



Arab Observatory for Policy Evaluation Institutionalization

ANNUAL REPORT 2025

Index of Evaluation Institutionalization in the Arab Region

Annual Report 2025: the institutionalisation of public policy evaluation in
eleven Arab League states

Dr Anis Ben Brik



2025

Annual report 2025

Index of Evaluation Institutionalization in the Arab Region

Annual Report 2025: the institutionalisation of public policy evaluation in eleven Arab League states

Dr Anis Ben Brik

Index of Evaluation Institutionalization in the Arab Region. Annual Report 2025: the institutionalisation of public policy evaluation in eleven Arab League states.

Please cite this publication as: Ben Brik, A. (2025). Index of Evaluation Institutionalization in the Arab Region: Annual Report 2025. Arab Observatory for Policy Evaluation Institutionalization, <https://mena-evaluation.com/report/>.

The designations employed do not imply the expression of any opinion concerning the legal status of any country or territory. Scores describe institutional arrangements coded from documentary evidence; they do not measure the performance of governments or the quality of individual evaluations. The views expressed are those of the author.

© Anis Ben Brik 2025. Published under the Creative Commons Attribution 4.0 International licence (CC BY 4.0). Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

Contents

Foreword	6
Acknowledgements	7
Executive summary	8
The report at a glance	9
1. Introduction: evaluation as administrative capacity	10
2. Conceptual framework	12
3. Data and method	15
4. The composite index	17
5. The political system: anchoring and use	20
6. The social system: diffusion and demand	23
7. The profession system: professionalisation	25
8. Configurations across systems	27
9. Institutional architecture and gaps	35
10. Supply, demand and the administrative capacity pathway	40
11. Exploratory configurations	44
12. State profiles	47
13. Discussion: decoupling, supply and demand	50
14. Implications for evaluation policy and public administration	52
15. Limitations and future editions	54
16. Conclusion	56
Annex A. Indicators	58
Annex B. Coded scores by state	59
References	60
Figures	
Figure 4.1. Composite index of evaluation institutionalisation, by state	18
Figure 5.1. Political system: institutional anchoring (1A) and evaluation use (1B)	21
Figure 5.2. Political system indicators, by state	22
Figure 6.1. Social system indicators, by state	24
Figure 7.1. Profession system indicators, by state	26
Figure 8.1. Political, social and profession system scores, by state	28
Figure 8.2. Decomposition of the composite index into sub-dimension contributions	29
Figure 8.3. Decoupling between political and social systems (percentage points)	30
Figure 8.4. Political anchoring and professionalisation	31
Figure 8.5. All 16 indicators, by state	32
Figure 8.6. Diffusion of each indicator across states	33
Figure 9.1. Institutional configuration: share of each system in the state total	37
Figure 9.2. Where institutionalisation is uneven: gaps and leading system	39
Figure 10.1. Evaluation supply and demand	42

Figure 10.2. The administrative capacity pathway	43
Figure 11.1. Exploratory configurations: Ward clustering of the sixteen indicators	45

Tables

Table A.1. Annex A. Indicators	58
Table 2.1. The sixteen indicators	14
Table 4.1. Composite index and sub-dimension scores, by state (per cent)	19
Table 8.1. Mean scores by region (per cent)	33
Table 8.2. Indicator statistics: mean, dispersion and diffusion	34
Table 9.1. Pairwise gaps between systems (percentage points)	38
Table 11.1. Mean system scores by group	46
Table 14.1. Recommendations by actor	53
Table B.1. Annex B. Coded scores by state	59

Foreword

Public policy evaluation has become part of the vocabulary of administrative reform across the Arab region. Development plans, performance frameworks and donor agreements refer to it, and many ministries now house monitoring and evaluation units. Whether evaluation has become an institution, with rules that oblige it, organisations that carry it, a profession that sustains it and publics that demand it, is a different question, and it has rarely been measured.

This report measures it. It scores eleven Arab League states on sixteen indicators organised in three systems, political, social and professional, and reads the resulting configurations as evidence about how evaluation is anchored in public administration. The report does not rank governments. It identifies where institutionalisation is concentrated, where it is absent, and what each configuration implies for evaluation policy.

Dr Anis Ben Brik

Acknowledgements

This report was written by Dr Anis Ben Brik, with Hala Al Attar as project coordinator. The index applies the institutionalisation of evaluation framework of the Evaluation GLOBE; Prof. Dr. Reinhard Stockmann (Centre for Evaluation, Saarland University) granted permission to use the framework and reviewed the monograph whose country assessments and scores the index draws upon. The country assessments rest on the contributions of the country-chapter authors of the Edward Elgar and Policy Press volumes. Responsibility for any error rests with the author.

Executive summary

Evaluation is present in the organisation of Arab states but absent from their law, their oversight institutions and, in most states, their public sphere. Across eleven Arab League states the composite index averages 16.6% of its theoretical maximum and reaches 25.6% at most (Morocco). No state has adopted a national evaluation law or decree.

Key messages

- Institutionalisation is concentrated in the executive and the profession. The profession system averages 24.4% and the political system 18.6%; the social system averages 4.5% and is recorded in two states only.
- Organisation precedes mandate. Organisational embedding exceeds formal anchoring in 9 of 11 states: evaluation units exist where binding rules do not.
- Practice does not reach oversight. Routine evaluation exceeds its use in audit, parliamentary or civic arenas in 11 of 11 states; the pathway average falls from 38.6% for practice to 2.3% for use.
- Nine of the eleven states lie on the political–profession edge of the configuration: no societal system is recorded in them. Composite scores are close; the architecture behind them differs.
- Exploratory clustering distinguishes three configurations: a professionalised outlier (Morocco), four Gulf states with organised executive systems, and six states with thin institutional bases.

Way forward

Chapter 13 sets out reform directions: legal anchoring of evaluation obligations; a role for supreme audit institutions and parliaments in reviewing evaluation evidence; publication of evaluation reports to create societal demand; and the consolidation of the profession through education, associations and standards.

The report at a glance

11 Arab League states coded: six in the Gulf, two in the Levant, three in North Africa	16 indicators in four sub- dimensions and three systems	16.6% regional mean composite, per cent of the theoretical maximum	25.6% highest composite score (Morocco)
0 states with a national evaluation law or decree	9 of 11 states where organisational embedding exceeds formal anchoring	9 of 11 states with no societal system recorded	33 of 55 state pairs grouped alike by both clustering bases

Reader guide

Each indicator is coded on a five-point scale (0, 0.25, 0.5, 0.75, 1) from documentary evidence. Sub-dimension and system scores are means of their indicators, expressed as a percentage of the theoretical maximum; the composite is the unweighted mean of four sub-dimensions. Scores describe institutional arrangements, not the performance of governments. Differences of a few points between states should not be over-interpreted.

I

Introduction: evaluation as administrative capacity

This chapter sets out why the institutionalisation of evaluation is a question of administrative capacity, and what the index adds to existing assessments.

I. Introduction: evaluation as administrative capacity

Key messages

- Evaluation is treated as a component of administrative capacity, not as a technical exercise.
- The index measures institutions, not the quality of individual evaluations.
- It covers eleven Arab League states in its first edition.

Evaluation is a core component of administrative capacity. It is the institutional means by which governments generate systematic evidence on the relevance, effectiveness and efficiency of public interventions, and by which that evidence enters budgeting, policy design and accountability. Where evaluation is institutionalised, it becomes a routine of the state rather than an episodic project financed by development partners.

Comparative research has long documented the uneven institutionalisation of evaluation across countries (Furubo et al., 2002; Jacob et al., 2015). The Arab region, however, has remained largely outside these comparisons. The EvalMENA Index addresses this gap. It asks three questions: how far evaluation is institutionalised in the political, social and professional systems of Arab League states; which components of institutionalisation account for the differences between states; and what the resulting configurations imply for evaluation policy and public administration reform.

II

Conceptual framework

This chapter presents the three systems of institutionalisation and the sixteen indicators that measure them.

2. Conceptual framework

Key messages

- Institutionalisation is observed in the political, social and professional systems.
- Sixteen indicators are grouped in four sub-dimensions.
- The framework follows the Evaluation GLOBE, used with permission.

The index conceives institutionalisation as the process by which evaluation acquires rules, organisations, routines and shared expectations that give it stability beyond individual projects and actors. Organisational institutionalism reminds us that formal structures may be adopted for their legitimacy rather than their technical efficiency, and may therefore remain decoupled from practice (Meyer & Rowan, 1977). Isomorphic pressures, whether coercive, mimetic or normative, can spread evaluation units and policies without the capabilities that make them effective (DiMaggio & Powell, 1983; Pritchett et al., 2013).

Following the Evaluation GLOBE project (Stockmann et al., 2020), institutionalisation is observed in three systems. In the political system, evaluation is anchored in laws, decrees, organisational units and the audit function, and is used by parliament and across policy sectors. In the social system, it is known, discussed, published and demanded by citizens and civil society. In the profession system, it rests on academic education, professional communication, voluntary organisations for professional evaluation (VOPEs) and codified norms. Evaluation systems in this sense are configurations of supply and demand for evaluative knowledge (Leeuw & Furubo, 2008), and their use depends on the organisational and political conditions in which findings are received (Weiss, 1998).

