

## DEVELOPING OFFICIAL STATISTICS LITERACY: A PROPOSED MODEL AND IMPLICATIONS

Iddo Gal<sup>1</sup> and Irena Ograjenšek<sup>2</sup>

<sup>1</sup> University of Haifa, Department of Human Services, Israel

<sup>2</sup> University of Ljubljana, Faculty of Economics, Slovenia

[iddo@research.haifa.ac.il](mailto:iddo@research.haifa.ac.il)

*This paper advocates the need to develop a conceptual framework about official statistics literacy, and the knowledge bases that adults at large and non-specialists need in order to be critically aware of social and economic phenomena by understanding key aspects of official statistics. The absence of a widely accepted framework in this regard may affect the ability of official statistics providers and statistics educators to develop effective educational programs. We thus propose a model encompassing six key knowledge elements needed for statistical literacy related to official statistics. We then discuss implications, including the need to develop a digital textbook and a modular online course focused on the elements in the proposed model, implications for research, and other directions.*

### INTRODUCTION

In recent years, increased attention has been paid to the formal training of professional statisticians who work in national and international statistical systems, and sometimes to the training of other user groups both by national and international statistical offices, as well as other producers of official statistics (see e.g. MacCuirc, 2015). However, the provision of training or development of resources related to official statistics for wider, non-professional audiences and adults at large has been largely left aside. Even educational materials designed for professionals, i.e., statistics or economics majors, pay little attention to official statistics issues (Pfeffermann, 2015).

We were able to confirm Pfefferman's 2015 finding ourselves: a literature search we conducted in early 2017 did not identify a single current textbook that describes key knowledge bases which have to be emphasized in detail when educating statistics majors about official statistics. The only relevant textbook in this regard is Citro and Straf (2013), a US-based text that focuses both on key aspirations or expectations from an official statistics provider (for example, relevance to policy issues, credibility among data users, trust among data providers, independence from political and external influences), and on numerous important administrative and organizational practices and roles (such as mission clarity, confidentiality, continuous development of useful data, openness about sources, data limitations transparency, and more). These are of course core issues for all official statistics providers around the world, yet from the content point of view, they are not related to the comprehension of the actual products of statistics providers.

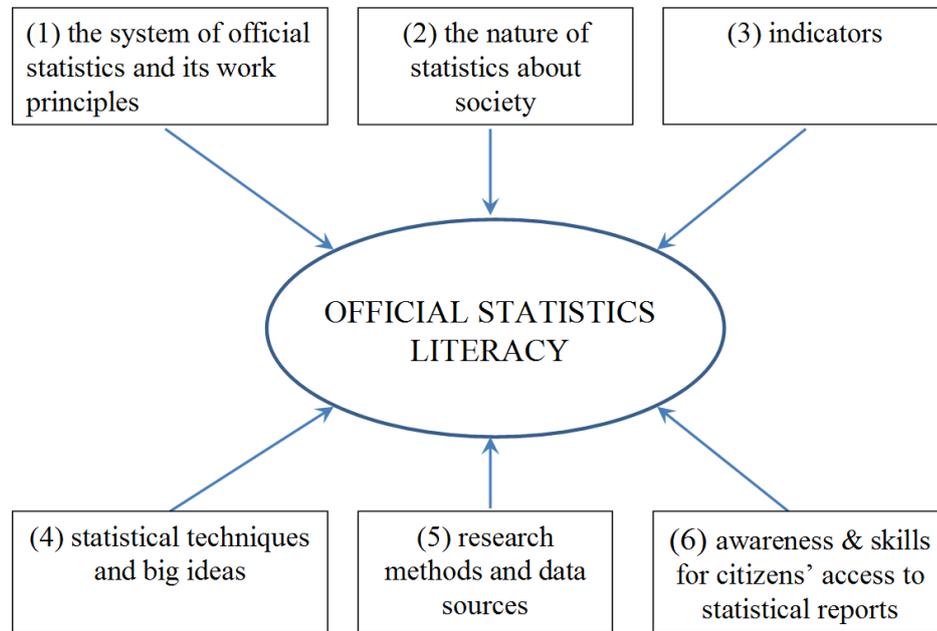
Hence, this paper focuses on specific issues that official statistics providers may face if they want to help non-specialist users or citizens develop the aspects of statistical literacy (Gal, 2002) that pertain to knowledge of, and engagement with, official statistics. For brevity, we refer to this desired knowledge base as *official statistics literacy* or *OSL*. We outline a model comprised of six possible building blocks of the desired knowledge bases that are specific to official statistics literacy. Due to space limits in this paper, here we only elaborate on three of its most critical elements. A detailed presentation of the model and its elements is available in Gal and Ograjenšek (2017). We conclude the paper by discussing selected key implications of our proposed model for official statistics providers and statistics educators, as well as for future research.

