 4. Adequate skills and training correlate with internal processes (r = 0.78, n = 52) and with organisational structure (0.69), organisational culture with internal processes (0.72) and quality assurance mechanisms with the formal demand for results reporting (0.72 n = 51). The correlations among the 12 building blocks are all positive, between 0.25 and 0.78, so institutions that rate one condition high tend to rate the others high as well.

Providers rate their formal conditions higher than the tacit conditions needed to use them, and the gap is a persistent one rather than widening. Across the 52 respondents the seven formal conditions average 6.2 on the ten-point scale and the five tacit conditions 5.6. Institutional change is usually described as slower on culture, skills and incentives than on structures and rules. Among the institutions that rated each block in both rounds (44, or 43 for the seven blocks one institution did not rate in 2020), organisational culture rose by 0.4 points, in line with policy and organisational structure and behind only information technology systems (+0.5). Adequate skills and training rose by 0.3, internal processes by 0.2 and dedicated resources by less than 0.1, while quality assurance mechanisms slipped slightly.

Providers elaborate on specific gaps throughout the survey. A bilateral provider from Group A indicates having produced guidance to encourage staff to adopt context-driven and adaptive approaches to both project design and implementation, but lists tools, systems and staff capacity as factors limiting the extent to which this happens in practice. A UN agency reports having produced corporate guidance including practical steps emphasising the need to use results information for learning, decision-making, and adaptive management, but signals that a shift in office culture is still required to embrace adaptive management and course-corrective practices. The same agency reports strong data availability whose use for steering is not yet consistently embedded in management routines. Another UN agency indicates having principles and examples of good practices in place but reports uneven implementation due to scarce incentives and competing priorities.

Figure 5. Formal and tacit conditions for managing for development impact, 2020 and 2026



Note: Scores are unweighted averages across all responding institutions. In 2020, seven blocks rest on 47 answers and five on 48; in 2026, every block rests on 52 answers except quality assurance mechanisms (51). The 2026 values include the calibrated assessment of one institution and the expert assessments of six institutions (2Annex A).

Dedicated resources and incentive structures are the two weakest dimensions of providers' institutional capability, both below 5 on the ten-point scale. Across all respondents, dedicated resources declined slightly between 2020 and 2026 (from 4.9 to 4.8, Figure 5). The two are correlated across institutions (r = 0.61), as most building blocks are: institutions that rate one low tend to rate the other low. Respondents' comments point to a shared source rather than two separate weaknesses: staff rewarded against financial targets rather than results, decisions driven politically rather than by results information, and competing priorities. Two different management problems sit beneath this. New requirements can be added faster than the capacity to meet them, as when a provider introduces a theory-of-change requirement without the staff time to apply it. Capacity can also fall under requirements that stay. Among the six UN agencies that answered both rounds, dedicated resources fell by 1 point on average, driven by sharp declines in two agencies, while their scores for results methodologies and information systems were broadly unchanged and policies improved slightly. A new reporting template would not address those problems.

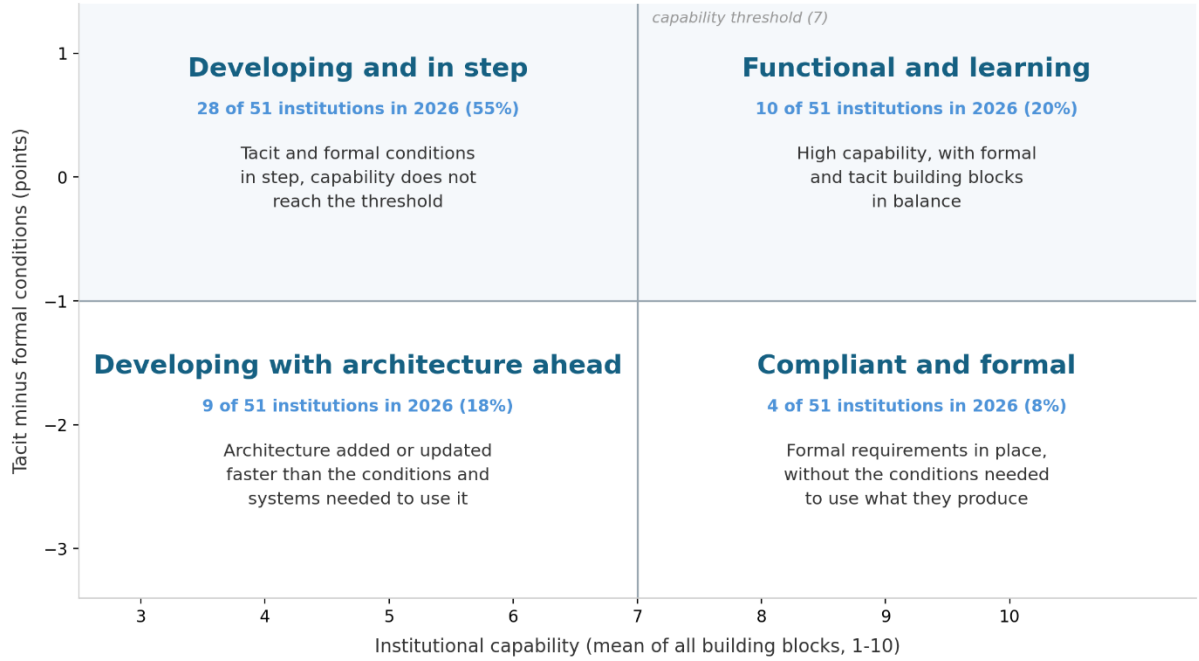
Analysing the overall capability (the mean of the 12 building blocks) and the balance between formal and tacit conditions, providers can be placed in four positions (Figure 6):

- **Developing and in step:** Institutions with capability on average below the good-enough threshold (7), with formal conditions less than a point ahead of tacit conditions. This represents 28 institutions

in 2026, and may reflect a results system still being built, a recent reform not yet visible in effect, or a generally low institutional capability.

- **Developing with architecture ahead:** Institutions with capability on average also below 7, but with formal conditions one point or more ahead of tacit conditions. Nine institutions are in this category in 2026. Architecture has been added or kept while the conditions needed to use it have not followed, either because requirements were added faster than capacity or because capacity fell under requirements that stayed. Frameworks exist, but the information they produce is not yet reliable enough, or used enough, to steer decisions.
- **Compliant and formal:** Institutions with capability at or above 7, with formal conditions one point or more ahead. Four institutions are in this category in 2026. Formal requirements are met and reporting is produced on schedule, while incentives, culture and skills trail. Where the previous position reflects a production problem, this one reflects a use problem despite efficient production.
- **Functional and learning:** Institutions with capability on average at or above 7, and with formal conditions less than a point ahead of tacit conditions. Ten institutions are in this category in 2026, five of which are UN agencies and three MDBs; there is no bilateral provider from Group B. On average, both formal and tacit conditions are scored above the "good enough" threshold. Evaluation findings reach decision makers, and staff have the time, capacity and mandate to respond.

Figure 6. Four positions of results management capability, 2026



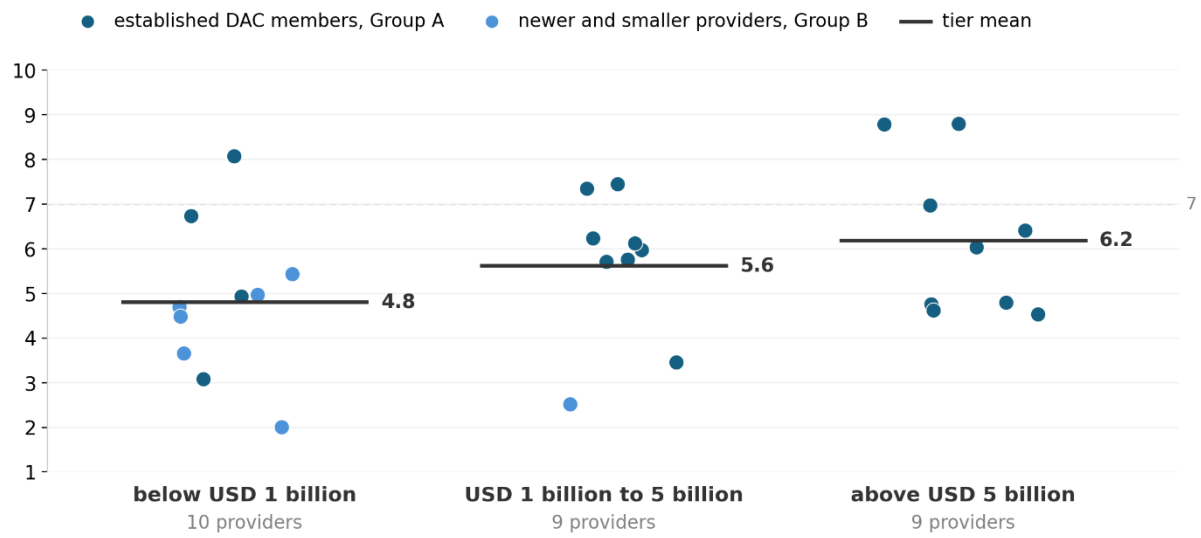
Note: Institutional capability is the mean of the 12 building blocks on the 1-10 scale; the threshold of 7 is the "good enough" point used. The vertical axis is the mean of the five tacit conditions (organisational culture, skills and training, incentives, quality assurance, actual use of results information) minus the mean of the seven formal conditions (policies, organisational structure, dedicated resources, internal processes, quality of methodologies, information technology systems, formal demand for results reporting). Positions on or below the line indicate formal architecture running one point or more ahead. An institution whose formal lead is exactly one point counts as ahead; any institution whose tacit conditions lead counts as in step. The one-point line is indicative. Eleven of the 51 institutions sit within a quarter of a point of it and six within a quarter of a point of the capability threshold (7), so the counts are a diagnostic rather than a classification. Self-assessments, except for six expert-assessed institutions and one calibrated institution (Annex A); 51 of the 52 institutions have ratings for all 12 blocks.

Comparison between 2020 and 2026 results shows that **the positions are overall stable**: of the 43 institutions that rated every block in both rounds, 33 (77%) are in the same position as in 2020. The ideal trajectory runs from *developing and in step* to *functional and learning*: institutions build capability while keeping formal and tacit conditions in step. Responses show a different path: **as capability rises, formal conditions are at risk of advancing faster than the tacit conditions required to use them**, thus pulling some institutions down the vertical axis as they move right along the horizontal one. Among the institutions that did move from one position to another, 40% moved from *developing and in step* to *developing with architecture ahead*, and 20% moved from *developing with architecture ahead* to *compliant and formal*.

Capability and balance require joint attention. Strengthening policies, structures, resources, processes and reporting systems raises measured capability, but without matching investment in the culture, skills, incentives and actual use of results information that make them work, a more elaborate architecture does not mean a more effective one. The tacit conditions are the ones that leave when people leave and they are also the first to erode when budgets fall.

Maturity, not programme size, separates stronger from weaker practice among bilateral providers. Compared by the size of their programmes, the larger bilateral providers report stronger practice on average, though the pattern reflects maturity rather than budget: among the 28 bilateral providers in Figure 7, the long-standing members of the DAC (Group A) average 6.0 on the practice score against 4.0 for newer and smaller providers (Group B). Data show that the size of a programme is not related to the quality of practice within either group, and the largest budgets in the survey include both the strongest results practice among bilateral providers and some of the weakest (2Annex A). What resources buy is machinery. The parts of results management that depend on systems, staff and money, above all a manageable and reliable results system, are the most closely associated with programme size, while the parts that depend on relationships with partner countries do not. Three implications follow. Smaller and newer providers will not close the gap by spending more: they need institutional support, shared tools and peer learning, which together make a collective approach worthwhile for them. Programme size does not by itself translate into a results system that is used: several of the largest organisations indicate having demanding reporting requirements alongside thin resources for acting on reported results. Any collective exercise should thus set its entry conditions in terms of practice, which any provider can meet, rather than in terms of budget, which sorts providers by the wrong criterion.