Table 2.1. The sixteen indicators

Code	Indicator	Sub-dimension
1A•1	National evaluation laws and strategies	1A Institutional anchoring
1A•2	National decrees on evaluation	1A Institutional anchoring
1A•3	Organisational embedding	1A Institutional anchoring
1A•4	Evaluation role of the supreme audit institution	1A Institutional anchoring
1B•1	Parliamentary use of evaluation	1B Evaluation use
1B•2	Sectoral spread of evaluation	1B Evaluation use
1B•3	Scope of evaluation practice	1B Evaluation use
2A•1	Use of evaluation by civil society	2A Societal diffusion
2A•2	Public knowledge of evaluation	2A Societal diffusion
2A•3	Media and public discourse	2A Societal diffusion
2A•4	Public availability of evaluation reports	2A Societal diffusion
2A•5	Civic demand for evaluation	2A Societal diffusion
3A•1	Academic education in evaluation	3A Professionalisation
3A•2	Professional communication and exchange	3A Professionalisation
3A•3	Organisation of the profession (VOPEs)	3A Professionalisation
3A•4	Professional norms and standards	3A Professionalisation

Note: per cent of the theoretical maximum.

III

Data and method

This chapter describes the coding, the scales and the aggregation from indicators to the composite index.

3. Data and method

Key messages

- Each indicator is coded from documentary evidence on a five-point scale.
- Sub-dimensions are equally weighted in the composite.
- Scores are expressed as a share of the theoretical maximum.

The first edition covers 11 states: six in the Gulf, two in the Levant and three in North Africa. Each of the 16 indicators is coded on a common scale from 0 (absent) to 1 (fully institutionalised) on the basis of the documented situation in each state, as analysed in the country chapters of Ben Brik (2025). Binary indicators take the values 0 or 1; three-level indicators take 0, 0.5 or 1; five-level indicators move in steps of 0.25.

Indicators are averaged into four sub-dimensions: institutional anchoring (1A, four indicators) and evaluation use (1B, three indicators) in the political system, societal diffusion (2A, five indicators) in the social system, and professionalisation (3A, four indicators) in the profession system. The composite index is the unweighted mean of the four sub-dimensions; the political system score is the mean of 1A and 1B. States are grouped as consolidating (22 per cent and above), emerging (14 to 22 per cent) and nascent (below 14 per cent).

IV

The composite index

This chapter presents composite scores and the distribution of states across three groups.

4. The composite index

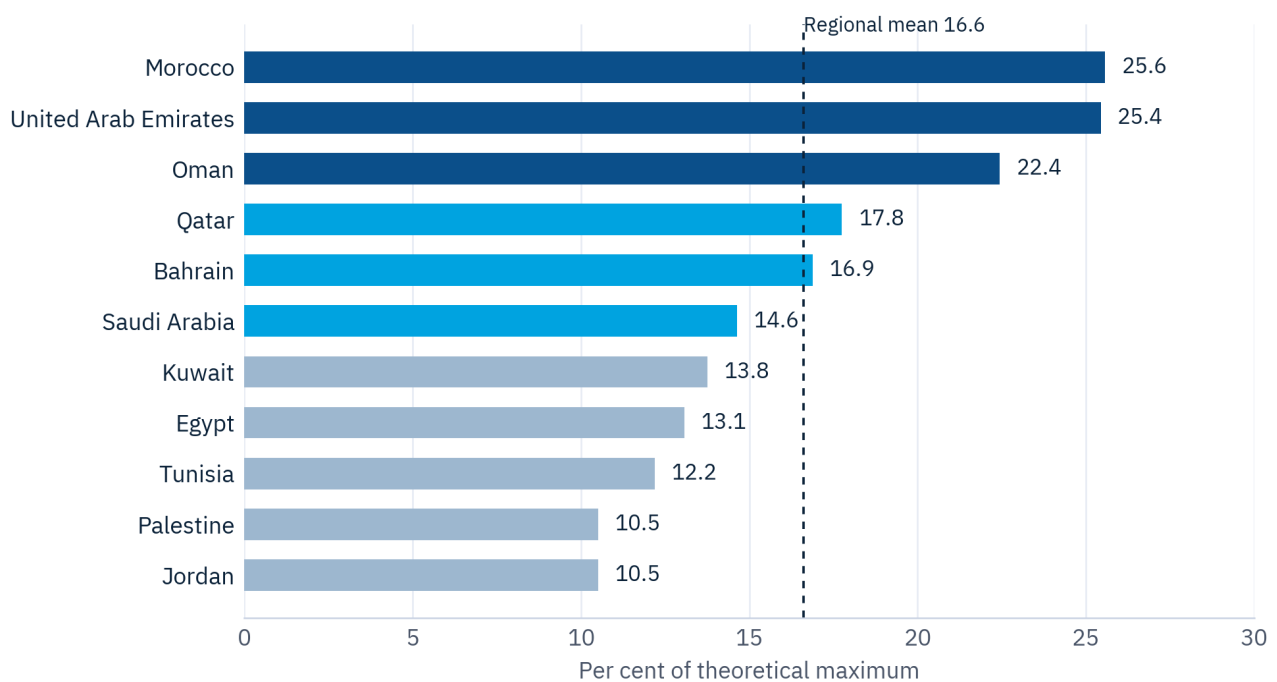
Key messages

- The regional mean is 16.6%; the highest score is 25.6% (Morocco).
- Scores are low in absolute terms and compressed across states.
- No state reaches a third of the theoretical maximum.

Composite scores are low in absolute terms and compressed across states. The most institutionalised evaluation system reaches 25.6% of the theoretical maximum, and the regional mean is 16.6%. The distance to full institutionalisation is therefore considerable everywhere, and the differences between states should be read as differences in the early stages of institution building.

Scores range from 10.5% to 25.6%. 3 states form the consolidating group (Morocco, United Arab Emirates and Oman), 3 the emerging group (Qatar, Bahrain and Saudi Arabia) and 5 the nascent group (Kuwait, Egypt, Tunisia, Jordan and Palestine). Every state lies below one third of the theoretical maximum, which places the region at an early stage of institution building. Use the controls to filter by region or to display any sub-dimension.

Figure 4.1. Composite index of evaluation institutionalisation, by state



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

The table reports the four sub-dimensions from which the composite is computed. Societal diffusion is zero in nine states, professionalisation varies most, and institutional anchoring never exceeds 25 per cent. The regional mean row provides the benchmark used in the state profiles.

Table 4.1. Composite index and sub-dimension scores, by state (per cent)

State	Region	Composite	1A	1B	2A	3A	Group
Morocco	North Africa	25.6%	13.0%	33.0%	0.0%	56.2%	Consolidating
United Arab Emirates	Gulf	25.4%	25.0%	33.0%	0.0%	43.8%	Consolidating
Oman	Gulf	22.4%	13.0%	33.0%	0.0%	43.8%	Consolidating
Qatar	Gulf	17.8%	13.0%	33.0%	0.0%	25.0%	Emerging
Bahrain	Gulf	16.9%	13.0%	17.0%	25.0%	12.5%	Emerging
Saudi Arabia	Gulf	14.6%	13.0%	33.0%	0.0%	12.5%	Emerging
Kuwait	Gulf	13.8%	13.0%	17.0%	25.0%	0.0%	Nascent
Egypt	North Africa	13.1%	13.0%	33.0%	0.0%	6.2%	Nascent
Tunisia	North Africa	12.2%	13.0%	17.0%	0.0%	18.8%	Nascent
Jordan	Levant	10.5%	0.0%	17.0%	0.0%	25.0%	Nascent
Palestine	Levant	10.5%	0.0%	17.0%	0.0%	25.0%	Nascent

Note: per cent of the theoretical maximum.

V

The political system: anchoring and use

This chapter examines institutional anchoring and the use of evaluation within the state.

5. The political system: anchoring and use

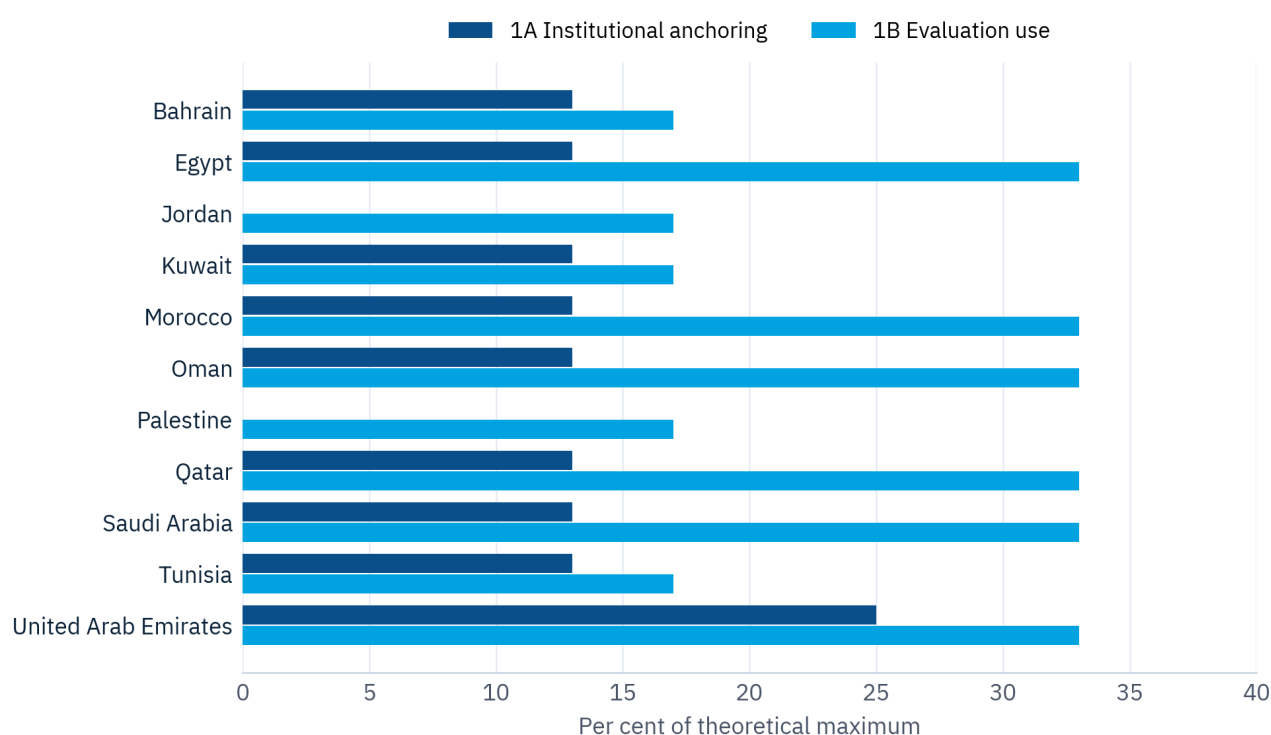
Key messages

- The political system averages 18.6%.
- No state has a national evaluation law or decree; no supreme audit institution has an evaluation role.
- Sectoral spread is the only indicator present in all states.