### A PROPOSED MODEL OF OFFICIAL STATISTICS LITERACY

The proposed model is outlined in Figure 1 and includes six elements which we argue are unique or encompass specific knowledge bases and skills that citizens at large and non-specialists, as well as specialists in statistics, need in order to make sense of official statistics, when they engage in their everyday or work lives with statistical information from official sources. Such information may be encountered, e.g., in digital media, when listening to news on TV, reading a

magazine article, or browsing the Internet and seeking information on topics of interest. A specific point of comparison pertains to the knowledge expected of students who have taken an introductory statistics course at the undergraduate level, which may be the last, and for some students the only, structured exposure to statistics (Moore, 1998; Meng, 2009).

Fig. 1. Proposed model of six building blocks (areas) of official statistics literacy



The logic behind identification of all six models and their full details can be found in Gal and Ograjenšek (2017). In this paper, given space constraints, we briefly elaborate on three elements (2, 3 and 5 in the model in Figure 1) which are most relevant for the discussion of educational ramifications and future collaborative action we would like to see initiated among official statistics providers and statistics educators:

- The nature of statistics about society (2nd element in the model)*: This element is based on new work by the ProCivicStat project, a collaboration by six universities in five countries (Germany, Hungary, Israel, Portugal, and the United Kingdom) funded by the European Commission's ERASMUS+ program (see [www.procivic.stat.org](http://www.procivic.stat.org)), which aims to promote civic engagement and understanding among young adults regarding 'civic statistics' about key societal phenomena. As part of the ProCivicStat work, Engel, Gal, and Ridgway (2016) claim that to be fully engaged, citizens need to understand 'civic statistics' with regard to past trends, present situations, and possible future changes in diverse areas of importance to society such as demographics, employment, wages, migration, health, poverty, access to services, education, human rights, and other domains. The ProCivicStat analysis argues that civic statistics have five general characteristics: they relate to (1) *multivariate phenomena* that are often intercorrelate and have non-linear relationships, and often to (2) data that are both analysed and reported in various levels of *aggregation or disaggregation*. They also involve (3) *dynamic data* that change or is updated over time. Furthermore, since data and findings about social phenomena are multivariate, aggregated at multiple levels, and dynamic, their description across time or comparison units requires the use of diverse types of representations. Hence, they may often be delivered through (4) *rich texts* and (5) *rich visualizations* that are broader in scope and at times more sophisticated compared with the limited range of ideas and representations included in introductory statistics classes.
- Indicators (3rd element in the model)*: Official statistics providers create key messages to decision makers and to the general public regarding levels or changes in dozens of

indicators (Horn, 1993), such as about unemployment level, child mortality, gross domestic product, educational attainment, unemployment, or income inequality (e.g., Gini coefficient). These and many other indicators in use by official statistics providers (e.g., “consumer price index”, “better life index”), are often not raw variables, such as those encountered in introductory statistics, but rather combinations of data elements that may be expressed as percentages, ratios, or numbers on arbitrary scales. The indices may be computed as simple rates, or be derived as complex aggregates of weighted elements. They may be based either on *objective data* (e.g., “consumer spending”) or *subjective data* (e.g., “consumer confidence”), and their definitions, underlying methodology, or modes of measurement may develop and change over time to reflect society’s needs for information (As an example, see a recent review of social indicators by Land & Michalos, 2017).

