

# Statistical and Data Literacy in Policy-Making<sup>1</sup>

Gaby Umbach

*Global Governance Programme, Robert Schuman Centre for Advanced Studies, European University Institute, Via Giovanni Boccaccio 121, 50133 Florence, Italy*

*E-mail: [gaby.umbach@eui.eu](mailto:gaby.umbach@eui.eu)*

**Abstract:** This paper offers conceptual reflections on statistical and data literacy in policy-making. It discusses the relevance of the use of statistics and data in politics and highlights their impact on policy-making. It underlines the need for and identifies key meanings of statistical and data literacy in policy-making. It also highlights how statistical and data literacy in policy-making is specific.

**Keywords:** Official statistics; data; statistical and data literacy; policy-making; evidence-informed policy-making; politics

## *Setting the scene*

In the shadow of post-factual contested politics, interest-driven political reasoning and partisan communication, legitimate policy-making depends more than ever on the capacity of political actors to make decisions based on transparent and accessible evidence. Adhering to the principles of evidence-informed policies, reliable information thus plays an indispensable role in democratic politics and decision-making. Within it, statistics and data<sup>2</sup> are key manifestations of quantitative evidence about the state and society. They are hence essential tools of evidence-informed policy-making.

With the ‘evidence turn’ in policy-making [1-8], statistics and data have become instruments of collective political action. They measure social reality, offer insights into progress in development and have the potential to identify correlations and trends relevant to informing policy design. As such, they are used to derive at, explain and justify policy choices. This link between statistics, data and politics turns quantitative evidence into a central policy instrument and makes its use in policy-making a specific feature of evidence-informed policy-making.

Such use of statistics and data in democratic policy-making is subject to the principles of legitimacy, transparency and accountability, but also to preference-building and negotiation. Within policy-making, statistics and data become subject to scrutiny. The political power that emerges from their use is open to contestation over the choice and quality of the evidence at hand. In this way, the use of statistics and data supports deliberation in political processes and fosters participatory structures between data producers and users that engage in promoting quantitative evidence as opportunity structures. The former do so to support evidence-informed policy-making, the latter to define and defend policy priorities. Through these systemic interlinkages, the use of statistics and data becomes open to multiple, strategic and value-based

---

<sup>1</sup> Extended version published as: Umbach, Gaby (2022): Statistical and Data Literacy in Policy-Making, (Pre-press) *Special Stream* on ‘Statistical and Data Literacy in Policy-making’ of the *Statistical Journal of the International Association of Official Statistics* (SJIAOS), 38(2), June.

<sup>2</sup> For the purpose of this paper and to reflect the practice of political decision-making, the terms ‘data’ and ‘statistics’ are used together to include the multiverse of metrological instruments that are used as quantitative evidence in policy-making. Such instruments include official statistics on the one hand and indicators, indices, composite indicators, and/or scoreboards prepared by other actors than NSO on the other.

considerations that support, defend and/or contest quantification choices as much as political interests [9].

Accompanying these developments, the landscape of data sources and the ways of using statistics and data in policy-making have grown significantly and become easily confusing for any non-data scientist engaged in public policy. These developments nurture literacy requirements that impact on and are impacted by political actors' capacity to select, evaluate and process data. Understanding statistics, working with data, analysing and arguing through data, therefore, became essential challenges for professionals working in public policy analysis, policy preparation, policy-making, decision-making, evaluation and scrutiny. To improve the impact of statistics and data on politics, it is hence important to analyse what constitutes statistical and data literacy in policy-making and how it can be promoted.

For the purpose of this paper, i.e. to inspire conceptual reflection on the topic, we define 'actors engaged in policy-making' to include public officials engaged in policy preparation and implementation as well as political decision-makers. We also pay attention to citizens as principals and recipients of public policies and the addressee of political communication. Despite this broad perspective, we are aware of the differentiation of literacy needs depending on the role and position of actors in policy-making and on their contribution to the policy process. For a further categorisation of statistical and data literacy demands according to actors' roles and positions see Schüller [10].

### *The meaning of statistical and data literacy*

Contemporary political and professional life requires distinct types of skills and literacies related to statistics and data [11]. These include among others information and media literacy; communication, and visualisation literacy; digital literacy; computational and machine learning literacy; statistical and ethical literacy.