The political system records a regional mean of 18.6%. Institutional anchoring (1A, mean 11.7%) rests almost entirely on organisational embedding: no state has a national evaluation law, strategy or decree, and supreme audit institutions nowhere exercise a formal evaluative mandate. Evaluation use (1B, mean 25.7%) reflects the spread of evaluation across policy sectors and, in 6 states, a broader scope of evaluative practice; parliaments do not use evaluation in any state.

Evaluation use (mean 25.7%) exceeds institutional anchoring (mean 11.7%) in most states: evaluation is practised across sectors more than it is mandated by binding rules. United Arab Emirates records full organisational embedding; Jordan and Palestine record none.

Figure 5.1. Political system: institutional anchoring (1A) and evaluation use (1B)



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

Four of the seven political indicators score zero in every state: national laws and strategies, national decrees, the evaluation role of the supreme audit institution and parliamentary use. The political system therefore rests on organisational embedding, sectoral spread and scope of practice.

Figure 5.2. Political system indicators, by state

	1A-1	1A-2	1A-3	1A-4	1B-1	1B-2	1B-3
Morocco			0.50			0.50	0.50
United Arab Emirates			1.00			0.50	0.50
Oman			0.50			0.50	0.50
Qatar			0.50			0.50	0.50
Bahrain			0.50			0.50	
Saudi Arabia			0.50			0.50	0.50
Kuwait			0.50			0.50	
Egypt			0.50			0.50	0.50
Tunisia			0.50			0.50	
Jordan						0.50	
Palestine						0.50	

Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

VI

The social system: diffusion and demand

This chapter examines whether evaluation reaches civil society, the media and the public.

6. The social system: diffusion and demand

Key messages

- The social system averages 4.5%.
- It is recorded in two states only: Bahrain and Kuwait.
- Demand for evaluation outside the state is the weakest link of institutionalisation.

The social system is the weakest component of institutionalisation, with a regional mean of 4.5%. Only Bahrain and Kuwait record any societal diffusion, at a uniform 25 per cent on each of the five indicators. Elsewhere, evaluation is not used by civil society, not discussed in the media, not published systematically and not demanded by citizens. In the terms of the evaluation use literature, findings lack the audiences and arenas in which evidence acquires political salience.

Societal diffusion is confined to Bahrain and Kuwait, each scoring 25 per cent on all five indicators. The uniformity of these scores points to a limited but generalised public presence of evaluation rather than to a developed civic demand.

Figure 6.1. Social system indicators, by state

	2A-1	2A-2	2A-3	2A-4	2A-5
Morocco					
United Arab Emirates					
Oman					
Qatar					
Bahrain	0.25	0.25	0.25	0.25	0.25
Saudi Arabia					
Kuwait	0.25	0.25	0.25	0.25	0.25
Egypt					
Tunisia					
Jordan					
Palestine					

Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

VII

The profession system: professionalisation

This chapter examines education, professional exchange, associations and standards.

7. The profession system: professionalisation

Key messages

- Professionalisation averages 24.4% and varies most across states.
- It drives most of the difference in composite scores.
- No state has adopted professional norms or standards.

Professionalisation records a regional mean of 24.4% and the widest dispersion of the three systems, from 0 to 56.2%. Academic education in evaluation is the most developed indicator, reaching the full score in Morocco, United Arab Emirates, Oman and Qatar. Professional communication is strongest in United Arab Emirates, Oman and Jordan. Organised VOPEs exist in Morocco, Jordan and Palestine, and no state has adopted codified professional standards or certification.

Academic education is the leading component of professionalisation, while professional norms score zero everywhere. Morocco combines full academic education with the only fully developed VOPE in the index, which accounts for its leading composite position.

Figure 7.1. Profession system indicators, by state

	3A-1	3A-2	3A-3	3A-4
Morocco	1.00	0.25	1.00	
United Arab Emirates	1.00	0.75		
Oman	1.00	0.75		
Qatar	1.00			
Bahrain	0.50			
Saudi Arabia	0.50			
Kuwait				
Egypt		0.25		
Tunisia	0.50	0.25		
Jordan		0.75	0.25	
Palestine	0.50	0.25	0.25	

Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

VIII

Configurations across systems

This chapter reads the three systems together and asks how they combine within each state.

8. Configurations across systems

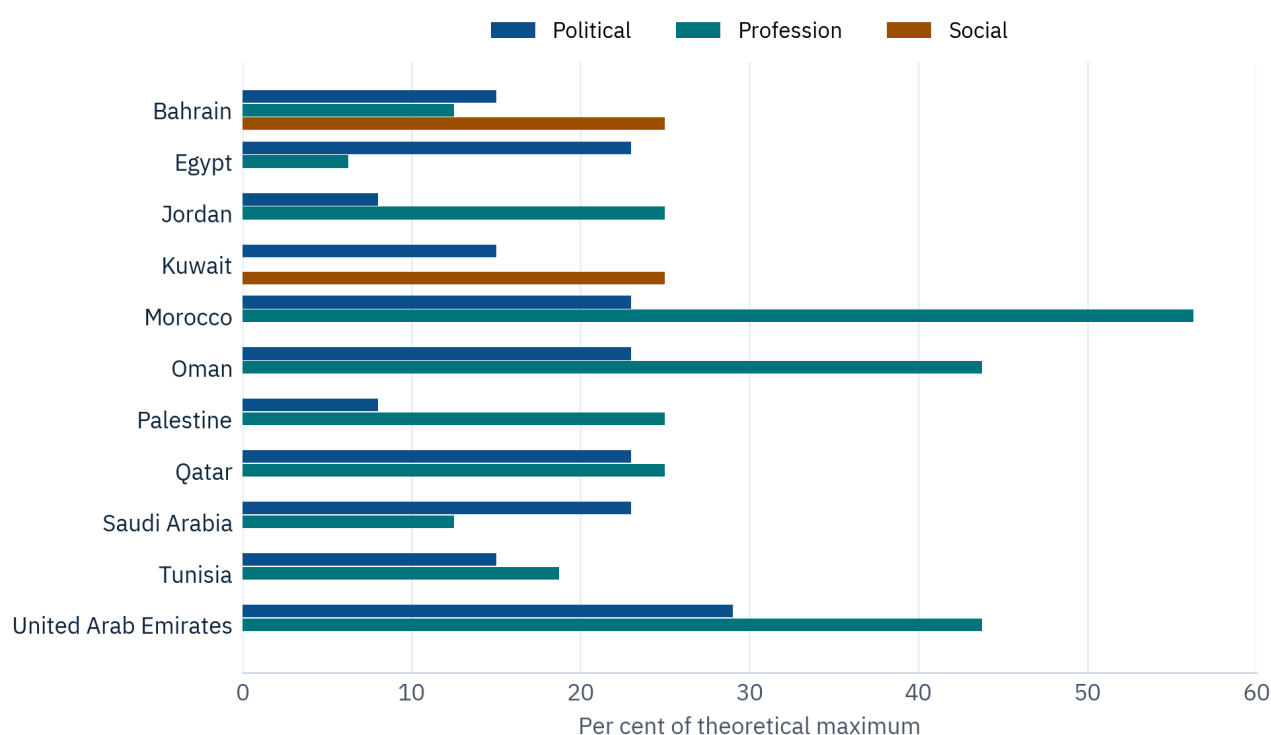
Key messages

- Formal structures sit in the executive; professional capacity varies; societal demand is largely absent.
- Political anchoring and professionalisation are only loosely related.
- Regional means conceal distinct national configurations.

Read together, the three systems describe a common configuration: formal structures in the executive, a partially developed profession and an absent demand side. The decomposition of the composite shows that the leading states owe their position to professionalisation; the gap between the political and social systems shows the extent of decoupling between executive structures and societal uptake; and the coverage of indicators shows which components of institutionalisation are missing across the entire region.

The profession system shows the widest range between states (0 to 56.2%), the political system a narrower range (8.0% to 29.0%) and the social system the narrowest (0 to 25 per cent). Differences in composite scores are therefore driven primarily by professionalisation.