Figure 7. Practice score by the size of the ODA programme, bilateral providers, 2026

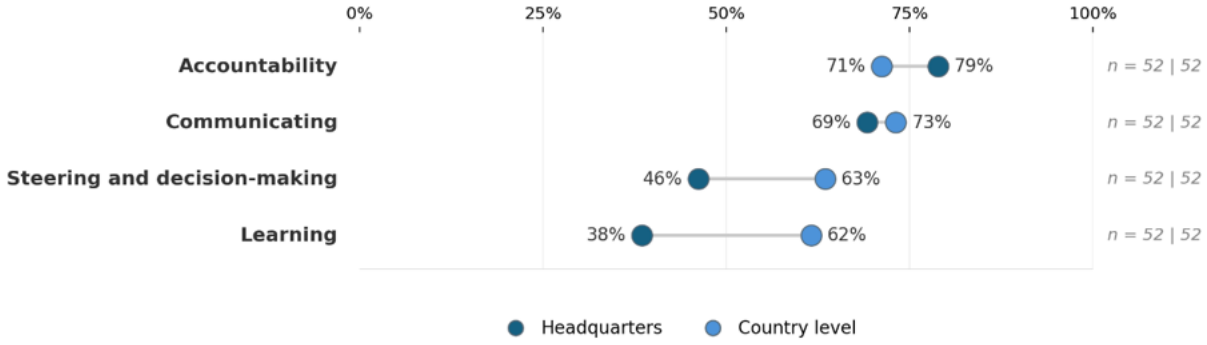


Note: Practice score is the mean of the six principle scores on the ten-point scale. Twenty-eight of the 32 bilateral respondents are shown; tiers cut at USD 1 billion and USD 5 billion. Where two institutions from the same country responded, one is shown, since ODA is reported at country level; one provider does not report ODA to the OECD. Tier means rise with size; within each tier the spread is wide, and the largest tier holds both the highest bilateral score (8.8) and scores below 4.6. Within Group A the rank correlation of size and practice is −0.07. Self-assessments, except for expert-assessed institutions (Annex A); ODA is the mean of 2023, 2024 and 2025 (preliminary data), total ODA on a grant-equivalent basis.

Before the dashboard: accountability is routine, learning is not

Results information is used most routinely for what institutions must do and least routinely for what would change what they do next. At headquarters, 79% of respondents use it systematically or in a majority of cases for accountability and 69% for communicating, while 38% use it for learning. At country level the shares are 71% for accountability and 62% for learning, and use for steering and decision-making is also more widespread there (63% at country level against 46% at headquarters). Communication is routine at both levels (69% at headquarters and 73% at country level) (Figure 8). The 2020 baseline found the same pattern: results information appears to inform learning and adjustment closer to operations, while use at corporate level remains oriented towards accountability.

Figure 8. Use of results information by purpose, headquarters and country level, 2026



Note: Share of the 52 respondents reporting use “systematically” or “in a majority of cases”.

Learning and steering are markedly more common at country level than at headquarters. In 2026, 62% of institutions use results information for learning systematically or in a majority of cases at country level, against 38% at headquarters, and 63% use it for steering and decision-making, against 46% (Figure 8). Bilateral providers in Group A show this contrast most clearly: 17 of 24 use results information for learning at country level, compared with nine at headquarters, and 18 use it for steering at country level, compared with 11. Comments from Group A providers point to mechanisms that bring results closer to operations. One provider has placed monitoring, evaluation and learning officers in a growing number of projects and anchored results in country-level bodies such as steering committees and programme management boards. Another reports continual follow-up of results at project level, with adjustments made where relevant, and a third notes that many of its country offices use results data for planning and decisions. A fourth describes new governance processes that support the use of evidence at project, programme and country level, while acknowledging that this remains “a work in progress”.

The data suggest that learning from results is most closely related to organisational culture (rank correlation 0.57 at headquarters and 0.46 at country level) **and organisational structure** (0.57 and 0.49), **followed by dedicated resources** (0.54 and 0.42) **and incentives** (0.51 and 0.35). These figures are

associations in self-assessed data, not proof of cause and effect: they nonetheless suggest that efforts to strengthen learning need to address culture alongside resources and incentives.

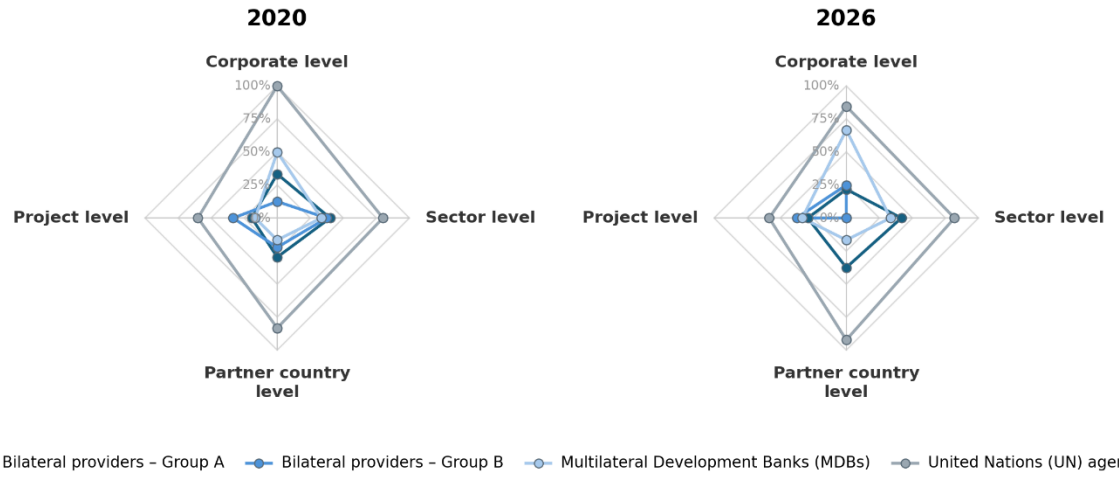
The SDG indicator framework was never used in depth, and is now losing ground

Beyond a set of broad and ambitious goals, the 2030 Agenda gave the development co-operation community a shared results infrastructure, including an indicator framework with agreed definitions, methodologies and designated custodian agencies, currently comprising 234 unique indicators (UN, 2026[4]). This gave the development co-operation community a common language for describing the same result, removed the need for each provider to define its own metrics, justified sustained investment in national statistical systems and offered a politically neutral basis for describing progress across countries at any income level. The 2026 survey shows that commitment to the goals is largely intact, but subsequent layers of their implementation struggle to reach providers' management systems, and this is especially true for bilateral providers.

Previous OECD research documented the shallowness of that alignment well before 2026. The 2020 baseline survey found that, apart from UN agencies, most providers made limited use of SDG indicators in their results frameworks, least of all at project level. More than half had a corporate results framework aiming fully or partly at SDG results, but only a third used SDG indicators systematically or frequently in their country-level frameworks. Alignment of corporate results frameworks also relied on official, derived or proxy indicators to varying degrees (Guerrero-Ruiz, Schnatz and Verger, 2021[2]). The OECD comparative assessment of SDG alignment in seven partner countries and 56 providers found that adoption thinned as it moved from corporate to project level. At country level, providers used the SDGs as a narrative and a source of legitimacy without extensively adapting their results frameworks. The assessment identified recurring obstacles: limited steering from headquarters, internal pressure for short-term attributable results, weak results-based management systems and gaps in SDG data (OECD, 2023[5]). The 2024 toolkit on Effective Results Frameworks indicates that alignment to goals alone is the easiest approach and the weakest, and carries the risk of "SDG-washing" (OECD, 2024[3]).

The 2026 survey confirms these findings. Alignment holds at the level of goals and narrative but reduces as it approaches indicators and delivery. A third of institutions align systematically to the SDGs at the corporate level, and this level of alignment reduces to only a fifth at the project level. Among institutions assessed in both rounds, SDG integration is unchanged at corporate level (43%) and at sector level (43%), and higher at partner-country level (32% to 36%) and project level (24% to 38%), and the higher 2026 value at corporate level reflects the more than doubling of UN respondents. Providers describe the pattern in their own words: a bilateral provider from Group A notes that the SDGs are clearly underlined in its international development policy, and the indicator framework less so. Another maintains broad linkages to the goals but limited links to targets, because it reports thematically at aggregate level to serve its parliament and does not articulate its contribution to SDG results in partner countries.

Figure 9. Extent to which results frameworks incorporate SDG indicators by level, 2020 and 2026



Note: The graph shows the percentage combining "systematically" and "in a majority of cases". Shares rest on 44, 44, 48 and 46 respondents in 2020 and on 51, 50, 52 and 51 in 2026, at corporate, sector, partner-country and project level respectively.

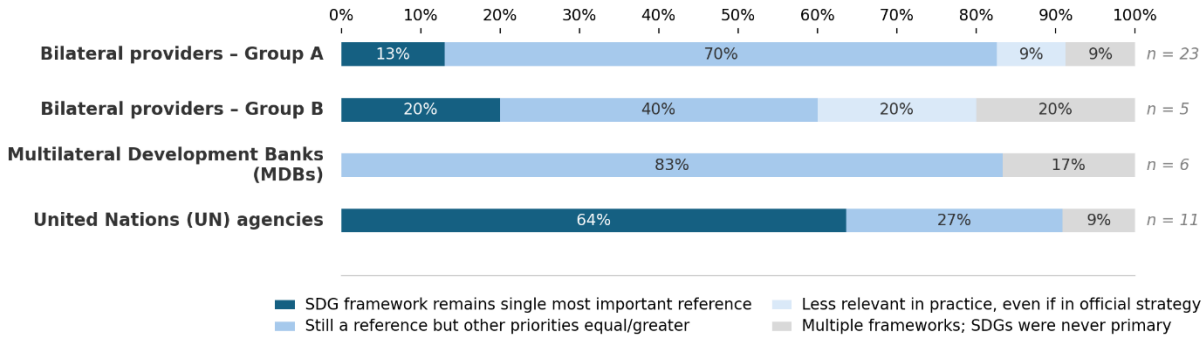
Depth of SDG alignment follows provider type. The United Nations system champions the integration of the SDGs into results frameworks, unsurprisingly, given that the UN system is the custodian of the SDG agenda. Eleven of the 13 responding UN agencies incorporate SDG indicators "systematically" or "in a majority of cases" at the corporate level, and 12 of 13 do so at the partner-country level. Among bilateral providers, the corresponding share is around one in four at corporate level, and no bilateral provider from Group B reports it at sector and partner-country levels. Four of six MDBs report it at the corporate level but this share drops at other levels. The overall picture points to a framework that is not equally widespread across development co-operation providers, suggesting it is becoming more reliant on the institutions whose mandates require them to utilise it.

The depth of alignment may also be linked to institutional capability, and providers find routes to alignment other than incorporating SDG indicators in their results frameworks. A bilateral provider from Group A reports SDG management capabilities that are restricted to a statistical classification of budget allocations by SDG rather than to SDG targets and indicators. Another reports that its implementing agency has developed a set of standard indicators largely derived from the SDGs, while the non-governmental organisations it funds retain the freedom to select their own indicators. A third relies heavily on multilateral contributions through UN agencies that are strongly aligned with the SDG framework.

In parallel, the role of the framework has changed in recent years, reflecting changes in the development co-operation community. Of the 46 institutions answering this question, only 24% still treat it as their single most important reference, and over half of them are UN agencies while none are MDBs. While the majority of respondents (57%) describe it as still a reference but with other priorities of equal or greater weight, 13% report that multiple frameworks apply and that the SDGs were never primary, and the remaining 7% report that it has become less relevant in practice even where it remains in official strategy (Figure 10).

Figure 10. Role of the SDG framework in development providers' results management systems, 2026

"Since 2021, the strategic environment for development co-operation has evolved. Which of the following best describes how this has affected the role of the SDG indicator framework in your organisation's results management?"



Growing domestic and accountability requirements have been reshaping the role of the SDG framework. A bilateral provider from Group A reports that the SDG framework remains the basis of its results management, while national interests defined through its foreign and security policy have become an important reference alongside it. Two bilateral providers from Group B report that security and stability in their own region now sit alongside their SDG-related objectives, and one of them has developed indicators for that part of its co-operation.

The political nature of the SDGs makes the framework easier to embrace at the level of formal commitments, while making consistent operational integration more difficult to sustain. A bilateral provider from Group B reports that the SDG framework is embedded in its current development policy, but that the weight of the framework at implementation level has fallen notably since the policy was adopted. During a joint reflection session, participants observed that where strategic documents are negotiated among governments, language is drafted broadly enough to accommodate divergent positions, which allows formal alignment to be sustained while operational priorities move elsewhere.

Methodological choices and institutional priorities can limit reliance on the SDG framework. A bilateral provider from Group A maintains a small set of standard indicators covering all relevant sectors, deliberately without targets and used predominantly for communication, to limit the distortions that easily measurable standard indicators can produce. An MDB reports that its work is driven by the objectives in its own mandate rather than by the SDG framework, and that the framework's priority has fallen further over the past year. Another MDB has adopted a single cross-cutting corporate priority as its organising reference.