Indicators are widely used by official statistics providers to report on a wide range of issues, hence their understanding is critical for all citizens. Although they are seemingly included in the broad description of the prior element “the nature of statistics about society”, we highlight indicators as a separate aspect of official statistics, because while they are prevalent in public and political discourse, indicators are hardly ever described or analyzed in textbooks and statistics curricula for non-specialists (though see Haack, 1979), or in resources related to teaching research methods (Gal, 2007).

- *Research methods and data sources (5th element in the model)*: We first note that knowledge bases related to methodological issues are often spread between the discipline of statistics and the domain loosely called ‘research methods’. There is an overlap between them (Gal, 2007; Meng, 2009). Consequently, there are long-standing debates as to where statistics ends and research methods begin. What statisticians view as fitting under ‘methodology and enquiry processes’ may only cover some elements of what experts from other disciplines may have in mind when they discuss research methods (Gal & Ograjenšek 2010; Ograjenšek & Gal, 2016). Hence, in many academic departments the teaching of research methods is provided by discipline-based instructors and is not part of service courses in statistics. Furthermore, at university, the learning of research methods is spread over multiple degree levels (e.g., undergraduate, graduate, doctoral), and is organized in diverse ways across different academic institutions and departments (Deem & Lucas, 2006).

Regardless of the existing diversity in how research methods (and statistics) are taught, however, the logic of the statistical enquiry process (Wild & Pfannkuch, 1999) or the PPDAC (problem, plan, data, analysis, conclusion) cycle (MacKay & Oldford, 2000) is likely to be encountered when students learn research methods. Consequently, some students may learn about surveys vs. experiments, sampling and randomization, some aspects of measurement or questionnaire design, or sources affecting internal and external validity of different research designs. In contrast, official statistics providers make use of a wider range of data sources and methods for data collection. Examples include the use of a national census, the increasing role of administrative records or public registers, and the many potential types of ‘big data’ (Daas et al., 2015) that accumulate from sources that fall outside the traditional distinction between surveys and experiments. Further, even when samples are used by official statistics providers, they are usually utilized on a large scale or a cycling basis (e.g., social surveys, employment surveys, employer-based or enterprise surveys) and involve weighting issues if a whole country or sector is to be represented. Given the repeated nature of many official surveys or data-collection efforts and the high-stakes nature of the findings derived from them, issues related to various error sources such as sample design, non-response, or respondent bias that determine data quality or credibility, receive much attention in official statistics.

## DISCUSSION AND PROPOSED COURSES OF ACTION

In recent decades, discussions in the professional literature about the connections between official statistics providers and statistics educators (e.g., Gal, 2002; chapters in Sanchez, 2008)

have focused in large part on how official statistics providers can facilitate improvement of generic statistics education at the school or university level. Within this framework, official statistics providers have been contributing to teachers' professional development by offering datasets, lesson plans, ideas for projects and poster competitions, and other resources that can inform class activities or highlight the importance of official statistics. Further, de Smedt (2016) describes directions related to enhancing support and explanations for users regarding interpretation of displays and published statistics, and other options.

Going beyond the directions for action noted above, in Gal & Ograjenšek (2017) we discuss various implications of our proposed model of OSL with its six elements. It is obviously important to continue existing collaborations between official statistics providers and school-level educators as noted by sources discussing the development of statistical literacy at school level as illustrated above. However, we believe the vision of systematically promoting official statistics literacy for non-specialists and for the general adult population (including actions in countries with characteristics that differ from the few that have spearheaded educational services and activities in statistics education) requires an examination of additional directions from a long-range future collaborative perspective. Our analysis of the literature underlying the proposed model indicates that unique efforts are needed to promote official statistics literacy, since regular statistics education normally does not highlight the unique aspects of elements in our model, such as the nature of data about society, the complex nature of indicators, etc.