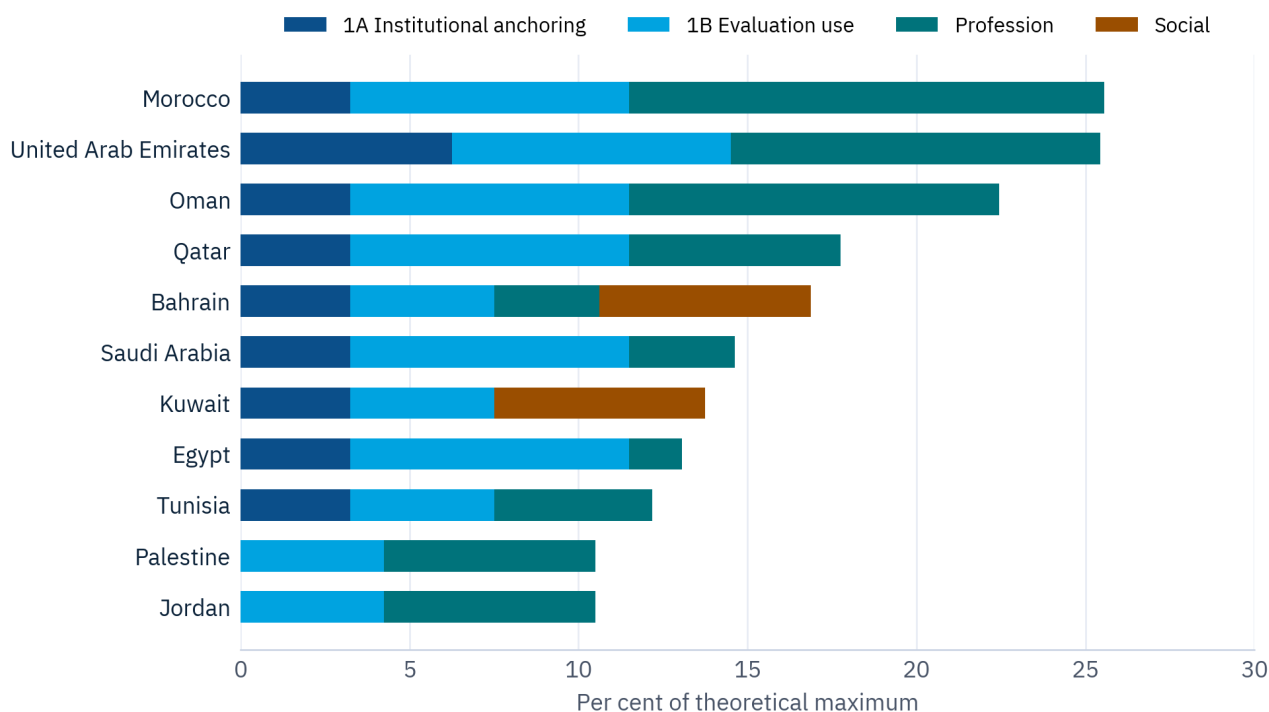
Figure 8.1. Political, social and profession system scores, by state



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

Each segment shows a sub-dimension score divided by four, so that segments sum to the composite. Professionalisation contributes 14.1 points to the composite of Morocco, more than any other sub-dimension in any state.

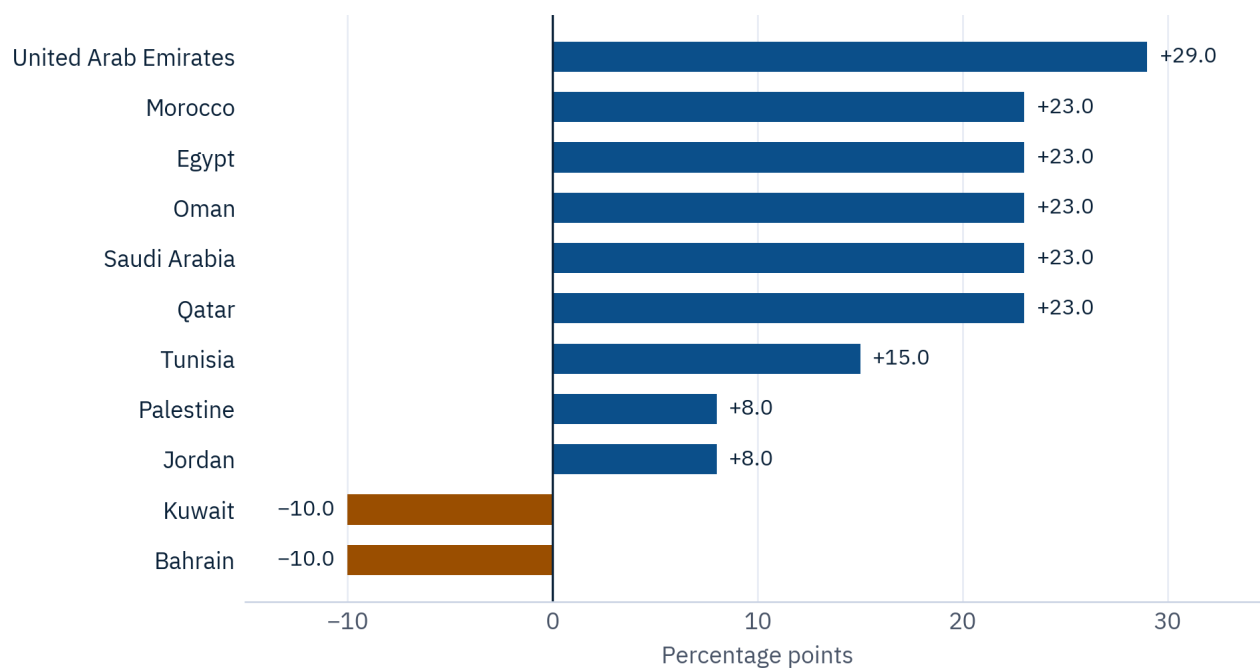
Figure 8.2. Decomposition of the composite index into sub-dimension contributions



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

The political system exceeds the social system in 9 states, by up to 29.0 points. Only Bahrain and Kuwait show the reverse, because societal diffusion is recorded there while political anchoring is weaker. The gap is an indicator of decoupling between executive structures and the social uptake of evaluation.

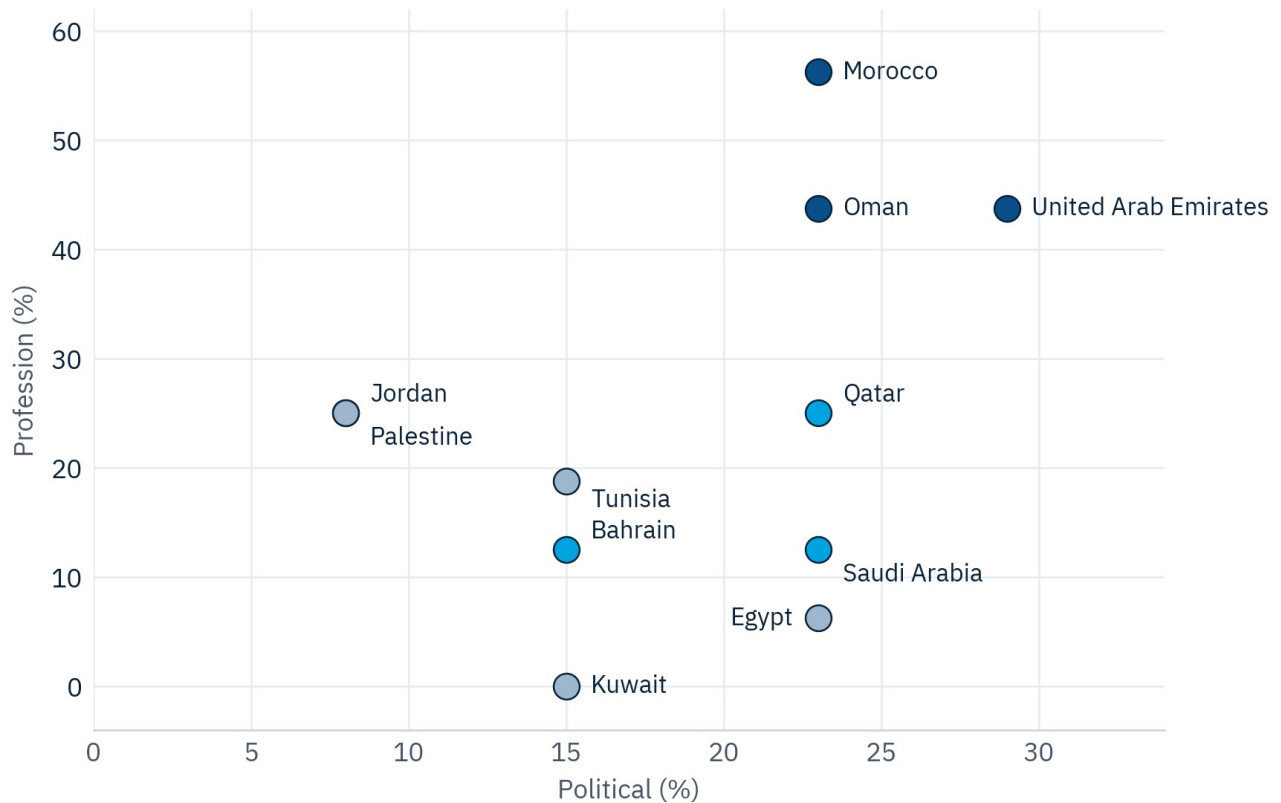
Figure 8.3. Decoupling between political and social systems (percentage points)



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: *Evaluation GLOBE* (Stockmann et al., 2020), used with permission.