An ever-increasing group of individuals and professionals need to be statistics and data literate. Apart from statisticians, these are – to name only the most obvious ones – policy, business and data analysts; data scientists; data (protection) officers; data architects and database administrators. The skills they require range from data collection and production; to data analysis, interpretation and visualisation, trends analysis and predictive analytics; to development of data strategies, creation of institutional data cultures and monitoring of data protection regulations.

These literacy needs are yet not only relevant for professionals dealing with statistics and data. Also citizens require statistical and data literacy to assess the validity and accuracy of quantitative evidence used in political and media communication and, ultimately, to judge and control the quality of democratic political decision-making [12, 13]. As Engel correctly concluded, an "enlightened citizenry that is empowered to study evidence-based facts and that has the capacity to manage, analyze and think critically about data is the best remedy for a world that is guided by fake news or oblivious towards facts." [14] In the same vein, Sharma underlines that "[h]aving a good grasp of social statistics can help citizens deal with a complex array of issues and participate actively in public debates and assert their rights." [15] and Koga recommends that "citizens must be capable of evaluating such statistical information thoroughly before making any decisions." [16]

When it comes to defining statistical and data literacy, a uniform definition is missing [see 15, 16] and there is a huge body of literature to consult to identify its meaning [see, among many others, 13, 15 to 19]. In a narrow understanding, data literacy includes the ability to critically read, analyse, process, interpret and argue with data. A broader perspective captures knowledge, skills and value dimensions and also relates to methodological, technical, and

socio-cultural capacities. It embraces ethical concerns about data production and use; the ability to preserve and protect data; to communicate data in a contextualised manner; and to understand the quality of data, data sources, methodologies, analytics, techniques, and technologies applied to analyse data.

In view of combining aspects of statistical and data literacy, we follow the assessment of Gould who includes data literacy in an augmented definition of statistical literacy to reflect the increased role and relevance of data in everyday life. While he argues that the set of knowledge defining statistical literacy often includes a functional differentiation between “the needs of consumers of statistics from those of producers of statistics” [13], he underlines that such definition “falls far short ... of what is required for life in modern democracies.” [13] He, therefore, proposes an extended definition of statistical literacy that includes elements of data literacy, such as “understanding who collects data” and how; “understanding issues of data privacy and ownership”; “create basic descriptive representations of data to answer questions about real-life processes”; and “understanding the importance of the provenance of data” [13]. This paper reflects the extended role of both statistics and data in society and policy-making and relate to Gould’s augmented definition of statistical literacy.

### *Is statistical and data literacy specific in policy-making?*

In view of omnipresent and seemingly omnipotent technologies and advances in open data, Giovannini concludes that “the ideal of the ‘fully informed decision maker’ should be a reality.” [20] He yet also highlights that “[u]nfortunately, this is far from the case. As Einstein put it, ‘information is not knowledge’ and although citizens are bombarded by information on a constant basis, this bombardment does not necessarily bring about knowledge.” [20] The essential facilitator turning information into knowledge for policy-making and beyond is not only open access to statistics and data, it more importantly so is statistical and data literacy.

As outlined above and in a previous articles [9], the use of statistics and data in policy-making is an essential means of doing politics and an indispensable ingredient of evidence-informed policy. It is a political instrument that can support the emergence and/or preservation of power. In a metrological and post-metrological way, it impacts both the social construction of knowledge that is regarded as relevant in and for politics (in contrast to experience) and on governance processes that rely on statistics and data as sources of factual evidence. Both effects are of relevance when looking at why statistical and data literacy in policy-making is important and specific. They show that policy-making requires particular capacities to understand statistics and data in context and certain ‘meta-skills’ to identify the implications of their use in policy-making.

Looking at the knowledge effects of statistics and data in politics, those engaged in policy-making need to be able to identify the side effects of using statistics and data. They need to be sensitive to the fact that statistics and data, as quantitative approximations to reality, represent knowledge constructs themselves (*‘knowing through data’*) and that these constructs are used to prioritise the public good instead of merely representing ‘neutral facts’. Therefore, the selection of statistics and data for policy-making requires ethical reflections and justification. Actors engaged in policy-making need to understand that, apart from their descriptive (*‘observing through data’*) and diagnostic power (*‘scrutiny by data’*), statistics and data also possess normative power (*‘stipulating by data’*) in politics. They also need to understand that, within policy-making, statistics and data support political deliberation (*‘arguing through data’*), motivate and justify political prioritisation (*‘agenda-setting by data’*) and can serve as dominant political knowledge (*‘framing through data’*). They need to understand that the use of statistics and data can be influenced by politically strategic and value-based considerations