With the above in mind, we outline below several possible directions and some additional ideas that could be implemented in the international collaborative setting:

1. *Textbook + Digital teaching modules.* We propose the development of a textbook on official statistics geared towards statistics majors as well as non-majors who may study selected topics in statistics, that focus on the six elements included in the proposed model. We note that there are many more non-majors than majors who take only introductory statistics, and the provision of an accessible textbook may be the first step to helping educational institutions develop new modules or whole courses related to official statistics that are currently lacking. Further, we propose the development of an MOOC or a collection of digital (video and audio) teaching modules for entry-level majors, non-majors, and other groups of interest among the general public. It is hard to expect a single official statistics provider to shoulder responsibility and allocate resources related to both initiatives outlined above, although it would be technically possible. Important ideas in this regard have been presented by Harraway and Forbes (2013), who have also developed distance-education modules on some topics in New Zealand.

Both initiatives thus call for an international collaborative effort of official statistics providers, statistics educators, specialists in applied fields that rely on official statistics when discussing major concepts inherent to their disciplines, and other stakeholders. Such an effort can, of course, benefit from existing materials and frameworks developed in the context of existing diploma and degree programs listed in the previous sections of this article. Textbook developers participating in this collaborative effort could build on experiences gained within the framework of the Phare project, which resulted in the modular online Course on the European Economic Statistics (Bregar et al., 2000). In addition, large professional associations with an international outreach and long-standing interest and activities in statistics education can also facilitate collaborations and the long-term development of a textbook and a MOOC. Key actors may be the International Statistical Institute (ISI) and its relevant divisions (the International Association for Statistics Education - IASE and the International Association for Official Statistics - IAOS) as well as the Royal Statistical Society (RSS), the American Statistical Association (ASA), and others.

2. *Content of instruction.* While the directions discussed above relate to ideas involving official statistics providers, there is much to rethink regarding the topics included in introductory statistics courses, if one aims to also develop students' official statistics literacy as envisioned in this paper and in Gal and Ograjenšek (2017).

In this brief paper we only have room to highlight two examples that stem from our proposed model and illustrate areas that go beyond the content of regular introductory courses:

- It is essential to pay more attention to *multivariate phenomena*, including sub-issues such as co-variation or interactions over time, aggregation and disaggregation of response variables within different population subgroups, and so forth. Strong arguments regarding the need to attend to multivariate thinking and confounding of variables have been presented by Schield (e.g., 2016), based on cumulative work and experiences related to the Statlit Project. Further, based on work at the ProCivicStat project, Engel et al. (2016) present additional ideas in this regard, and additional suggestions and lesson plans and related resources can be found at [www.procivicstat.org](http://www.procivicstat.org).
- *Indicators* are a key aspect of official statistics, yet receive little attention in regular textbooks and curricula for introductory statistics. The lack of attention to indicators may in part be a result of the fact that some indicators are comprised of *qualitative* variables or derived in part via qualitative methods, or otherwise their meaning and interpretation may require qualitative thinking. Possibly, such qualitative issues are not receiving the same attention from statistics educators, compared to quantitative variables and quantitative issues. Hence, Ograjenšek and Gal (2016) point to various ways to increase the attention to qualitative variables during instruction, which can serve both the understanding of indicators as well as of key methodological issues.

3. *Future research*. Finally, it is essential to point to the need for further research on multiple fronts. First, as part of establishing the base for developing an international textbook and accompanying MOOC elements proposed above, there is a need for comparative study of official statistics curricula elements across the globe both at secondary and tertiary level, as useful resources may exist in languages other than English. This could help shorten the path to creating solid materials. Further, there is a need to embark on a consensus-building process, perhaps using a Delphi methodology, in order to ensure that the textbook and digital teaching modules reflect the perceptions of key stakeholders regarding needed content, curricular sequencing, etc.