Providers are re-anchoring their results frameworks individually at a time when no successor to the 2030 Agenda has been agreed. The Pact for the Future invites the high-level political forum convened under the auspices of the General Assembly, meeting in September 2027, to consider how sustainable development is advanced by 2030 and beyond (UN, 2024, pp. annex, para. 31(b)[6]). In this context, providers are making decisions on corporate frameworks, guidance, information systems and indicators mostly individually, without co-ordination or shared conventions.

The cost of an uncoordinated transition falls first on partner countries. Only 37% of institutions report that their monitoring approaches maximise the use of partner countries' monitoring and statistical systems, and 56% report aligning country-level results frameworks with partner-country frameworks at indicator level. Reliance on country systems is not rising. Among institutions that answered both the 2020 and 2026 surveys, alignment of country-level results frameworks with partner-country frameworks at indicator level

has increased, while the use of partner-country monitoring systems has not changed. Providers are increasingly measuring the same things as partner countries, but they still largely produce the data themselves. The 2011 Busan commitments and SDG indicator 17.15.1 track both aspects of using country-owned results frameworks, and on this evidence progress has been limited to one of them. The SDG framework, built and maintained multilaterally, is unevenly being replaced by frameworks defined by each provider. Each institution re-anchoring on its own is rational for that institution and expensive for the development co-operation community and for the national statistical systems the framework justified investing in.

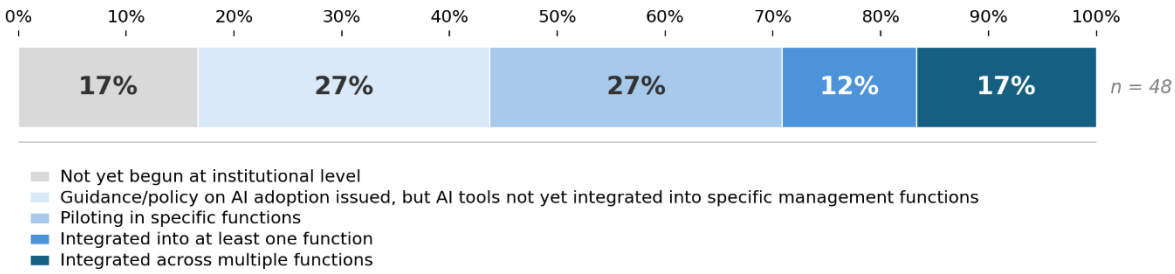
In this scenario, the useful question is **which of the functions played by the SDG framework** (such as the common language across providers and partner countries, the indicator framework of agreed definitions, methods and custodians, the investment in country data infrastructure, and the neutral basis for describing progress) **can be protected until and beyond 2030**.

AI has arrived in results systems as pilots, mostly where capability was already high

Institutions are adopting AI in their management processes, but for most of them adoption has not gone beyond guidance or piloting, and integration is concentrated in multilateral organisations. Of the 48 institutions answering the AI-related questions, over half have reached at least the piloting stage: 27% are piloting AI in specific functions, 12% have integrated it into at least one function and 17% across multiple functions. However, 27% have issued guidance or policy without integrating tools, and 17% have not begun at institutional level (Figure 11). Ten of the 14 institutions reporting integration are MDBs or UN agencies, out of the 18 MDBs and UN agencies responding to the question and against four of the 29 bilateral respondents.

Figure 11. Development providers' institutional approaches to using AI in management processes, 2026

"Which of the following best describes your organisation's current institutional approach to using artificial intelligence (AI) in its management processes?", n = 48.

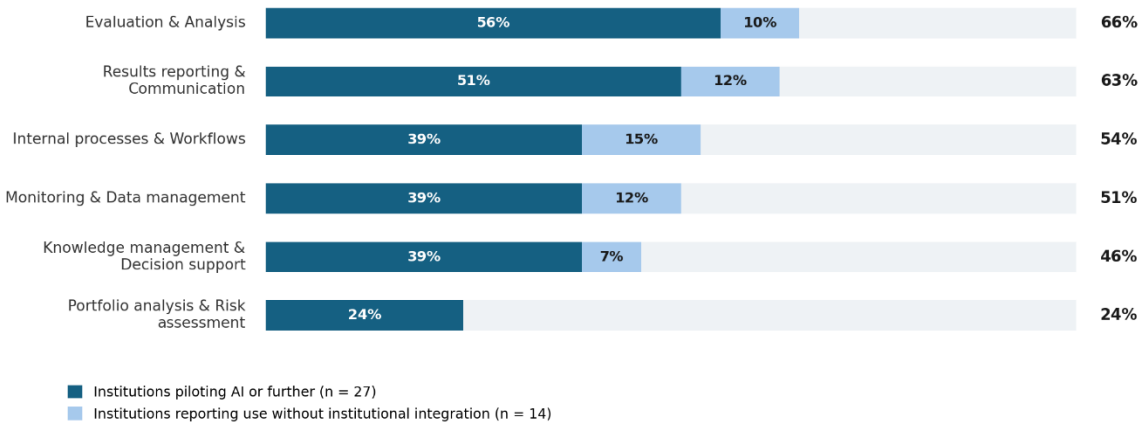


Informal use runs ahead of institutional adoption. Thirteen of the 21 institutions that indicate not reaching piloting stage nonetheless report using AI in at least one management function. Across all respondents, 83% report AI in use in at least one function. For a sizeable group of organisations, the tools are already inside the institution while, at institutional level, adoption has not gone beyond guidance or has not yet started. This pattern of individual use running ahead of organisational adoption is also described in the OECD working paper on AI in development co-operation (OECD, 2026[7]).

AI use tends to be quicker when there is abundant input data for AI-powered tools, even if the need for quality and efficiency gains is not necessarily the highest. AI is most used for evaluation and analysis (66%) and results reporting and communication (63%), followed by internal processes and workflows (54%), monitoring and data management (51%) and knowledge management and decision support (46%). It is least used for portfolio analysis and risk assessment (24%) (Figure 12). Monitoring and data management, where the data are structured, ranks fourth. This suggests that the abundance of documents, rather than the value of the information for learning and decision-making, is what initially draws AI. The pattern matches the knowledge tools that several institutions have built independently on curated collections of evaluation and programme documents (OECD, 2026[7]). The survey, however, did not ask what the outputs changed.

Figure 12. Use of AI across management functions, 2026

"In which management functions is your organisation currently using artificial intelligence (AI)?"

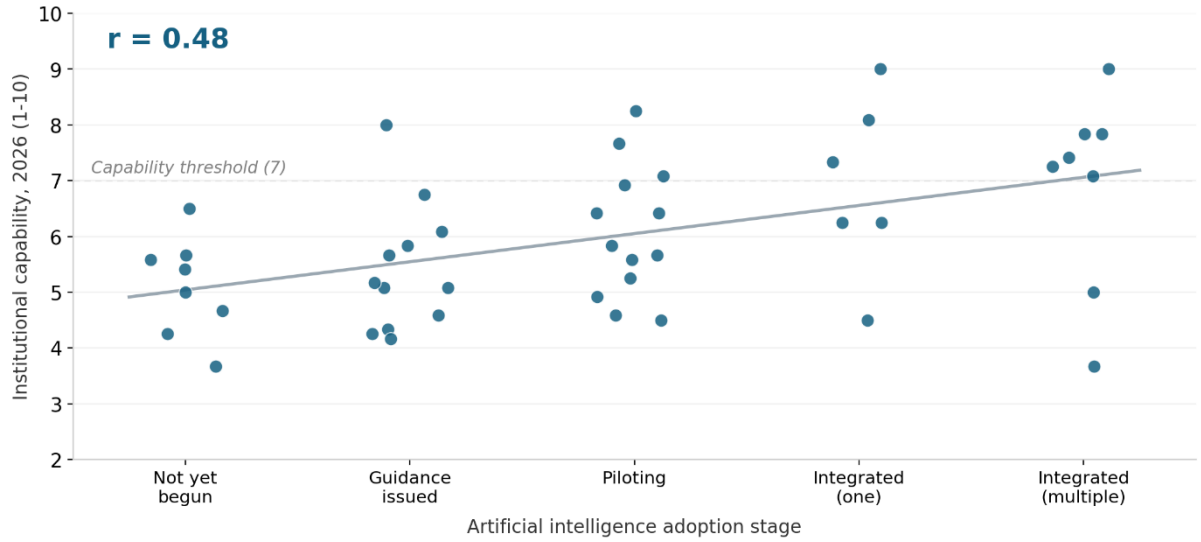


Note: Share of the 41 institutions answering the question on management functions. Institutions could select more than one function, so shares do not sum to 100. The darker segment covers institutions reporting piloting in specific functions, integration into at least one function, or integration across multiple functions (n = 27). The lighter segment covers institutions reporting neither (n = 14), of which eight had issued guidance or policy on AI without integrating tools into specific functions, five reported not having begun at the institutional level, and one did not answer the question on institutional adoption.

AI enters institutions where results information is used primarily for accountability rather than for learning. As 79% of respondents use results information "systematically" or "in a majority of cases" for accountability at headquarters and only 38% for learning, AI is first deployed on evaluation, reporting and communication, which are the areas providers already perform best. Whether AI use strengthens learning and decision-making depends on the arrangements the institutions develop for using what is produced using AI.

If AI adoption catalyses existing capability, it risks widening the distribution rather than narrowing it. Institutions with stronger results management systems report more advanced stages of AI adoption. Institutional capability and stage of AI adoption are positively associated (r = 0.48 on the 47 institutions that rated all 12 blocks; rank correlation 0.46). Of the 13 institutions in Figure 13 at or above the capability threshold of 7, 12 are piloting AI or have integrated it into at least one function and one has not yet reached the piloting stage.

Figure 13. AI adoption and institutional capability, 2026



Note: n = 47, as one of the 48 institutions answering the question on institutional approach did not rate all 12 blocks and is not shown. Institutional capability is the mean of the 12 building blocks on a 1-10 scale. The capability threshold of 7 corresponds to the "good enough" point. The stages are ordered categories, not equal intervals.

The same pattern holds across provider types. MDBs report AI in use in 4.2 of the six management functions on average, UN agencies in 3.2, bilateral providers from Group A in 2.2 and bilateral providers from Group B in 0.5.

Two windows for a collective approach to AI adoption are open. The first concerns interoperability, and it is open because leading providers are building and using their systems at the same time. The OECD working paper documents how several institutions have converged independently on the same design for AI knowledge tools: a curated corpus of evaluation or programme documents, answers traceable to their sources, explicit limits on reliability, and filters by theme, geography or period (OECD, 2026[7]). Conventions on definitions, metadata and what counts as AI-assisted reporting are cheaper to agree while such systems are being built than to retrofit afterwards. They would allow institutions with fewer resources to build on shared infrastructure rather than from scratch, and they would spare partner countries a further layer of incompatible systems.

The second window concerns governance. Informal use sets its own norms before a defined framework exists. During a peer learning session, providers observed that where the authorising environment feels insecure, risk-averse voices can prevent a considered discussion, leaving staff to use tools informally, without shared rules or agreed good practice, and that adoption proceeded more openly where AI was framed as a support to staff rather than as a replacement. The working paper points out that a minimum viable institutional framework for AI governance should include rules on which data may be used with which tools, an approved tool list, guidance on prompting, output verification and citation, human oversight calibrated to what is at stake, an incident reporting path, vendor due diligence, a sandbox for experimentation, training, adoption measurement and a mechanism for sharing lessons (OECD, 2026[7]).