that champion adequate rather than best available quantitative evidence (*'proving through data'*). Sensitivity to potential feedback loops between the production of statistics and data and policy development is especially important to understand that quantification is political (*'data for policy'*, *'policy from data'*) and that it, therefore, needs to adhere to ethical principles and quality standards. This overall awareness needs to be coupled with the insight that statistics and data can define what they are meant to describe instead of 'only' measuring a certain reality (*'understanding data'*) and that oversimplification and difficulties to develop common understandings of complex social phenomena impact data collection and selection. These issues demand interlinkages between and responsiveness of statistics' production and the policy context, which data shall inform [9, 21]. Actors engaged in policy-making need to acquire the skills and capacity to work critically with these new forms of knowledge.

In terms of governance effects, actors engaged in policy-making require skills to analyse how statistics and data are used by whom for what purpose in what type of policy-making contexts (*'data as evidence'*). They need to be able to understand that statistics- and data-based processes impact governance (*'governing by data'*, *'governance for data'*), affect who participates in policy-making and that statistics and data directly take over different governance functions (*'collective action through data'*, *'steering by data'*, *'ruling by data'*, *'controlling with data'*, *'advocacy through data'*). Moreover, they need to be sensitive to the power dimension and impact of statistics and data in interest-formation, preference-building (*'community-building through data'*) and negotiations (*'(ab)using data'*) and to the fact that their use is interest-driven and context-dependent (*'data in evidence arenas'*). Therefore, special attention is required to the fact that statistics and data have become a target of instrumentalisation and politicisation which can endanger democratic decision-making. As analysed previously in this journal [9], Greece, Tanzania, and the World Banks' Doing Business indicators are negative examples and cases in point here.

### *What to learn for policy-making?*

The above considerations show that statistical and data literacy for policy-making goes beyond statistical-numerical-technical knowledge, skills and capacities. It is as much about understanding the context in and purpose for which statistics and data are used as it is about accepting the limitations of statistics and data. It requires a particular transparency-of-evidence-commitment and ethical standards to enhance the legitimacy of political decision-making instead of jeopardising it. Acknowledging this demand, institutional practice across levels of governance increasingly underlines the relevance of statistical and data literacy in and for evidence-informed policy-making. Within the European context, for instance, the Joint Research Centre of the European Commission's 2017 skills map for evidence-informed policy-making [22] explicitly refers to statistical and data literacy under various categories: "Communicating Scientific Knowledge"; "Evidence gap mapping"; "Digital data visualisation"; and "Infographic design". [22] As specific requirements, the map formulates "Basic skills for interpretation of quantitative and qualitative data" and "Improved understanding of statistical data in various formats and visualisations". [22]

Given the effects and perils of the use of statistics and data in policy-making, the responsibility delegated to actors engaged in politics by their electorates or institutions and the political power exerted by collective political action, statistical and data literacy is an essential prerequisite for 21st-century policy-making. Political actors need to be statistically and data literate to be able to use statistics and data for the public good. They need to understand the potential of and sensitivities arising from the use of statistics and data. They need to have the skills to identify their misuse. This means that they need to understand the relevance of interests,

preferences, motives, values, ethics and goals for the use of statistics and data in policy-making. This, in turn, obliges them to invest in their cognitive, affective, and socio-emotional competencies to raise their awareness to the context of using statistics and data in a politically, economically, socially and culturally embedded way.