States with similar political system scores differ widely in professionalisation: at a political score of 23 per cent, professionalisation ranges from 6.2% to 56.2%. Political anchoring is thus a weak predictor of professional development. Change the axes to explore other pairs of measures.

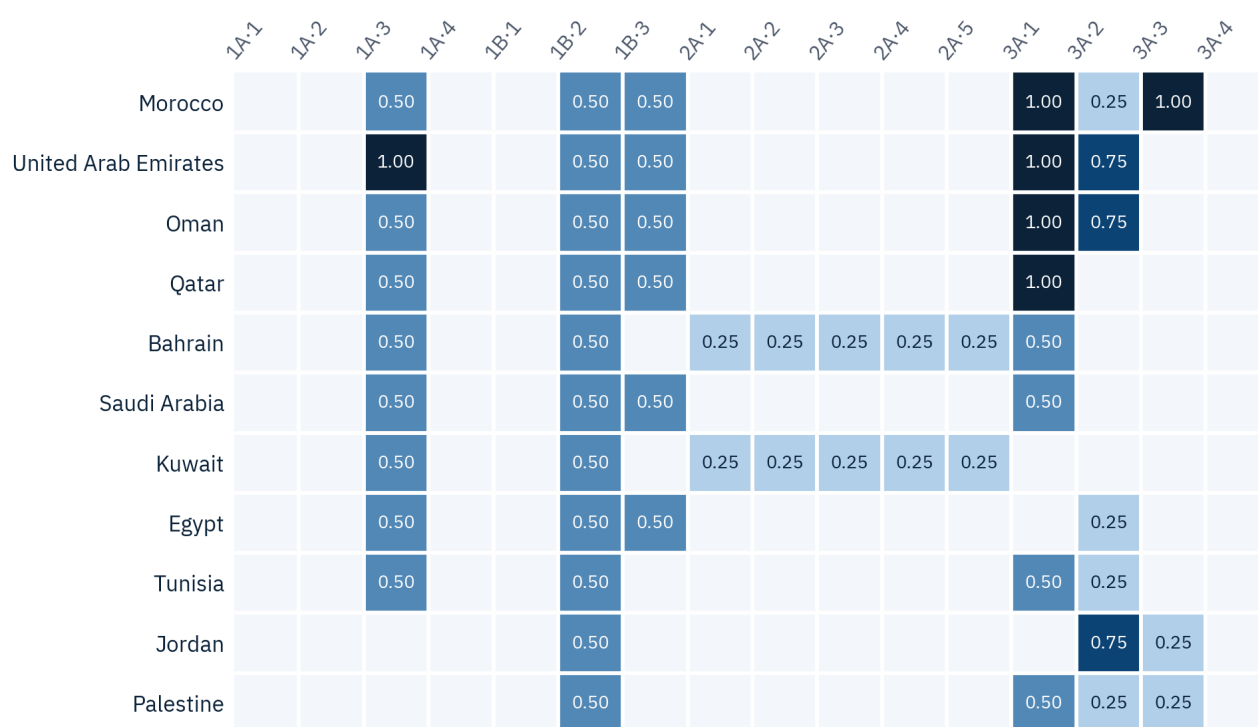
Figure 8.4. Political anchoring and professionalisation



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: *Evaluation GLOBE* (Stockmann et al., 2020), used with permission.

The heatmap shows the full coding. Columns without shading identify the components missing across the region; rows show each state's configuration.

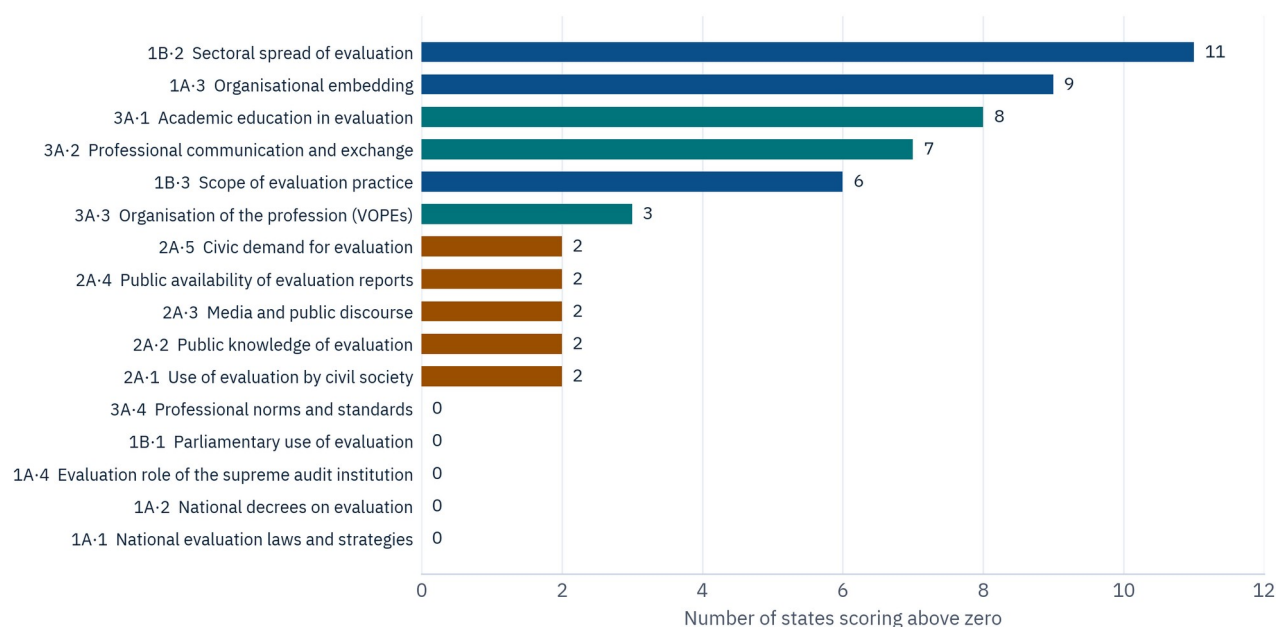
Figure 8.5. All 16 indicators, by state



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

Sectoral spread is the only indicator present in all 11 states; organisational embedding is present in nine. Five indicators are absent everywhere, and societal indicators are present in two states only. The profile of diffusion confirms an executive-centred pattern of institutionalisation.

Figure 8.6. Diffusion of each indicator across states



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

Unweighted regional means should be read with caution given the small number of states per region, particularly the two Levant states. The Gulf records the only societal diffusion in the index; North Africa records the highest mean professionalisation.

Table 8.1. Mean scores by region (per cent)

Region	Composite	Political	Social	Profession
Gulf	18.5%	21.3%	8.3%	22.9%
Levant	10.5%	8.0%	0.0%	25.0%
North Africa	16.9%	20.3%	0.0%	27.1%

Note: per cent of the theoretical maximum.

Indicators with a mean of zero and no dispersion carry no discriminating information in this edition but identify the priorities of evaluation policy. Discriminating indicators are concentrated in the profession system.

Table 8.2. Indicator statistics: mean, dispersion and diffusion

Indicator	Mean	SD	States above zero
1A·1 National evaluation laws and strategies	0.00	0.00	0
1A·2 National decrees on evaluation	0.00	0.00	0
1A·3 Organisational embedding	0.45	0.26	9
1A·4 Evaluation role of the supreme audit institution	0.00	0.00	0
1B·1 Parliamentary use of evaluation	0.00	0.00	0
1B·2 Sectoral spread of evaluation	0.50	0.00	11
1B·3 Scope of evaluation practice	0.27	0.25	6
2A·1 Use of evaluation by civil society	0.05	0.10	2
2A·2 Public knowledge of evaluation	0.05	0.10	2
2A·3 Media and public discourse	0.05	0.10	2
2A·4 Public availability of evaluation reports	0.05	0.10	2
2A·5 Civic demand for evaluation	0.05	0.10	2
3A·1 Academic education in evaluation	0.55	0.40	8
3A·2 Professional communication and exchange	0.30	0.30	7
3A·3 Organisation of the profession (VOPEs)	0.14	0.29	3
3A·4 Professional norms and standards	0.00	0.00	0

Note: per cent of the theoretical maximum.

IX

Institutional architecture and gaps

This chapter examines the architecture of institutionalisation: the share of each system in the total of a state, and the gaps between systems.

9. Institutional architecture and gaps

Key messages

- 9 of 11 states lie on the political–profession edge: no societal system is recorded in them.
- The widest gap separates the profession and social systems in Morocco (56.2 points).
- States with similar composite scores have different architectures.