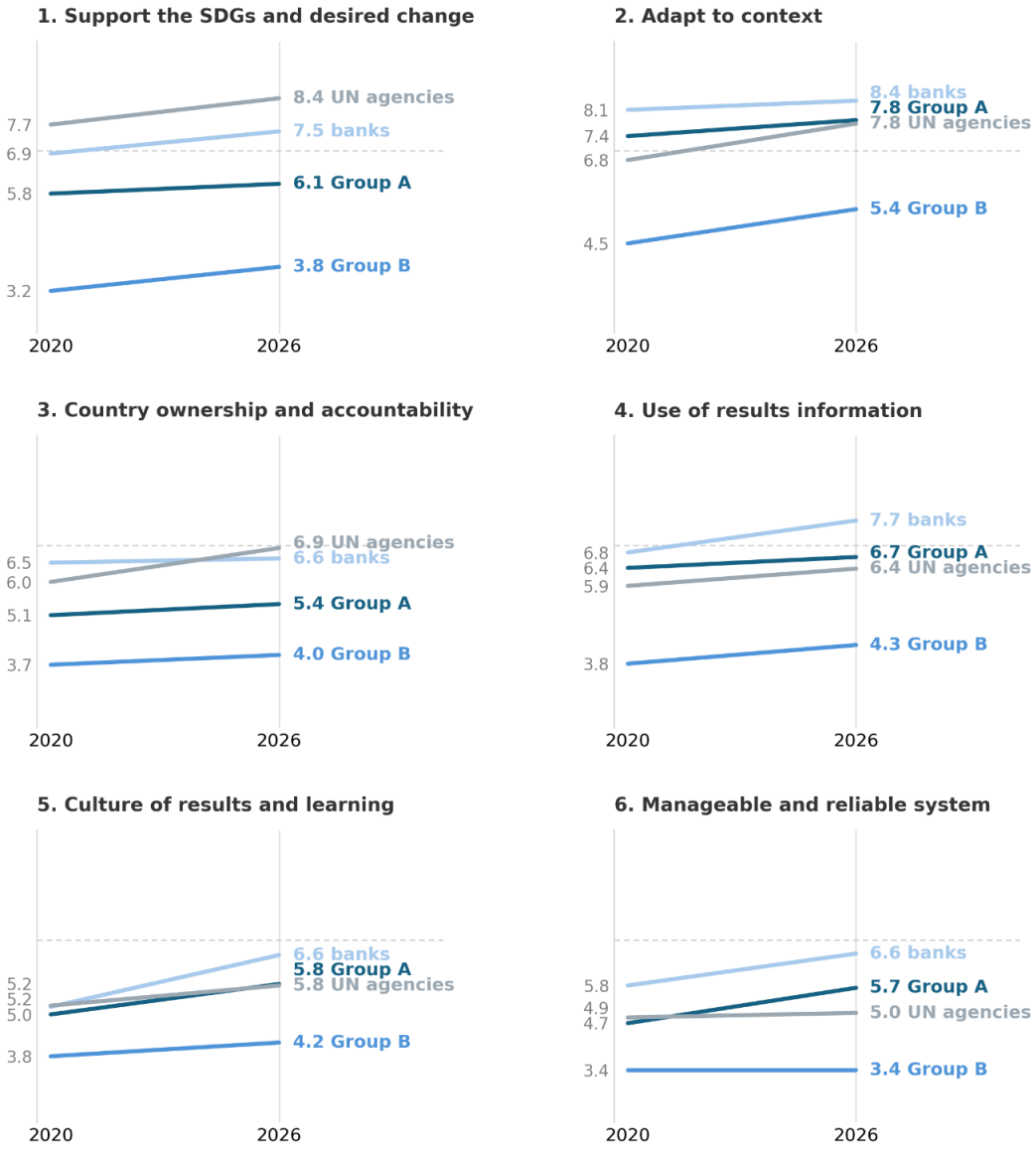
These findings reflect institutional status and practices at a specific point in time within a rapidly evolving field. Both questions were asked for the first time in 2026, so no baseline comparison is available and no trend can be inferred. Responses are self-assessed, and the question on institutional approach captures an organisation's formal position rather than the extent of use. The survey did not ask whether

AI has changed a decision, saved time or improved quality, and no respondent reports an evaluation of its own AI use yet.

2 Findings by Guiding Principle

At the time of their approval, the development co-operation community was lagging on all the OECD DAC Guiding Principles on Managing for Sustainable Development Results, below the level the 2021 baseline report treated as good enough. These principles were assessed via a number of detailed items. This section examines each Guiding Principle in turn.

Figure 14. Mean score by provider type and Guiding Principle, institutions answering both rounds, 2020-2026



Note: Mean of the principle's items on the four-point scale, rescaled to 1-10 as in Figure 1, for the institutions that answered both rounds (24 bilateral providers from Group A, eight bilateral providers from Group B, six multilateral development banks, and six UN agencies). Self-assessments in both years, except for the six institutions expert-assessed in 2026 (Annex A). Dotted line at the "good enough" threshold of 7.

Guiding Principle 1: Support sustainable development goals and desired change

Figure 15. GP1 trajectory table by provider type, 2020-2026

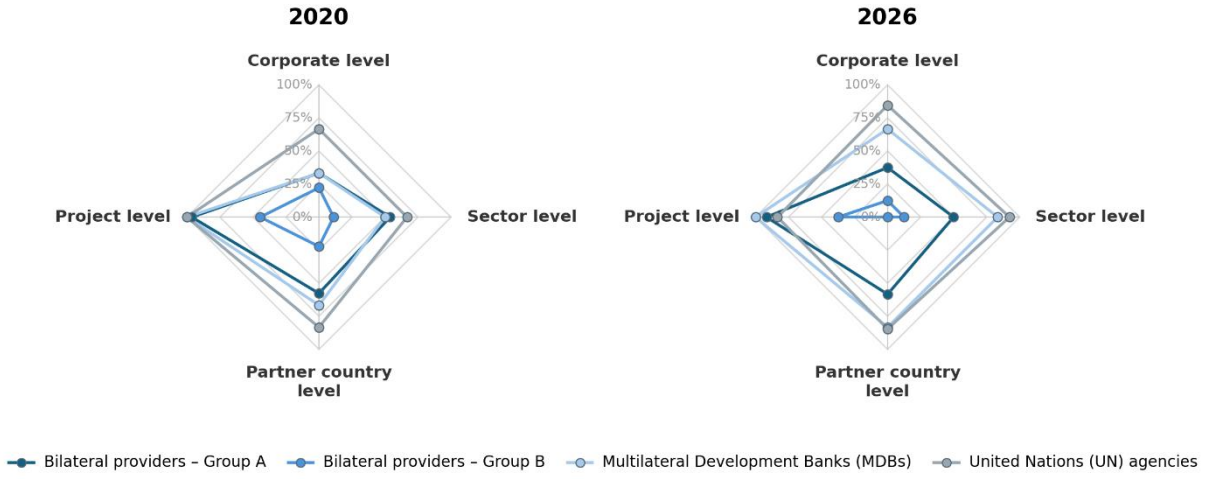
Survey item	Group A	Group B	MDBs	UN agencies
Guiding Principle 1 – Support SDGs and desired change				
Results Frameworks are developed at corporate level	→ 50%	↗ **50%**	→ 100%	→ 100%
Results Frameworks are developed at sector level	↗ **50%**	→ 12%	→ 50%	↗ 100%
Results Frameworks are developed at partner country level	↗ **67%**	→ 12%	→ 83%	→ 83%
Results Frameworks are developed at project level	→ 92%	→ 88%	→ 100%	→ 100%
To what extent do your results frameworks incorporate SDG results indicators at corporate level?	↘ 22%	↗ 25%	↗ 67%	→ 100%
To what extent do your results frameworks incorporate SDG results indicators at sector/thematic level?	→ 42%	↘ **0%**	→ 33%	↗ 100%
To what extent do your results frameworks incorporate SDG results indicators at partner country level?	↗ 38%	↘ 0%	→ 17%	↗ 100%
To what extent do your results frameworks incorporate SDG results indicators at project level?	↗ 29%	↗ 38%	↗ 33%	↗ 80%
Established methodologies (e.g. logical framework, theory of change) are used to define the results chain at corporate level	→ 38%	→ 12%	↗ **67%**	↗ 83%
Established methodologies (e.g. logical framework, theory of change) are used to define the results chain at sector/thematic level	→ 50%	↗ 12%	↗ **83%**	↗ **100%**
Established methodologies (e.g. logical framework, theory of change) are used to define the results chain at partner country level	→ 58%	↘ 0%	↗ 83%	↗ 100%
Established methodologies (e.g. logical framework, theory of change) are used to define the results chain at project level	↘ 92%	→ 38%	→ 100%	→ 100%
The design of the organisational structure is geared towards achieving development results.	↗ 54%	↗ **38%**	→ 83%	→ 83%
In line with the integrated approach of the SDGs, cross-sector approaches are promoted and included into measurement, learning and reporting systems.	↗ **50%**	↗ 12%	↗ 83%	→ 50%

↗ Strong increase ↗ Slight increase → Stable ↘ Slight decrease ↘ Strong decrease

Note: Share of respondents selecting one of the top two response categories, i.e. "systematically" or "in a majority of cases" ("to a great extent" for items on an extent scale), unweighted and calculated separately for each provider type, among the 44 institutions that responded to both rounds. The percentage beside each arrow is the 2026 share. The thresholds separating a strong change, a slight change and a stable position are scaled to the number of respondents in each provider type, using the standard error of the change with a small-sample adjustment: a change counts as strong from one standard error and as slight from half of one. For the item on corporate results frameworks, which uses a yes / under development / no scale, the share is of institutions answering yes.

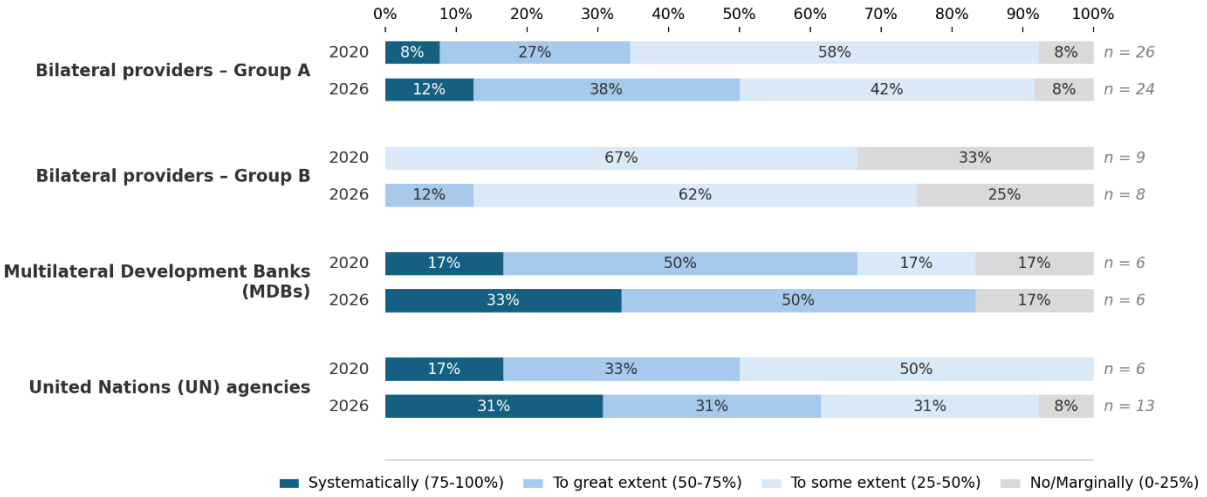
Established methodologies such as logical frameworks and theories of change are among the essential elements of management systems oriented towards outcomes and impact. They help organise objectives, activities and results and clarify the causal links and assumptions between them, providing a coherent basis for planning, implementation, monitoring and evaluation (OECD, 2024[3]). Data suggest that they are **mostly used at project level**, where 82% of providers apply them systematically or in a majority of cases, against 50% at corporate level. However, the gap between project and corporate levels has narrowed, partly because methodological practice extended upwards (corporate use rose from 36% to 50% of respondents, and from 37% to 44% among the 41 institutions answering the item in both rounds) and partly because project-level use slipped (from 87% to 82%) (Figure 16).

Figure 16. The extent to which established methodologies are used to define the results chain, by level, 2020-2026



In line with the integrated approach of the SDGs, **providers are increasingly promoting cross-sector approaches and including them in measurement, learning and reporting systems**. They have notably become predominant among MDBs (Figure 17). Some providers describe explicit cross-sector arrangements: one organises its sectoral plans and targets around a single cross-cutting area, and another runs cross-sector knowledge programmes in fields such as fisheries and taxation. Others report related shifts that, while not framed in these terms, may create conditions for cross-sector approaches, including a move from a sectoral to a country-based approach with sectoral results frameworks no longer prioritised.