## References

- [1] Baron J. A Brief History of Evidence-Based Policy. The ANNALS of the American Academy of Political and Social Science 2018;678(1):40–50.  
<https://doi.org/10.1177/0002716218763128>.
- [2] Cairney P. The politics of evidence-based policy making. Basingstoke: Palgrave MacMillan; 2016.
- [3] Head BW. Toward More “Evidence-Informed” Policy Making? Public Administration Review 2015;76(3):472–84. <https://doi.org/10.1111/puar.12475>
- [4] Greenhalgh T., Russell J. Evidence-based policymaking: a critique. Perspectives in Biology and Medicine 2009;52(2):304–318. <https://doi.org/10.1353/pbm.0.0085>
- [5] OECD. Building Capacity for Evidence-Informed Policy-Making: Lessons from Country Experiences. Paris: OECD Publishing; 2020.
- [6] Parkhurst J. The politics of evidence: from evidence-based policy to the good governance of evidence. Taylor & Francis; 2017.
- [7] Solesbury W. Evidence based policy: Whence it came and where it’s going. London: ESRC UK Centre for Evidence Based Policy and Practice; 2001.
- [8] United Nations Department of Economic and Social Affairs. Areas of Work: Evidence-based Policy | Capacity Development 2020.  
<https://www.un.org/development/desa/capacity-development/what-we-do/areas-of-work/evidence-based-policy/>.
- [9] Umbach G. Of Numbers, Narratives and Challenges: Data as Evidence in 21st Century Policy-Making, Special Feature on Governing by the Numbers – Statistical Governance. Statistical Journal of the IAOS 2020;36(4):1043-1055, <https://doi.org/10.3233/SJI-200735>.
- [10] Schüller K. Future Skills: a Framework for Data Literacy: Competence Framework and Research Report, Hochschulforum Digitalisierung. Working Paper 53, 2020.
- [11] Misra A. Advancing Data Literacy in the Post-Pandemic World: A Primer to Catalyse Policy Design and Action. Draft discussion paper for the 2021 PARIS21 Annual Meetings; Supervision: Johannes Jütting (PARIS21) and Diego Kuonen (Statoo Consulting & University of Geneva), Paris: 2021:  
[https://paris21.org/sites/default/files/inline-files/DataLiteracy\\_Primer\\_0.pdf](https://paris21.org/sites/default/files/inline-files/DataLiteracy_Primer_0.pdf).
- [12] Carmi, E., Simeon J. Y., Lockley, E., Pawluczuk A. Data Citizenship: Rethinking Data Literacy in the Age of Disinformation, Misinformation, and Malinformation’. Internet Policy Review 2020;9(2):1-22. <https://doi.org/10.14763/2020.2.1481>.
- [13] Gould R. Data Literacy is Statistical Literacy. Statistics Education Research Journal 2022;16(1):22-25, <https://doi.org/10.52041/serj.v16i1.209>
- [14] Engel, J. Statistical Literacy for Active Citizenship: A Call for Data Science Education. Statistics Education Research Journal 2022;16(1):44-49,  
<https://doi.org/10.52041/serj.v16i1.213>.
- [15] Sharma S. Definitions and models of statistical literacy: a literature review. Open Review of Educational Research 2017;4(1):118-133,  
<https://doi.org/10.1080/23265507.2017.1354313>

- [16] Koga, S. Characteristics of statistical literacy skills from the perspective of critical thinking. *Teaching Statistics* 2022;(early view)1–9, <https://doi.org/10.1111/test.12302>.
- [17] Frank, M., Walker, J. Some Key Challenges for Data Literacy. *The Journal of Community Informatics* 2020;12(3):232-235, <https://doi.org/10.15353/joci.v12i3.3288>.
- [18] Schield, Milo. 2011. ‘Statistical Literacy: A New Mission for Data Producers’. *Statistical Journal of the IAOS* 2011;27(3,4):173–83, <https://doi.org/10.3233/SJI-2011-0732>.
- [19] Gal, I. Adults’ statistical literacy: Meanings, components, responsibilities. *International Statistical Review* 2002;70(1):1–25, <https://doi.org/10.1111/j.1751-5823.2002.tb00336.x>
- [20] Giovannini, E. Statistics and politics in a ‘knowledge society’. *Social Indicators Research* 2008;86(2):177-200, [10.1007/s11205-007-9137-z](https://doi.org/10.1007/s11205-007-9137-z).
- [21] Malito D.V., Umbach G., Bhuta N. *The Palgrave Handbook of Indicators in Global Governance*. Palgrave Macmillan; 2018.
- [22] Joint Research Centre. *Skills for Evidence-Informed Policy Making: Continuous Professional Development Framework*. Brussels: European Commission, Joint Research Centre; 2017;  
[https://ec.europa.eu/jrc/communities/sites/default/files/10\\_2017\\_ec\\_jrc\\_skills\\_map\\_evidence-informed\\_policymaking\\_final.pdf](https://ec.europa.eu/jrc/communities/sites/default/files/10_2017_ec_jrc_skills_map_evidence-informed_policymaking_final.pdf).