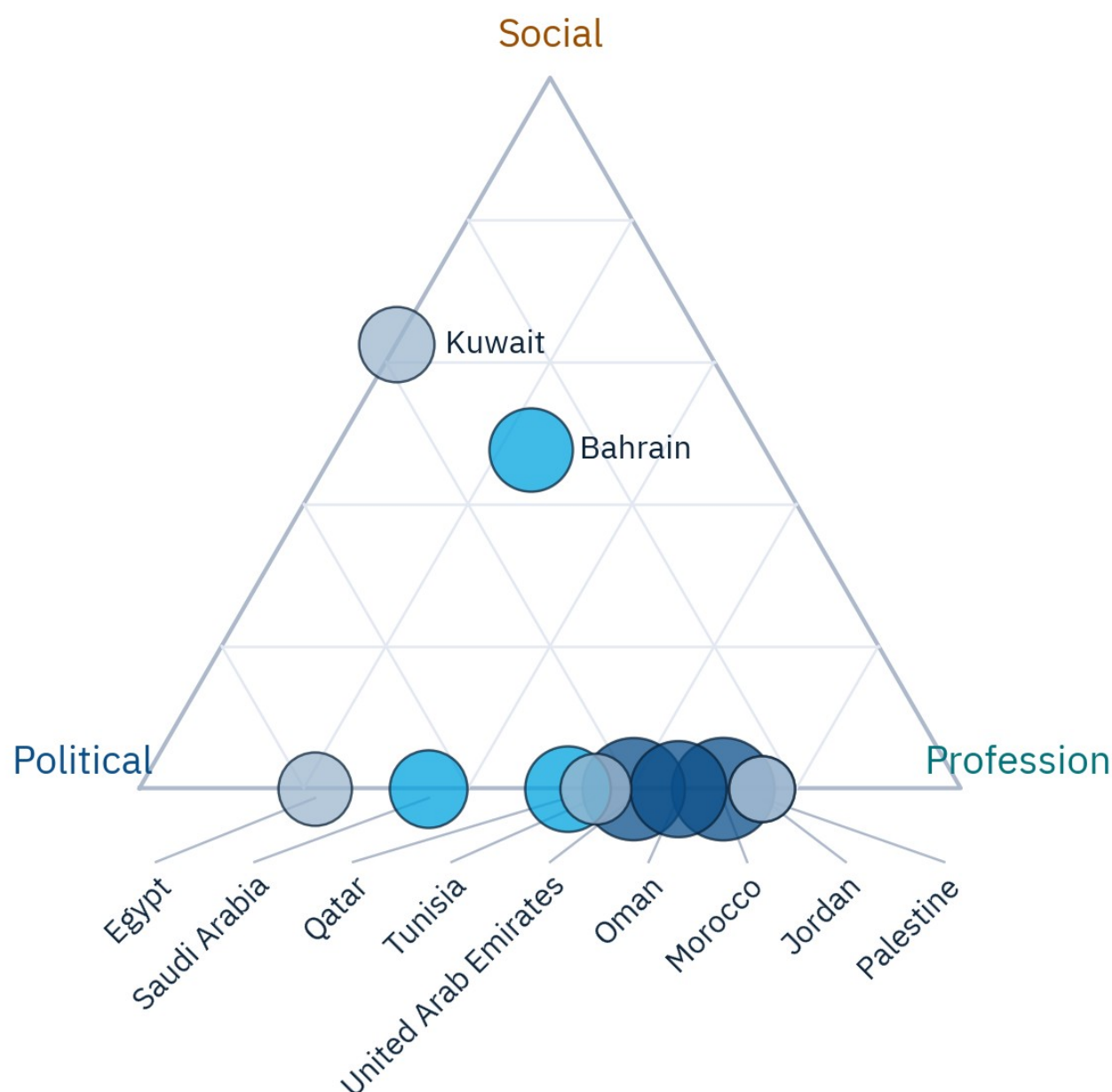
The composite index measures how much evaluation is institutionalised; it does not show how institutionalisation is distributed. The share of each system in the total of a state describes that architecture. In 9 of the 11 states the social system scores zero, so their institutionalisation rests entirely on the political and professional systems; only Bahrain and Kuwait record a societal component, and in both it forms the largest share.

Along the political–profession edge, position separates states in which professional capacity leads from states anchored mainly in the executive. Egypt and Saudi Arabia sit closest to the political pole; Morocco, Jordan and Palestine closest to the profession pole. The pairwise gaps between systems make the imbalance explicit. The largest gap in the region separates the profession and social systems in Morocco (56.2 points), followed by the United Arab Emirates (43.8 points).

The implication is that similar composite scores conceal different institutional problems. A state whose score rests on executive structures requires professional and societal capacity; a state whose score rests on the profession requires anchoring in the state and in its oversight institutions.

The configuration chart places each state by the share of its three systems in its total, with circle area proportional to the composite score. It answers a question the composite cannot: whether institutionalisation rests on the state, on the profession or on society. Most states sit on the lower edge, where the social share is zero.

Figure 9.1. Institutional configuration: share of each system in the state total



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

The table reports the three pairwise gaps for every state, with the system that leads. It is the numerical basis for the configuration chart.

Table 9.1. Pairwise gaps between systems (percentage points)

State	Political– Profession	Political–Social	Profession–Social	Leading system
Bahrain	+2.5	–10.0	–12.5	Social
Egypt	+16.8	+23.0	+6.2	Political
Jordan	–17.0	+8.0	+25.0	Profession
Kuwait	+15.0	–10.0	–25.0	Social
Morocco	–33.2	+23.0	+56.2	Profession
Oman	–20.8	+23.0	+43.8	Profession
Palestine	–17.0	+8.0	+25.0	Profession
Qatar	–2.0	+23.0	+25.0	Profession
Saudi Arabia	+10.5	+23.0	+12.5	Political
Tunisia	–3.8	+15.0	+18.8	Profession
United Arab Emirates	–14.8	+29.0	+43.8	Profession

Note: per cent of the theoretical maximum.

The pairwise gaps give the size and direction of the imbalance between systems in each state. The profession–social gap is the widest in most states, which identifies societal demand, not professional supply, as the missing component.

Figure 9.2. Where institutionalisation is uneven: gaps and leading system

	Political–Profession	Political–Social	Profession–Social
Bahrain	2.5 Political	10.0 Social	12.5 Social
Egypt	16.8 Political	23.0 Political	6.2 Profession
Jordan	17.0 Profession	8.0 Political	25.0 Profession
Kuwait	15.0 Political	10.0 Social	25.0 Social
Morocco	33.2 Profession	23.0 Political	56.2 Profession
Oman	20.8 Profession	23.0 Political	43.8 Profession
Palestine	17.0 Profession	8.0 Political	25.0 Profession
Qatar	2.0 Profession	23.0 Political	25.0 Profession
Saudi Arabia	10.5 Political	23.0 Political	12.5 Profession
Tunisia	3.8 Profession	15.0 Political	18.8 Profession
United Arab Emirates	14.8 Profession	29.0 Political	43.8 Profession

Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

X

Supply, demand and the administrative capacity pathway

This chapter regroups existing indicators to ask whether evaluation operates as an administrative function or as part of a wider accountability system, and where the pathway from mandate to use breaks.

10. Supply, demand and the administrative capacity pathway

Key messages

- Supply exceeds demand in 10 of 11 states; horizontal accountability is absent everywhere.
- The pathway rises from mandate (0.0%) to practice (38.6%) and collapses at use (2.3%).
- Organisation precedes mandate in 9 states; practice exceeds governance use in 11.

Supply and demand are regroupings of existing indicators, not a new index. Supply combines administrative capacity (laws, decrees, organisational embedding, sectoral spread and scope) and professional capacity (education, exchange, associations and norms). Demand combines horizontal accountability (the evaluation role of the supreme audit institution and parliamentary use) and societal demand (the five indicators of the social system). Supply exceeds demand in 10 of 11 states; in Kuwait societal demand exceeds supply. Horizontal accountability scores zero throughout the region.

The administrative capacity pathway maps the indicators to four stages: mandate (laws and decrees), organisation (organisational embedding and the audit role), practice (sectoral spread and scope) and use (parliamentary and civil society use). The regional means trace the break: 0.0% for mandate, 22.7% for organisation, 38.6% for practice and 2.3% for use. Learning, the fifth stage in the sequence, is not measured by the index.

Two forms of decoupling follow. Organisational embedding exceeds formal anchoring in 9 of 11 states: evaluation units exist where binding rules do not. Routine practice exceeds governance use in 11 of 11 states: evaluation is carried out across sectors, but its findings do not enter audit, parliamentary or civic arenas.

The stages and their indicator mapping are the author's for this edition. They reorganise the coded evidence; they do not add new data.

The supply–demand chart sets indicators of capacity to produce evaluation against indicators of demand for its use. Bars to the left measure supply, bars to the right demand. The absence of any right-hand bar in nine states is the principal finding.