Figure 17. The extent to which cross-sector approaches are promoted and included into measurement, learning and reporting systems, 2020-2026



Established methodologies are concentrated where activities are designed rather than where corporate objectives are set. Among the 24 bilateral providers from Group A, frequent use of established methodologies runs at 92% at project level and 38% at corporate level; among the eight bilateral providers from Group B, three report it at project level and one at corporate level. Project methodologies are thus more widespread than the methods that explain how activities contribute to corporate objectives: for example, a provider from Group B indicates that "most of the indicators remained activity-based or output-based. Strategic objectives were not sufficiently linked to specific results of projects." In contrast, 11 of the

13 UN agencies report extensive use of corporate methodologies, and one describes linking its corporate framework to country programmes and collective objectives.

Of the 28 comments provided by respondents on this principle, 17 qualify the progress they report. Some describe new frameworks whose operations are still being developed; others distinguish broad alignment with the SDGs from the use of their targets and indicators. The management task is to make the intended contribution legible from project choices through to corporate priorities, while retaining the country context.

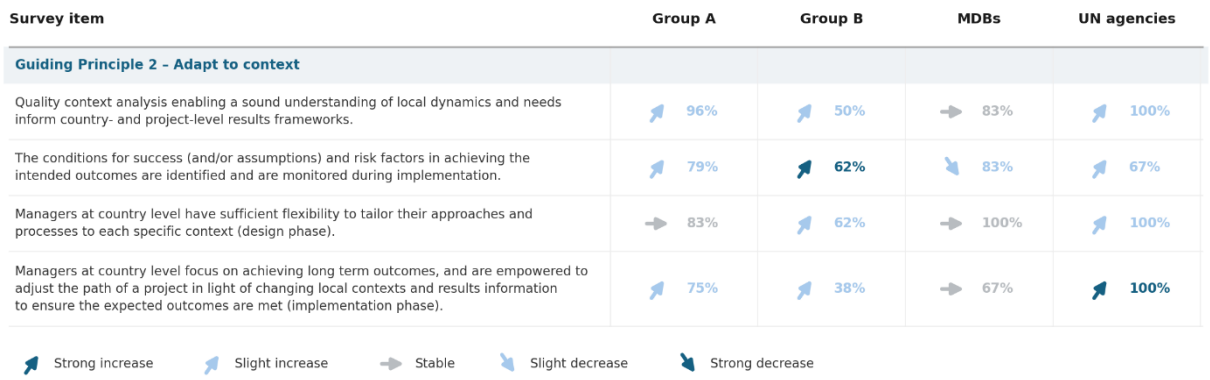
Table 2. GP1: mean score by provider type, institutions answering both rounds

Provider type	Institutions in both rounds	2020	2026
Bilateral providers/Group A	24	5.8	6.1
Bilateral providers/Group B	8	3.2	3.8
Multilateral development banks	6	6.9	7.5
United Nations agencies	6	7.7	8.4

Note: Mean of the principle's frequency items on the four-point scale, rescaled to 1-10 as in Figure 1, for the institutions that answered both rounds. Self-assessments in both years, except for the six institutions expert-assessed in 2026 (Annex A).

Guiding Principle 2: Adapt to context

Figure 18. GP2 trajectory table by provider type, 2020-2026

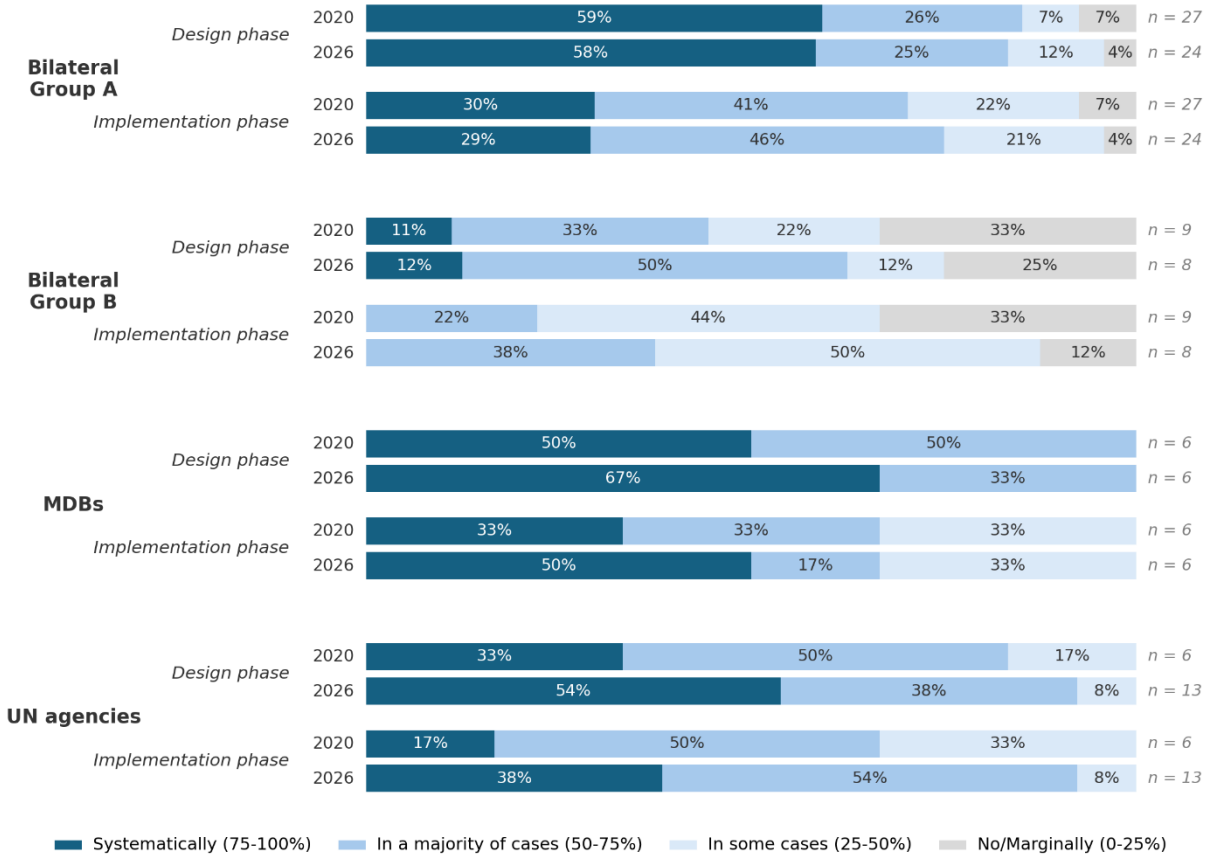


Note: Share of respondents selecting one of the top two response categories, i.e. "systematically" or "in a majority of cases" ("to a great extent" for items on an extent scale), unweighted and calculated separately for each provider type, among the 44 institutions that responded to both rounds. The percentage beside each arrow is the 2026 share. The thresholds separating a strong change, a slight change and a stable position are scaled to the number of respondents in each provider type, using the standard error of the change with a small-sample adjustment: a change counts as strong from one standard error and as slight from half of one.

Providers were asked if **managers at country level have sufficient flexibility at design phase** to tailor their approaches and processes to each specific context, and **at implementation phase**, to focus on achieving long-term outcomes and adjust project paths considering changing local contexts and results information. **Managers at country level are granted more flexibility at the design phase than at the implementation phase.** In 2026, 85% of the 52 respondents report frequent flexibility at design against 71% at implementation: this difference holds among bilateral providers and development banks (Figure 19). For example, a bilateral provider from Group A explains that it has been working on monitoring risks during implementation in recent years only, making it a less formalised process than assessing risks

during project design. It adds that while project outcome and impact objectives can be changed during implementation, it requires consultations with the administration and/or the political level.

Figure 19. Flexibility for managers at country level at design and implementation phase, 2020-2026



Adapting to context remains the strongest reported area of results management and practical constraints differ across providers. Among the 52 respondents, 87% report that quality context analysis informs frameworks systematically or in a majority of cases. Frequent flexibility is less widespread, at 85% during design and 71% during implementation, which directs attention to what happens after a project has been approved, when changing course affects budgets, agreements and expectations. Group A institutions generally report substantial discretion at country level, with 18 of 24 reporting frequent implementation flexibility. Group B providers report a narrower space. Five of eight report frequent flexibility during design and three during implementation. A Group A provider identifies the combination of constraints: "Guidance to staff encourages a contextually-driven and adaptive approach to both design and implementation. However, tools, systems and staff capacity challenges limit the extent to which it happens in practice. Nonetheless, there are pockets of good practice."

Adaptation ultimately depends on how decisions are taken. Managers need clarity on which changes fall under their authority, which require approval, and how results information enters that judgement. Improving this process does not imply weaker accountability; it makes accountability compatible with an evidence-based change of course. This links the strong reported performance on adapting to context to the finding that learning from results information is more routine at country level than at headquarters.

Table 3. GP2: mean score by provider type, institutions answering both rounds

Provider type	Institutions in both rounds	2020	2026
Bilateral providers/Group A	24	7.4	7.8
Bilateral providers/Group B	8	4.5	5.4
Multilateral development banks	6	8.1	8.4
United Nations agencies	6	6.8	7.8

Note: Mean of the principle's frequency items on the four-point scale, rescaled to 1-10 as in Figure 1, for the institutions that answered both rounds. Self-assessments in both years, except for the six institutions expert-assessed in 2026 (Annex A).

Guiding Principle 3: Enhance country ownership, mutual accountability and transparency

Figure 20. GP3 trajectory table by provider type, 2020-2026

Survey item	Group A	Group B	MDBs	UN agencies
Guiding Principle 3 – Country ownership				
Country-level results frameworks are aligned with partner countries' results frameworks, including at indicator level.	→ 54%	→ 25%	→ 50%	↗ **100%**
The organisation's planning cycles at country level are synchronised with partner country planning cycles.	→ 33%	→ 12%	↘ 67%	→ 67%
Country and project level results frameworks are developed in a participatory way, and inclusive involvement is sought in problem analysis, design, planning, monitoring, evaluation and use of results.	→ 58%	→ 50%	↗ 83%	↗ **100%**
Monitoring approaches maximise use of partner countries' monitoring and statistical systems.	→ 25%	→ 12%	→ 67%	→ 33%
Efforts are made to strengthen countries' statistical and monitoring capacities where needed.	↗ 42%	→ 25%	↘ 67%	→ 33%
Mutual accountability is actively pursued at country- and project-level with regular dialogue among partners.	↗ **83%**	↗ 50%	↗ 83%	→ 67%
Results frameworks, information and data are high quality, timely, and publicly available.	↗ **67%**	↗ 12%	→ 67%	↗ 50%

↗ Strong increase ↗ Slight increase → Stable ↘ Slight decrease ↘ Strong decrease

Note: Share of respondents selecting one of the top two response categories, i.e. "systematically" or "in a majority of cases" ("to a great extent" for items on an extent scale), unweighted and calculated separately for each provider type, among the 44 institutions that responded to both rounds. The percentage beside each arrow is the 2026 share. The thresholds separating a strong change, a slight change and a stable position are scaled to the number of respondents in each provider type, using the standard error of the change with a small-sample adjustment: a change counts as strong from one standard error and as slight from half of one.

Despite being among the lowest-scoring principles in 2020 and thus featuring wider room for improvement, **Guiding Principle 3 is the principle where providers have shown the smallest progress among the 44 institutions that answered both rounds (+0.37), with Guiding Principle 1 (+0.44) and Guiding Principle 4 (+0.45) close behind**.

Public availability of results frameworks, information and data has risen sharply, from 31% to 56% of respondents (+25 percentage points among the 44 in both rounds), but publication does not always extend to the reporting against the frameworks. A Group A provider describes improved publication arrangements while recognising that publishing frameworks does not always mean publishing the reporting against them. Another makes clear that a lower self-assessment can reflect a more demanding understanding of ownership: "Though we are scoring lower than the previous time we had to

fill out this survey, we are better aware of non-alignment and dominance of our own results frameworks vis-a-vis those of partner countries. We wish to see this changed in the coming years but still face institutional bottlenecks to achieve this."

Use of partner systems has not moved. Among the 52 respondents, 56% report aligning country-level frameworks with partner frameworks at indicator level systematically or in a majority of cases, and 37% report maximising the use of partner monitoring and statistical systems with the same frequency. Among the 43 institutions with answers in both rounds, alignment rose from 20 to 24 while frequent use of partner systems remained at 13. Among Group B providers, frequent alignment is reported by two of eight and use of partner monitoring systems by one of eight; one describes participation and monitoring organised around projects and the requests of partners, with weaker arrangements above that level. An MDB reports stronger alignment of its country frameworks with those of partners, and a UN agency describes more participatory country programming. The UN column in the trajectory table is especially sensitive to composition. Among the six agencies in both rounds, frequent use of partner monitoring systems is unchanged at two of six and synchronisation of planning cycles at four of six. A common language supports ownership when it makes country systems easier to use, rather than giving providers another way to describe their own results.