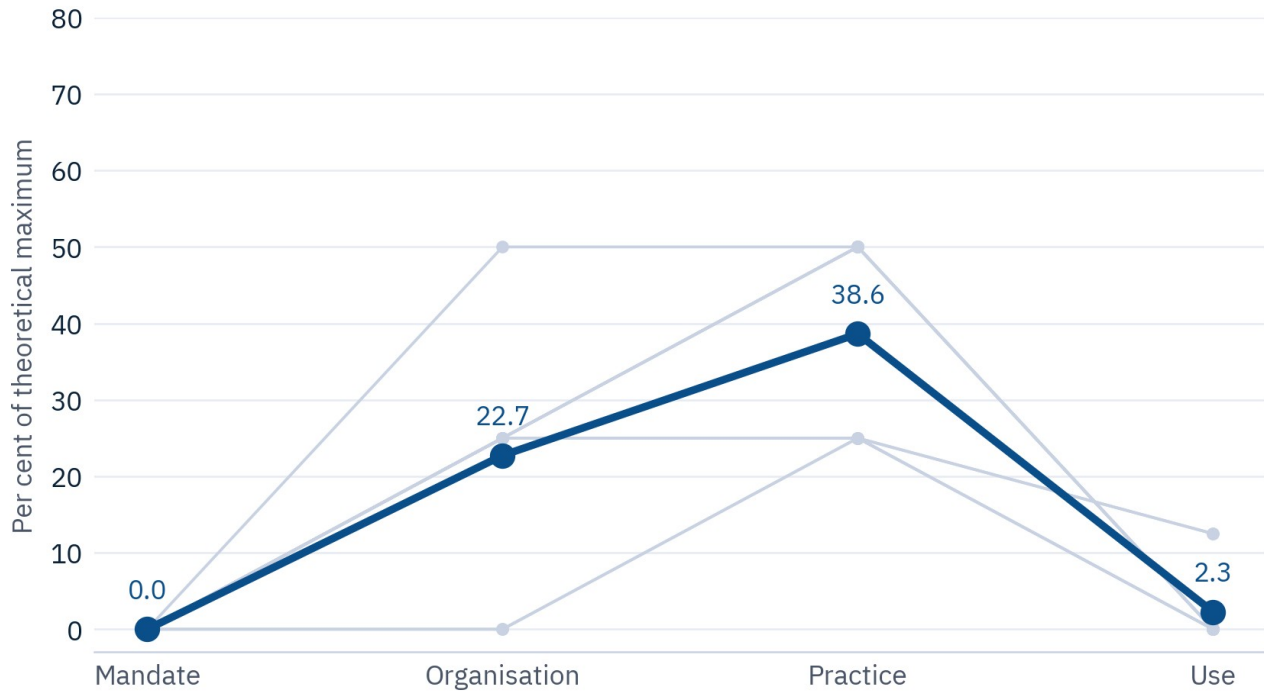
Figure 10.1. Evaluation supply and demand



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: *Evaluation GLOBE* (Stockmann et al., 2020), used with permission.

The pathway traces the regional mean, and each state in grey, across the four stages from mandate to use. The line rises through organisation and practice and falls at use: evaluation is carried out, but its findings do not reach oversight.

Figure 10.2. The administrative capacity pathway



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: *Evaluation GLOBE* (Stockmann et al., 2020), used with permission.

XI

Exploratory configurations

This chapter groups states by the similarity of their coded profiles. With eleven states, the groups are exploratory, not a regional classification.

II. Exploratory configurations

Key messages

- Three configurations emerge from Ward clustering of the sixteen indicators.
- Morocco stands alone as a professionalised configuration.
- The two clustering bases agree on 33 of 55 state pairs.

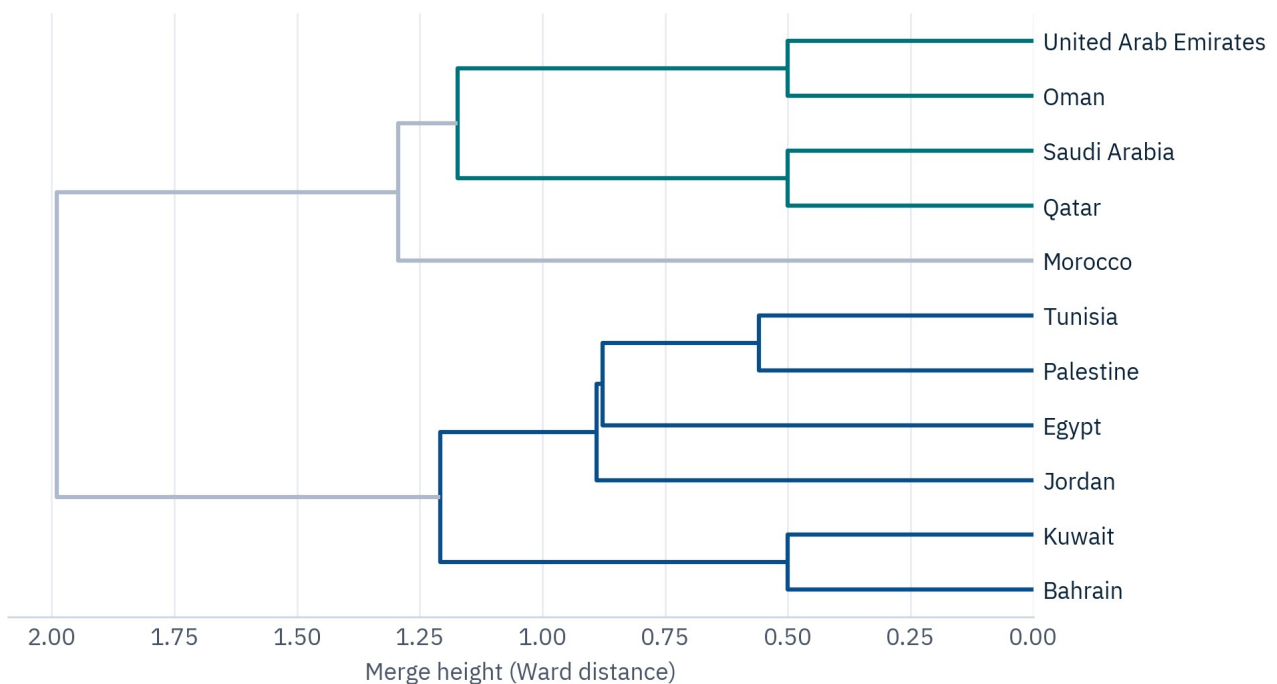
Hierarchical clustering with Ward linkage on the sixteen indicators groups the states into three configurations. Group 1 (Bahrain, Kuwait, Jordan, Palestine, Egypt and Tunisia) combines thin institutional bases in all three systems: political 14.0%, profession 14.6% and social 8.3% on average. Group 2 is Morocco alone, with the most professionalised system in the region (profession 56.2%, political 23.0%, social 0.0%). Group 3 (Qatar, Saudi Arabia, Oman and United Arab Emirates) combines organised executive systems with developing professional capacity (political 24.5%, profession 31.2%, social 0.0%).

The groups are sensitive to the clustering basis. At three groups, clustering on the sixteen indicators and on the four equally weighted sub-dimensions agrees on 33 of 55 state pairs. The five societal indicators move together, so the indicator basis weights the social system five times; the sub-dimension basis mirrors the equal weighting of the composite.

The configurations are therefore exploratory. They are useful as a way of reading similarity among states, not as a regional typology.

The dendrogram shows how states merge as similarity decreases. The height at which branches join measures the distance between configurations; the three coloured branches are the groups discussed in the text.

Figure II.1. Exploratory configurations: Ward clustering of the sixteen indicators



Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: *Evaluation GLOBE* (Stockmann et al., 2020), used with permission.

Table 11.1. Mean system scores by group

Group	States	Political	Profession	Social
1	Bahrain, Kuwait, Jordan, Palestine, Egypt and Tunisia	14.0%	14.6%	8.3%
2	Morocco	23.0%	56.2%	0.0%
3	Qatar, Saudi Arabia, Oman and United Arab Emirates	24.5%	31.2%	0.0%

Note: per cent of the theoretical maximum.

XII

State profiles

A short profile of each state, with a link to its country brief.

12. State profiles

The following profiles summarise the institutional configuration of each state from the coded indicators. They are descriptive and do not rank governments; small differences carry little substantive weight at these score levels.

Bahrain

Bahrain (Gulf) scores 16.9%, above the regional mean, in the Emerging group. Its strongest sub-dimension is 2A Societal diffusion (25.0%). It scores above zero on 8 of 16 indicators.

Country brief: mena-evaluation.com/briefs/bahrain-en.pdf

Egypt

Egypt (North Africa) scores 13.1%, below the regional mean, in the Nascent group. Its strongest sub-dimension is 1B Evaluation use (33.0%). It scores above zero on 4 of 16 indicators.

Country brief: mena-evaluation.com/briefs/egypt-en.pdf

Jordan

Jordan (Levant) scores 10.5%, below the regional mean, in the Nascent group. Its strongest sub-dimension is 3A Professionalisation (25.0%). It scores above zero on 3 of 16 indicators.

Country brief: mena-evaluation.com/briefs/jordan-en.pdf

Kuwait

Kuwait (Gulf) scores 13.8%, below the regional mean, in the Nascent group. Its strongest sub-dimension is 2A Societal diffusion (25.0%). It scores above zero on 7 of 16 indicators.

Country brief: mena-evaluation.com/briefs/kuwait-en.pdf

Morocco

Morocco (North Africa) scores 25.6%, above the regional mean, in the Consolidating group. Its strongest sub-dimension is 3A Professionalisation (56.2%). It scores above zero on 6 of 16 indicators, reaching the full score on Academic education in evaluation and Organisation of the profession (VOPEs).

Country brief: mena-evaluation.com/briefs/morocco-en.pdf

Oman

Oman (Gulf) scores 22.4%, above the regional mean, in the Consolidating group. Its strongest sub-dimension is 3A Professionalisation (43.8%). It scores above zero on 5 of 16 indicators, reaching the full score on Academic education in evaluation.

Country brief: mena-evaluation.com/briefs/oman-en.pdf

Palestine

Palestine (Levant) scores 10.5%, below the regional mean, in the Nascent group. Its strongest sub-dimension is 3A Professionalisation (25.0%). It scores above zero on 4 of 16 indicators.

Country brief: mena-evaluation.com/briefs/palestine-en.pdf

Qatar

Qatar (Gulf) scores 17.8%, above the regional mean, in the Emerging group. Its strongest sub-dimension is 1B Evaluation use (33.0%). It scores above zero on 4 of 16 indicators, reaching the full score on Academic education in evaluation.

Country brief: mena-evaluation.com/briefs/qatar-en.pdf

Saudi Arabia

Saudi Arabia (Gulf) scores 14.6%, below the regional mean, in the Emerging group. Its strongest sub-dimension is 1B Evaluation use (33.0%). It scores above zero on 4 of 16 indicators.