However, **participatory approaches appear increasingly embedded in programme design**. The share of institutions reporting that mutual accountability is actively pursued at country and project level through regular dialogue with partners in at least a majority of cases rose from 60% in 2020 to 75% in 2026. A UN agency reports that mutual accountability mechanisms have been reinforced through more regular, structured policy and results dialogues at country level. A Group A provider's partner country plans, published in 2025, were all consulted with partner governments, and activity-level results frameworks are routinely developed in a participatory manner and used as a basis for dialogue. Another provider does not have country-level results frameworks but lets its partners develop results frameworks for individual grants while providing input through dialogue. Comments also point to limits. For a Group A provider, increased dialogue does not consistently make use of results and evidence, and some aspects of its strategies are not shared or discussed with partners. Another provider reports that mutual accountability and dialogue with public authorities have diminished in some countries with complex contexts, such as civil war or coups d'état, while another notes that traditional results frameworks are difficult to apply in the fluid context of its support to Ukraine.

Table 4. GP3: mean score by provider type, institutions answering both rounds

Provider type	Institutions in both rounds	2020	2026
Bilateral providers/Group A	24	5.1	5.4
Bilateral providers/Group B	8	3.7	4.0
Multilateral development banks	6	6.5	6.6
United Nations agencies	6	6.0	6.9

Note: Mean of the principle's frequency items on the four-point scale, rescaled to 1-10 as in Figure 1, for the institutions that answered both rounds. Self-assessments in both years, except for the six institutions expert-assessed in 2026 (Annex A).

Guiding Principle 4: Maximise the use of results information for learning and decision-making

Figure 21. GP4 trajectory table by provider type, 2020-2026

Survey item	Group A	Group B	MDBs	UN agencies
Guiding Principle 4 - Use of results information for learning and decision-making				
Results information from M&E is used for learning and decision making at all levels.	→ 46%	↗ 25%	→ 50%	→ 33%
Qualitative information is collected to complement quantitative data.	↗ **79%**	→ 38%	↗ 83%	→ 67%
The purpose and use of results information are clearly specified to staff and partners.	↗ **79%**	↗ 25%	↗ 83%	↗ **83%**
At headquarter level, results information is used for accountability	→ 83%	↗ **62%**	→ 83%	↘ 67%
At headquarter level, results information is used for communicating	→ 67%	→ 25%	↗ 100%	↗ 83%
At headquarter level, results information is used for learning	↘ 38%	↗ 12%	↗ 83%	→ 33%
At headquarter level, results information is used for steering/decision-making	→ 46%	↗ 25%	↗ **67%**	→ 33%
At country level, results information is used for accountability	↗ **83%**	↘ 12%	→ 83%	→ 67%
At country level, results information is used for communicating	↗ 71%	↗ 38%	→ 83%	→ 100%
At country level, results information is used for learning	↗ 71%	↗ **38%**	↗ 83%	→ 50%
At country level, results information is used for steering/decision-making	→ 75%	→ 25%	↗ 67%	↗ 50%

↗ Strong increase ↗ Slight increase → Stable ↘ Slight decrease ↘ Strong decrease

Note: Share of respondents selecting one of the top two response categories, i.e. “systematically” or “in a majority of cases” (“to a great extent” for items on an extent scale), unweighted and calculated separately for each provider type, among the 44 institutions that responded to both rounds. The percentage beside each arrow is the 2026 share. The thresholds separating a strong change, a slight change and a stable position are scaled to the number of respondents in each provider type, using the standard error of the change with a small-sample adjustment: a change counts as strong from one standard error and as slight from half of one.

The purpose for which information is used remains the central distinction in results management. Guiding Principle 4 improved by 0.45 points among the 44 institutions that answered both rounds. Its overall score combines accountability, communication, learning and steering, and the complete use matrix in Figure 8 shows a sharper management problem than that average can reveal. Routine use for accountability reaches 79% at headquarters compared with 38% for learning, and steering and decision-making also reach more institutions at country level (63%) than at headquarters (46%). The distinction is not between institutions that use results and those that do not, but rather between purposes that have become routine and purposes for which use is still intermittent.

Results by provider type show that the constraint is not the same everywhere. Among the 24 Group A providers, routine learning is reported by 17 at country level and nine at headquarters, and the headquarters count fell from 12 to nine among the same institutions. One Group A provider describes a practical investment in the connection to implementation: "This shift has been reinforced by the progressive integration of MEL officers within a growing number of projects, embedding results expertise closer to implementation." For Group B, three of eight report routine learning at country level and one of eight at headquarters; one provider explains that its information is available mainly at project level, which limits its use for choices across the institution. An MDB describes using a revised corporate results framework for steering. These examples point to different needs, from aggregation across projects to a corporate routine for steering on the evidence already available.

The same institutions can improve how they specify the purpose of information without yet changing how consistently it is used. Among the 52 current respondents, 69% report clearly specifying purpose and use to staff and partners, and 71% report collecting qualitative information alongside quantitative data. Beyond having these inputs, what matters is connecting those inputs to identifiable

decisions. A portfolio review, for instance, needs a question, evidence that speaks to it and designated responsibility for the resulting decision. More reporting, including AI-assisted reporting, will not close the gap without arrangements for learning and response.

Table 5. GP4: mean score by provider type, institutions answering both rounds

Provider type	Institutions in both rounds	2020	2026
Bilateral providers/Group A	24	6.4	6.7
Bilateral providers/Group B	8	3.8	4.3
Multilateral development banks	6	6.8	7.7
United Nations agencies	6	5.9	6.4

Note: Mean of the principle's frequency items on the four-point scale, rescaled to 1-10 as in Figure 1, for the institutions that answered both rounds. Self-assessments in both years, except for the six institutions expert-assessed in 2026 (Annex A).

Guiding Principle 5: Foster a culture of results and learning

Figure 22. GP5 trajectory table by provider type, 2020-2026

Survey item	Group A	Group B	MDBs	UN agencies
Guiding Principle 5 – Culture of results and learning				
Leadership communicates the purpose of RBM and uses results information for overall programme management and policy influencing at a political level.	↗ 46%	↗ 38%	↗ 67%	→ 67%
Leadership promotes a culture of results and learning by providing appropriate guidance. [2]	↗ **54%**	↗ 25%	↗ 67%	↗ 50%
Leadership promotes a culture of results and learning by providing tailored incentives and resources for managing for sustainable development results. [2]	↗ **54%**	→ 12%	↗ 67%	↘ 17%
All relevant staff are trained in RBM approaches and methods.	↗ 62%	↗ 12%	↗ 67%	→ 50%
There are room and mechanisms for peer learning on successes and failures.	↗ 58%	↗ 50%	↗ 67%	→ 33%
The organisation values careful experimentation in support of innovation as a way to foster better results.	↗ **42%**	↗ 25%	→ 50%	↘ 67%

↗ Strong increase ↗ Slight increase → Stable ↘ Slight decrease ↘ Strong decrease

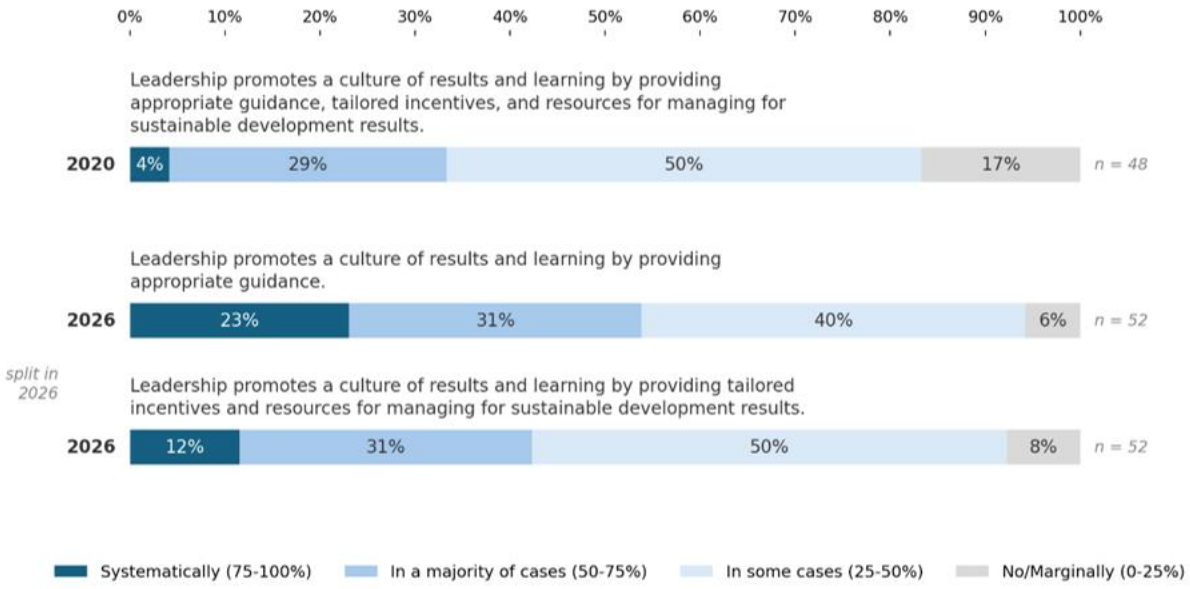
Note: Share of respondents selecting one of the top two response categories, i.e. "systematically" or "in a majority of cases" ("to a great extent" for items on an extent scale), unweighted and calculated separately for each provider type, among the 44 institutions that responded to both rounds. The percentage beside each arrow is the 2026 share. The thresholds separating a strong change, a slight change and a stable position are scaled to the number of respondents in each provider type, using the standard error of the change with a small-sample adjustment: a change counts as strong from one standard error and as slight from half of one. In 2020, the survey asked one item for guidance and for incentives; that single answer serves as the baseline for both rows.

Guiding Principle 5 does not rank among the strongest areas of alignment, but **it records the largest advance of the six Guiding Principles between the two rounds**, rising by 0.79 points on the ten-point scale among the 44 institutions that answered both rounds. Respondents point to identifiable investments made since the baseline. MDBs describe new corporate functions dedicated to outcomes, knowledge and innovation, an updated corporate results framework accompanied by revised guidance on country and project-level frameworks, and a first institutional impact report. UN agencies report a dedicated results-based management hub with expanded training, a rights and results-based management curriculum deployed organisation-wide and a roadmap under a newly appointed executive head. Among bilateral providers, comments describe agency-specific training and guidance material, a corporate training

programme developed since 2021, a new results model accompanied by capacity building and staff training scheduled alongside the development of an overarching logical framework.

Leadership is providing guidance on results-based management faster than it is providing incentives and resources. The three elements combined in a single question in the 2020 round – guidance, incentives and resources – do not necessarily move together within an institution. Hence, the 2026 questionnaire separates leadership guidance from the incentives and resources made available to staff (Figure 23). The distinction proves substantive: among the 52 institutions, leadership provides appropriate guidance systematically or in a majority of cases for 54%, against 42% for tailored incentives and resources. Thirteen institutions rate guidance above incentives and resources and only two rate it below, with seven reporting that guidance is provided in a majority of cases or systematically while incentives and resources remain at some cases or less. Because the baseline round recorded a single combined response, this finding describes the 2026 position rather than a change over time.

Figure 23. Leadership support for a culture of results and learning, 2020-2026



Note: Guidance, incentives and resources were one single question in 2020, which serves as a baseline for both questions in 2026.

This confirms at item level the divergence between formal and tacit conditions discussed earlier in this report. The building blocks point the same way: across all respondents, dedicated resources is the only block whose average score falls between the two rounds (Figure 5), and incentive structure remains the lowest-scoring block of all (as discussed in Chapter 1). Comments articulate the same sequencing. A UN agency observes that the institution has "the correct principles and excellent examples of good practice", while overall implementation remains uneven "due to scarcity of incentives and competing priorities". Another reports that corporate guidance and tools have been progressively strengthened and better integrated into planning systems, while sustaining a learning culture "remains dependent on managerial behaviours, incentives, and time allocation", particularly following reductions to the organisation's financial capacity in 2025. A Group A bilateral provider notes that results-based management guidelines were updated in 2023 while personnel resources have been reduced. A second describes an institution that values careful experimentation in principle but that the practice "requires skills and resource".

Leadership engagement is advancing, but staff time and capacity remain the operative constraint. Providers report overall progress in leadership communicating the purpose of results-based management and using results information for programme management and policy influencing. An MDB reports that impact culture and buy-in "has definitely improved among staff, senior management and the board since 2021". A Group A bilateral provider states that ownership is strong at all levels of leadership, including close monitoring of an improved focus on results. A Group B provider notes that despite being at an early stage of developing a systematic approach, leadership nonetheless appreciates the importance and role of results-based management. A UN agency attributes part of its progress to more consistent leadership messaging on results-based management and its use for both programme management and external engagement.

What constrains progress is the capacity to act on that engagement, and provision differs sharply across provider types. Routine training in results-based management is noted in 15 of the 24 Group A providers, nine of the 13 UN agencies and four of the six MDBs, but in only one of the eight Group B providers. Sixteen of the 23 comments related to this principle concern staff numbers, skills, turnover or competing priorities, across all four provider types. A Group B bilateral provider notes that regular staff rotation means results systems and know-how require continuous support. A Group A provider describes human resources that are insufficient in numbers and competence, high turnover between government areas, and rigid public administration rules on performance evaluation and careers as structural obstacles the agency "cannot solve alone". Another Group B provider reports an established results-based planning framework that is not yet fully implemented in practice, citing the need to strengthen professional capacities and update the planning toolkit. Even where provision is strongest the gap persists: among the six UN agencies that answered both rounds, the number reporting routine incentives and resources fell from two to one.

The advances recorded under this principle indicate that a results culture is being consolidated, but not necessarily that it is geared towards learning. This parallels the evidence under Guiding Principle 4, which suggests that consolidation is occurring around accountability rather than learning. Comments from organisations that rated accountability above learning in uses of results information describe this asymmetry directly. A Group A bilateral provider states that learning "is surely promoted", while the use of results information within that learning "is hardly practiced", with evaluations by the independent evaluation unit serving as the principal learning moments. An MDB reporting improved buy-in at all levels adds that operationalisation and feedback loops remain a challenge. A Group B provider describes an existing results-based management culture that is not yet fully implemented in practice, and a UN agency contrasts sound principles with uneven implementation. Experimentation is the point at which a commitment to learning becomes costly: it is the lowest-scoring item under this principle and asks something different of an organisation than other items do. Communicating the purpose of results-based management, issuing guidance and delivering training are decisions leadership can take and implement centrally, while valuing experimentation means accepting in advance that some interventions will not work, which is harder to reconcile with the accountability requirements institutions face.

Table 6. GP5: mean score by provider type, institutions answering both rounds

Provider type	Institutions in both rounds	2020	2026
Bilateral providers/Group A	24	5.0	5.8
Bilateral providers/Group B	8	3.8	4.2
Multilateral development banks	6	5.2	6.6
United Nations agencies	6	5.2	5.8

Note: Mean of the principle's frequency items on the four-point scale, rescaled to 1-10 as in Figure 1, for the institutions that answered both rounds. Self-assessments in both years, except for the six institutions expert-assessed in 2026 (Annex A). The 2020 questionnaire asked one item for guidance and for incentives; its answer is used as the baseline for each of the two 2026 items.

Guiding Principle 6: Develop a results system that is manageable and reliable

Figure 24. GP6 trajectory table by provider type, 2020-2026

Survey item	Group A	Group B	MDBs	UN agencies
Guiding Principle 6 – Manageable and reliable results systems				
Monitoring and evaluation systems are adequately resourced to meet the needs of the organisation.	↗ 46%	↘ 25%	↗ 67%	→ 33%
Monitoring and evaluation systems are set up to produce credible quantitative and qualitative data and evidence that meets the needs and capacities of the development organisation and its partners.	→ 67%	→ 38%	↗ 67%	↗ 67%
Results based management is supported by adequate information systems, able to manage quantitative and qualitative results information.	→ 38%	→ 0%	↗ 83%	→ 50%
The overall RBM system is user friendly, lean and reviewed regularly to fit its purpose.	↗ **50%**	↘ 0%	↗ **100%**	↗ 50%

↗ Strong increase ↗ Slight increase → Stable ↘ Slight decrease ↘ Strong decrease

Note: Share of respondents selecting one of the top two response categories, i.e. "systematically" or "in a majority of cases" ("to a great extent" for items on an extent scale), unweighted and calculated separately for each provider type, among the 44 institutions that responded to both rounds. The percentage beside each arrow is the 2026 share. The thresholds separating a strong change, a slight change and a stable position are scaled to the number of respondents in each provider type, using the standard error of the change with a small-sample adjustment: a change counts as strong from one standard error and as slight from half of one.

Manageability and reliability remain the lowest-scoring principle, at 5.5 on the ten-point scale among the 52 institutions, although the principle improved by 0.66 points among the 44 institutions that answered both rounds. A system can become more elaborate without becoming easier to use. Among the 52 institutions, for 62% monitoring and evaluation systems produce credible evidence systematically or in a majority of cases, but fewer (42%) have adequate information systems with the same frequency, and 42% have adequately resourced monitoring and evaluation. A system may therefore generate useful evidence while struggling to organise, maintain or use it consistently. The credibility item is also a self-assessment, not an external assessment of the information produced. The share reporting a system that is user-friendly, lean and regularly reviewed is 50% among the 52 institutions. On the matched set of 44, this rose from 34% to 48%.

Provider types face different priorities. Among the 24 Group A providers, 12 have a routinely manageable system but nine have adequate information systems. One describes systems that have improved while still serving learning at individual activity level more readily than learning across the portfolio. For the eight Group B providers, none reports adequate information systems or a lean, user-friendly system in a majority of cases or systematically. One describes project and finance tools that would need adapting to support an institution-wide results approach. All six MDBs report routinely manageable systems, but the category still contains gaps in resourcing and information management. Among the 13 UN agencies, five report adequately resourced monitoring and evaluation, and the six continuing agencies show almost no change in their mean for this principle. Their apparent improvement in the full sample reflects the additional institutions answering in the later round.

The comments provide a practical distinction between a system arriving and a system working. Respondents describe new platforms, systems under construction and efforts to make existing arrangements easier to use. A UN agency describes collecting feedback from users: "These feedback mechanisms have directly informed iterative improvements aimed at increasing user-friendliness and

reducing complexity." Such feedback makes manageability an ongoing task; it can reveal redundant requirements, difficult interfaces and evidence that is collected but rarely used. The relevant test is whether staff and partners can obtain credible information for the decisions they face, at a cost they can sustain. Additional analytical capability, including AI, will be useful where it reduces a real burden or improves a decision process; adoption alone does not show that either has happened.

Table 7. GP6: mean score by provider type, institutions answering both rounds

Provider type	Institutions in both rounds	2020	2026
Bilateral providers/Group A	24	4.7	5.7
Bilateral providers/Group B	8	3.4	3.4
Multilateral development banks	6	5.8	6.6
United Nations agencies	6	4.9	5.0

Note: Mean of the principle's frequency items on the four-point scale, rescaled to 1-10 as in Figure 1, for the institutions that answered both rounds. Self-assessments in both years, except for the six institutions expert-assessed in 2026 (Annex A).

References

Guerrero-Ruiz, A., J. Schnatz and C. Verger (2021), *A Baseline Survey of the Guiding Principles on Managing for Sustainable Development Results*, OECD Publishing, https://doi.org/10.1787/f8974d89-en. [2]

OECD (2026), *AI adoption in development co-operation systems: from experimentation to institutional design*, OECD Publishing. [7]

OECD (2024), *Effective Results Frameworks for Sustainable Development: Achieving Impact by Design*, OECD Publishing, Paris, https://doi.org/10.1787/aaa4bb52-en. [3]

OECD (2023), *Achieving SDG Results in Development Co-operation – A Comparative Assessment [DCD(2023)29]*, Development Co-operation Directorate, https://one.oecd.org/document/DCD(2023)29/en/pdf. [5]

OECD (2021), *Achieving SDG Results in Development Co-operation: Summary for Policy Makers*, OECD Publishing, https://doi.org/10.1787/5b2b0ee8-en. [8]

OECD (2019), *OECD DAC The Guiding Principles on Managing for Sustainable Development Results (MfSDR)*, https://doi.org/10.1787/44a288bc-en. [1]

OECD (2019), *Sustainable Results in Development: Using the SDGs for Shared Results and Impact*, OECD Publishing, https://doi.org/10.1787/368cf8b4-en. [9]

UN (2026), *SDG Indicators*, https://unstats.un.org/sdgs/indicators/indicators-list/. [4]

UN (2024), *The Pact for the Future [A/RES/79/1]*, https://www.un.org/sites/un2.un.org/files/sotf-pact_for_the_future_adopted.pdf. [6]

Annex A. Survey methodology

Process

Between April and July 2026, the OECD conducted a survey to assess institutional progress against the OECD DAC Guiding Principles on Managing for Sustainable Development Results, following a baseline survey conducted in 2020[2].

The 2026 survey reproduces the questions of the baseline survey, with one item split in two (leadership guidance, and leadership incentives and resources) and three new questions (one on the role of the SDG framework and two on artificial intelligence). I, with questions focusing on the key elements of each principle. The survey concludes with a question on the overall strengths and constraints of the providers' ability to apply the Guiding Principles, referred to as the building blocks.

The 2026 survey was launched on 2 April 2026 and circulated in successive waves as institutional focal points were identified. A brief containing preliminary findings was circulated at the **OECD Conference on the Future of Development Co-operation** on 11-12 May 2026. Each responding institution received a **confidential development capability profile** visualising how it compares with peers on institutional dimensions covered by the self-assessment. In June 2026, the OECD convened participants for **two joint reflection sessions** for an exchange on the preliminary findings of the survey under the Chatham House rules.

Participation sample

The dataset for the 2026 assessment round comprises 52 institutions. Among these, 46 organisations provided self-assessments directly through the survey. In one case, responses were calibrated based on practice and documentation to ensure consistency with the evidence available and comparability across respondents. Additionally, six institutions that participated in 2020 were included in the 2026 round through expert assessments. The calibration and expert assessments were based on available data, practices and documentation.

As in the baseline round, responses were disaggregated by provider group, accounting for different organisational and institutional realities across providers. Four categories of respondents were identified:

- **Bilateral providers/Group A:** larger, longer-standing DAC members that started reporting official development assistance (ODA) statistics to the OECD before 2010;
- **Bilateral providers/Group B:** smaller or newer DAC members and providers outside the DAC;
- **Multilateral development banks (MDBs)**; and
- **United Nations (UN) agencies**.

[2] For more details on the baseline survey, see the report "A Baseline Survey of the Guiding Principles on Managing for Sustainable Development Results" (Guerrero-Ruiz, Schnatz and Verger, 2021[2]).

One regional financial institution also provided responses to the survey. It is included in all figures and statistics for the full sample, but not in those disaggregated by provider group.

Table 8. Participation by provider type8 sets out participation numbers by provider group across the two rounds, including the number of institutions assessed in both rounds. The list of institutions can be found in Annex B.

Table 8. Participation by provider type

	Baseline round	2026 round	Both rounds (matched sample)
Bilateral providers/Group A	27	24	24
Bilateral providers/Group B	9	8	8
Multilateral development Banks	6	6	6
United Nations agencies	6	13	6
Regional financial Institution	0	1	0
Total	**48**	**52**	**44**

Note: The matched sample represents the number of institutions that responded to both rounds of the survey. The baseline column counts the 48 baseline respondents. Four baseline respondents did not take part in 2026, and eight institutions joined, seven of them UN agencies. The 44 institutions in both rounds are the basis for the matched comparisons.

Methodological considerations

Coding and Guiding Principle means: Frequency answers were scaled from “no or marginally (0-25%)” = 1, “in some cases (25-50%)” = 2, “in a majority of cases (50-75%)” = 3 and “systematically (75-100%)” = 4, and extent answers as “to some extent” = 2 and “to a great extent” = 3. Missing answers remain missing. The items on the role of the SDG framework, on AI use and application, and the comments are not included in the principle means. A principle mean averages all item responses of the institutions concerned, so that each answered item carries the same weight. For some analyses, the 1-4 scale was transformed into a 1-10 scale. On Guiding Principle 5, the 2020 questionnaire combined guidance, incentives and resources in one question. The 2026 survey separated those. When comparing both rounds, the single 2020 answer is used as the baseline for each of the two 2026 items.