Country brief: mena-evaluation.com/briefs/saudi-arabia-en.pdf

Tunisia

Tunisia (North Africa) scores 12.2%, below the regional mean, in the Nascent group. Its strongest sub-dimension is 3A Professionalisation (18.8%). It scores above zero on 4 of 16 indicators.

Country brief: mena-evaluation.com/briefs/tunisia-en.pdf

United Arab Emirates

United Arab Emirates (Gulf) scores 25.4%, above the regional mean, in the Consolidating group. Its strongest sub-dimension is 3A Professionalisation (43.8%). It scores above zero on 5 of 16 indicators, reaching the full score on Organisational embedding and Academic education in evaluation.

Country brief: mena-evaluation.com/briefs/united-arab-emirates-en.pdf

XIII

Discussion: decoupling, supply and demand

This chapter interprets the evidence and sets out what it cannot establish.

13. Discussion: decoupling, supply and demand

Key messages

- Evaluation is organised before it is mandated, and practised before it is used.
- Supply of evaluation is not matched by demand from oversight institutions or society.
- The coding describes institutions; it does not establish their effects.

Three interpretations follow from the evidence. First, the absence of legal anchoring alongside organisational embedding is consistent with institutional decoupling: evaluation units and references to evaluation in planning documents coexist with the absence of binding obligations to evaluate, to publish or to respond to findings (Meyer & Rowan, 1977). Second, the prominence of professionalisation in the leading states suggests a normative pathway of institutionalisation, in which universities, professional networks and VOPEs build capacity ahead of the state (DiMaggio & Powell, 1983). Third, the near-total absence of societal diffusion and parliamentary use indicates that evaluation has not yet become an instrument of horizontal or social accountability.

These configurations carry a risk familiar from the state capability literature: organisations may acquire the forms of evaluation without the functions, and reforms may reward the adoption of evaluation units rather than the use of evaluative evidence (Pritchett et al., 2013). The distance between mandates and use is therefore the central object of evaluation policy in the region.

XIV

Implications for evaluation policy and public administration

This chapter sets out reform directions and recommendations by actor.

14. Implications for evaluation policy and public administration

Key messages

- Anchor evaluation in law and link it to budget and oversight.
- Create demand through publication, audit and parliamentary review.
- Consolidate the profession through education, associations and standards.

(i) Anchor evaluation in law or regulation that specifies obligations to evaluate major programmes, to publish reports and to issue management responses. (ii) Assign evaluative mandates to supreme audit institutions and parliamentary committees, so that evaluation feeds horizontal accountability. (iii) Institutionalise the publication of evaluation reports and management responses to create the conditions for societal diffusion and civic demand. (iv) Support the professionalisation pathway through degree programmes, professional journals and VOPEs, and develop national or regional standards and competency frameworks. (v) Build evaluation into budgeting and planning cycles, so that use becomes a routine of public management rather than an exception.

Table 14.1. Recommendations by actor

Priority	Action	Responsible actors	Intended outcome
Legal anchoring	Adopt a law, decree or national strategy that sets obligations to evaluate, assigns responsibilities and requires the use of findings	Centre of government; legislature	A formal mandate for evaluation in all eleven states
External oversight	Give supreme audit institutions a mandate to review the quality and use of evaluations	Supreme audit institutions; parliaments	Evaluation evidence enters external audit
Parliamentary use	Require evaluation evidence in budget and legislative review and transmit reports to committees	Parliaments; ministries of finance	Findings used in legislative scrutiny
Societal demand	Publish evaluation reports and open channels for civil society and media to use them	Ministries; evaluation units; civil society	Public availability and use of reports
Professional consolidation	Develop academic programmes, support evaluation associations and adopt professional standards	Universities; VOPEs; government evaluation units	An organised profession with shared norms

Note: per cent of the theoretical maximum.

XV

Limitations and future editions

This chapter states the limits of the first edition and the priorities for the next.

15. Limitations and future editions

Key messages

- Coding rests on documentary evidence and ordinal scales.
- Eleven states limit statistical inference.
- Future editions will extend coverage and add measures of use.

The index relies on documentary coding of the situation in each state and on ordinal scales treated as interval measures in aggregation. Equal weighting of sub-dimensions is a normative choice that future editions will test through sensitivity analysis. Coverage is limited to 11 states, and the first edition provides a baseline rather than a trend. Subsequent editions will extend coverage, publish the coding evidence indicator by indicator, and track change over time.

XVI

Conclusion

16. Conclusion

This report has measured the institutionalisation of public policy evaluation in eleven Arab League states. Evaluation is organised in the executive and supported, unevenly, by a developing profession; it is not anchored in law, not reviewed by oversight institutions and, in most states, not visible to the public. The composite scores are low and close; the architectures behind them differ.

The practical lesson is that institutionalisation requires demand as much as supply. Evaluation units and trained evaluators will not by themselves make evaluation an institution; obligations in law, review by audit institutions and parliaments, and publication to citizens will. The index will be repeated, and this edition sets the baseline against which change can be judged.

Annex A. Indicators

Table A.1. Annex A. Indicators

Code	Indicator	System
1A•1	National evaluation laws and strategies	Political system
1A•2	National decrees on evaluation	Political system
1A•3	Organisational embedding	Political system
1A•4	Evaluation role of the supreme audit institution	Political system
1B•1	Parliamentary use of evaluation	Political system
1B•2	Sectoral spread of evaluation	Political system
1B•3	Scope of evaluation practice	Political system
2A•1	Use of evaluation by civil society	Social system
2A•2	Public knowledge of evaluation	Social system
2A•3	Media and public discourse	Social system
2A•4	Public availability of evaluation reports	Social system
2A•5	Civic demand for evaluation	Social system
3A•1	Academic education in evaluation	Profession system
3A•2	Professional communication and exchange	Profession system
3A•3	Organisation of the profession (VOPEs)	Profession system
3A•4	Professional norms and standards	Profession system

Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

Annex B. Coded scores by state

Table B.1. Annex B. Coded scores by state

State	1A- 1	1A- 2	1A- 3	1A- 4	1B- 1	1B- 2	1B- 3	2A- 1	2A- 2	2A- 3	2A- 4	2A- 5	3A- 1	3A- 2	3A- 3	3A- 4
Bahrain	0.00	0.00	0.50	0.00	0.00	0.50	0.00	0.25	0.25	0.25	0.25	0.25	0.50	0.00	0.00	0.00
Egypt	0.00	0.00	0.50	0.00	0.00	0.50	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00
Jordan	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.75	0.25	0.00
Kuwait	0.00	0.00	0.50	0.00	0.00	0.50	0.00	0.25	0.25	0.25	0.25	0.25	0.00	0.00	0.00	0.00
Morocco	0.00	0.00	0.50	0.00	0.00	0.50	0.50	0.00	0.00	0.00	0.00	0.00	1.00	0.25	1.00	0.00
Oman	0.00	0.00	0.50	0.00	0.00	0.50	0.50	0.00	0.00	0.00	0.00	0.00	1.00	0.75	0.00	0.00
Palestine	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.25	0.25	0.00
Qatar	0.00	0.00	0.50	0.00	0.00	0.50	0.50	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Saudi Arabia	0.00	0.00	0.50	0.00	0.00	0.50	0.50	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.00
Tunisia	0.00	0.00	0.50	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.25	0.00	0.00
United Arab Emirates	0.00	0.00	1.00	0.00	0.00	0.50	0.50	0.00	0.00	0.00	0.00	0.00	1.00	0.75	0.00	0.00

Source: Ben Brik (2025), *Index of Evaluation Institutionalization in the Arab Region*, first edition. Framework: Evaluation GLOBE (Stockmann et al., 2020), used with permission.

References

- Ben Brik, A. (2025). The institutionalization of evaluation in the Middle East and North Africa. Edward Elgar Publishing. <https://doi.org/10.4337/9781035359820>
- Ben Brik, A. (Ed.). (2025). Public policy evaluation in the Middle East and North Africa. Bristol University Press. <https://doi.org/10.46692/9781447375708>
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160.
- Furubo, J.-E., Rist, R. C., & Sandahl, R. (Eds.). (2002). International atlas of evaluation. Transaction Publishers.
- Jacob, S., Speer, S., & Furubo, J.-E. (2015). The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later. *Evaluation*, 21(1), 6–31.
- Leeuw, F. L., & Furubo, J.-E. (2008). Evaluation systems: What are they and why study them? *Evaluation*, 14(2), 157–169.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363.
- Pritchett, L., Woolcock, M., & Andrews, M. (2013). Looking like a state: Techniques of persistent failure in state capability for implementation. *Journal of Development Studies*, 49(1), 1–18.
- Stockmann, R., Meyer, W., & Taube, L. (Eds.). (2020). The institutionalisation of evaluation in Asia-Pacific. Palgrave Macmillan.
- Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21–33.



Arab Observatory for Policy Evaluation Institutionalization

Notes and sources

The designations employed do not imply the expression of any opinion concerning the legal status of any country or territory. Scores describe institutional arrangements coded from documentary evidence; they do not measure the performance of governments or the quality of individual evaluations. The views expressed are those of the author.

Please cite this publication as: Ben Brik, A. (2025). Index of Evaluation Institutionalization in the Arab Region: Annual Report 2025. Arab Observatory for Policy Evaluation Institutionalization, <https://mena-evaluation.com/report/>.