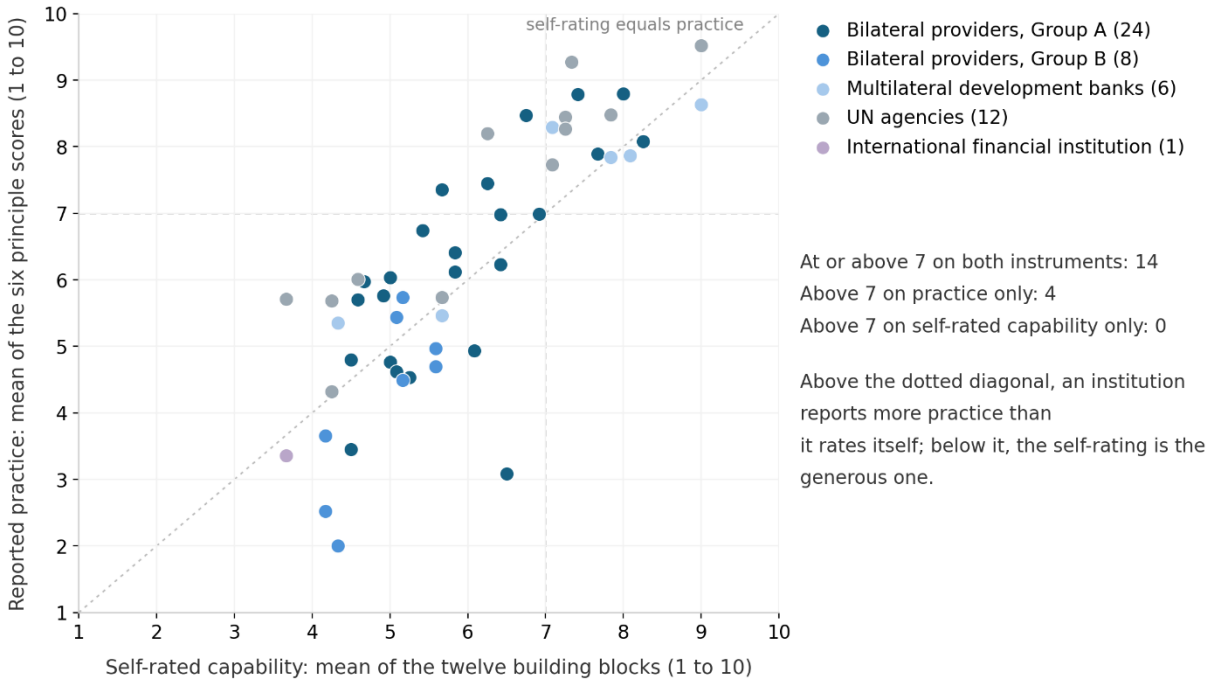
Building blocks, groupings and positions: The 12 building blocks use a 1-10 scale, with 7 and above considered the “good enough” threshold and 4 and below “insufficient”. Capability averages available blocks within an institution, then averages institutions. Forty-four respondents rated every block themselves; with the six expert assessments and the calibrated assessment, 51 institutions have ratings for all 12 blocks.

Calibration and expert assessments (seven institutions): Assessments started from the institutions’ survey responses and were then reviewed against documentary evidence, including OECD Development Co-operation Profiles, Multilateral Organisation Performance Assessment Network (MOPAN) assessments, evaluation reports, organisational strategies and other publicly available sources. Scores were updated only where there was credible evidence of changes in institutional practices or capacities. A key principle was to distinguish between organisational outputs (e.g. new strategies, policies, units or reporting systems) and outcomes (demonstrable changes in behaviour, decision-making or use of results). Where the initial self-assessments appeared inflated, expert judgement was used to recalibrate scores based on external evidence. All assessments were subjected to consistency checks and reviewed against comparable institutions to ensure plausibility and comparability across the sample.

Reading Guiding Principle scores and institutional capability together: The survey measures institutions in two ways: it asks how often they do specific things (i.e. the 46 items behind the six Guiding

Principles) and it asks them to rate their own institutional capability (i.e. the 12 building blocks). The two agree closely across the institutions, but not in the same way for everyone. On the one hand, institutions with strong practice tend to rate their capability below what their practice suggests; on the other hand, institutions with weak practice tend to rate it above. The newer and smaller bilateral providers are the only group whose self-rated capability exceeds its reported practice. Therefore, the results are more useful for diagnosing strengths and weaknesses than for producing a league table of institutional performance.

Figure 25. Capability against reported practice, 51 institutions, 2026



Note: Each dot is one institution: its self-rated capability on the 12 building blocks against its reported practice on the 46 survey items, 2026. Practice is the mean of the six principle scores on the ten-point scale; the frequency scales are converted to that scale, and the item on corporate results frameworks, which uses a yes / under development / no scale, is mapped onto it with "under development" at the midpoint. The item on the role of the SDG framework and the two items on artificial intelligence are excluded, as they do not measure practice against the Guiding Principles. Fifty-one of the 52 respondents are shown; one institution did not rate all 12 building blocks. Dotted lines at 7, the survey's "good enough" point; the diagonal marks equal values on the two instruments. The values of seven institutions are not self-reported: six were expert-assessed and one was calibrated (Annex A).

Size of programme (Figure 7)**:** Country ODA (grant equivalents, average of 2023-2025) is available for 28 of the 29 bilateral providers (32 respondents, with institutions from the same country counted once). Its rank correlation with the practice score, the mean of the six principle scores, is 0.43 across the 28 and −0.07 among the 21 Group A providers alone; with the two group means removed it is negative (−0.22). Three tiers cut at USD 1 billion and 5 billion give mean practice scores of 4.8 on ten providers, 5.6 on nine and 6.2 on nine, and the largest tier spans 4.5 to 8.8. By principle, the association is strongest for a manageable and reliable results system (0.47) and for supporting the SDGs (0.36), and weakest for country ownership (0.16).

Questions

Table 9. List of survey questions

Guiding Principle	Question
1	Results frameworks are developed at the following levels:
1	At Corporate Level
1	At Sector/Thematic Level
1	At Partner country Level
1	At project Level
1	To what extent do your results frameworks incorporate SDG results indicators?
1	At Corporate Level
1	At Sector/Thematic Level
1	At Partner country Level
1	At project Level
1	Established methodologies (e.g. logical framework, theory of change) are used to define the results chain at the following levels:
1	At Corporate Level
1	At Sector/Thematic Level
1	At Partner country Level
1	At project Level
1	Results Mainstreaming and Coherence:
1	The design of the organisational structure is geared towards achieving development results.
1	In line with the integrated approach of the SDGs, cross-sector approaches are promoted and included into measurement, learning and reporting systems.
1	Since 2021, the strategic environment for development co-operation has evolved. Which of the following best describes how this has affected the role of the SDG indicator framework in your organisation's results management?
1	Please comment here for any elaboration on the Principle 1 related questions and provide complementary information on any types of results frameworks that are not listed here (e.g. results frameworks designed for partnerships):
2	Thinking about your development co-operation portfolio in partner countries, which of the following practices are you adopting in order to adapt to context and be flexible?
2	Quality context analysis enabling a sound understanding of local dynamics and needs inform country- and project-level results frameworks.
2	The conditions for success (and/or assumptions) and risk factors in achieving the intended outcomes are identified and are monitored during implementation.
2	Managers at country level have sufficient flexibility to tailor their approaches and processes to each specific context (design phase).
2	Managers at country level focus on achieving long term outcomes, and are empowered to adjust the path of a project in light of changing local contexts and results information to ensure the expected outcomes are met (implementation phase).
2	Please comment here for any elaboration (and/or significant changes in your organisation's approach to this principle since 2020):
3	To enhance country ownership, mutual accountability and transparency:
3	Country-level results frameworks are aligned with partner countries' results frameworks, including at indicator level.
3	The organisation's planning cycles at country level are synchronised with partner country planning cycles.
3	Country and project level results frameworks are developed in a participatory way, and inclusive involvement is sought in problem analysis, design, planning, monitoring, evaluation and use of results.
3	Monitoring approaches maximise use of partner countries' monitoring and statistical systems.
3	Efforts are made to strengthen countries' statistical and monitoring capacities where needed.
3	Mutual accountability is actively pursued at country- and project-level with regular dialogue among partners.
3	Results frameworks, information and data are high quality, timely, and publicly available.
3	Please comment here for any elaboration (and/or significant changes in your organisation's approach to this principle since 2020):
4	In maximising the use of results information for learning and decision making:
4	Results information from M&E is used for learning and decision making at all levels.
4	Qualitative information is collected to complement quantitative data.
4	The purpose and use of results information are clearly specified to staff and partners.
4	At headquarters level, results information is used for:

4	Accountability
4	Communicating
4	Learning
4	Steering/Decision Making
4	At country level, results information is used for:
4	Accountability
4	Communicating
4	Learning
4	Steering/Decision Making
4	Please comment here for any elaboration (and/or significant changes in your organisation's approach to this principle since 2020):
5	To foster a culture of results and learning:
5	Leadership communicates the purpose of RBM and uses results information for overall programme management and policy influencing at a political level.
5	Leadership promotes a culture of results and learning by providing appropriate guidance.
5	Leadership promotes a culture of results and learning by providing tailored incentives and resources for managing for sustainable development results.
5	All relevant staff are trained in RBM approaches and methods.
5	There are room and mechanisms for peer learning on successes and failures.
5	The organisation values careful experimentation in support of innovation as a way to foster better results.
5	Please comment here for any elaboration (and/or significant changes in your organisation's approach to this principle since 2020):
6	To develop a results system that is manageable and reliable:
6	Monitoring and evaluation systems are adequately resourced to meet the needs of the organisation.
6	Monitoring and evaluation systems are set up to produce credible quantitative and qualitative data and evidence that meets the needs and capacities of the development organisation and its partners.
6	Results based management is supported by adequate information systems, able to manage quantitative and qualitative results information.
6	The overall RBM system is user friendly, lean and reviewed regularly to fit its purpose.
6	Please comment here for any elaboration (and/or significant changes in your organisation's approach to this principle since 2020):
	In reflecting upon the strengths and constraints for developing further your RBM system, where do you situate your organisation in the following areas? Rate from 1 "Major Constraints" to 10 "Major Strength". As reference points, 7 and above is considered good enough and 4 and below is considered insufficient.
	Policy/Policies
	Organisational structure
	Internal organisational culture
	Dedicated resources
	Adequate skills & training
	Internal processes
	Incentive structure
	Quality of results methodologies
	Quality assurance mechanisms
	Information technology systems (IT)
	Formal demand for results information reporting
	Actual use of results information
	Which of the following best describes your organisation's current institutional approach to using artificial intelligence (AI) in its management processes?
	In which management functions is your organisation currently using artificial intelligence (AI)? (Select all that apply)

Annex B. Survey respondents

Table 10. List of survey respondents, 2020 and 2026

Bilateral providers/Group A	Bilateral providers/Group B
Australia	Czech Republic
Austria	Estonia
Belgium (Enabel)	Iceland
Belgium (MFA)	Latvia
Canada	Lithuania
Denmark	Mexico
European Commission	Poland
Finland	Slovak Republic
France	Slovenia
Germany (BMZ)	**Multilateral development banks**
Greece	African Development Bank
Ireland	Asian Development Bank
Italy	European Bank for Reconstruction and Development
Korea	Inter-American Development Bank
Luxembourg (LuxDev)	Islamic Development Bank
Luxembourg (MFA)	World Bank Group
Netherlands	**United Nations agencies**
New Zealand	Food and Agriculture Organization (FAO)
Norway	International Labour Organization (ILO)*
Portugal	Office of the United Nations High Commissioner for Human Rights (OHCHR)*
Spain	Pan American Health Organization (PAHO)*
Sweden (MFA)	United Nations Office for the Coordination of Humanitarian Affairs (UN OCHA)*
Sweden (Sida)	United Nations Entity for Gender Equality and the Empowerment of Women (UN Women)*
Switzerland (SDC)	United Nations Development Programme (UNDP)
Switzerland (SECO)	United Nations Educational, Scientific and Cultural Organization (UNESCO)
United Kingdom	United Nations Children's Fund (UNICEF)
United States	United Nations Industrial Development Organization (UNIDO)
	United Nations Office on Drugs and Crime (UNODC)*
	World Food Programme (WFP)
	World Health Organization (WHO)*
	Regional financial institution
	Arab Fund for Economic and Social Development*

Note: Institutions marked by (*) did not participate in the 2020 round of the survey.