In addition, the introduction of new educational tools or directions, whether regarding multivariate phenomena, indicators, or other aspects highlighted in our model of OSL, has to be accompanied by more detailed research. For example, there is a need to examine the perceptions, value judgments, and reactions of learners and instructors alike, as well as the many ways in which learners operate within new digital learning spaces that are now an integral part of the landscape of learning statistics, and are especially important when students will be expected to interact with products of official statistics providers, or with new types of visualization tools. As Gal and Ograjenšek (2010) argue, research that aims to support learners and users of statistics has to combine both quantitative *and* qualitative research methods, in order to provide effective information to the agencies or actors that develop statistical information products.

#### ACKNOWLEDGEMENT

The first author was partially supported by a grant from the ERASMUS+ program of the European Commission for the *ProCivicStat* project. However, the opinions expressed in this article reflect the author's own views and not necessarily those of the sponsoring agency.

#### REFERENCES

- Bregar, L., Ograjenšek, I., & Bavdaž, M. (2000). Teaching economic statistics in a digital environment. In Ferligoj, A. and A. Mrvar (Eds). *New approaches in applied statistics*, 237-249, <http://mrvar.fdv.uni-lj.si/pub/mz/mz16/abst/bregar.htm>
- Citro, C. F., & Straf, M. E. (eds.). (2005). *Principles and Practices for a Federal Statistical Agency* (3rd ed). National Research Council.. Washington: National Academy Press.
- Daas, P.J., Puts, M.J., Buelens, B., & van den Hurk, P.A. (2015). Big data as a source for official statistics. *Journal of Official Statistics*, 31(2), 249-262.
- Deem, R., & Lucas, L. (2006). Learning about research: Exploring the learning and teaching & research relationship amongst rducational practitioners studying in higher education. *Teaching in Higher Education*, 11(1), 1-18.
- de Smedt, M. (2016). European statistics and eurostat's contribution to improving statistical literacy.

- In J. Engel (ED.), *Proceedings, IASE Roundtable on Promoting Understanding of Statistics About Society*, Berlin, [http://iase-web.org/Conference\\_Proceedings.php?p=Promoting\\_Understanding\\_of\\_Statistics\\_about\\_Society\\_2016](http://iase-web.org/Conference_Proceedings.php?p=Promoting_Understanding_of_Statistics_about_Society_2016).
- Engel, J., Gal, I., & Ridgway, J. (2016). Mathematical literacy and citizen engagement: the role of civic statistics. Paper presented at the *13th International Congress on Mathematics Education (ICME13)*.
- Gal, I. (2007). Research methods: reflections on teaching frameworks and research. In *Learning & Teaching of Research methods at University*. Murtonen, M., Rautopuro, J., & Väisänen, P. (Eds). Turku, Finland: Finnish Educational Research Association.
- Gal, I. (2002). Adult statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Gal, I., & Ograjenšek, I. (2017). Official statistics and statistics education: Bridging the gap. *Journal of Official Statistics*, 33(1), 79–100.
- Gal, I., & Ograjenšek, I. (2010). Qualitative research in the service of understanding learners and users of statistics. *International Statistical Review*, 78(2), 287-298.
- Haack, D. G. (1979). *Statistical Literacy: A Guide to Interpretation*. Duxbury Press/Wadsworth.
- Harraway, J.A., & Forbes, S. D. (2013). Partnership between national statistics offices and academics to increase official statistical literacy. *Statistical Journal of the IAOS*, 29(1), 31-40.
- Land, K. C., & Michalos, A. C. (2017). Fifty years after the social indicators movement: Has the promise been fulfilled?. *Social Indicators Research*, 1-34.
- MacCuirc, E. (2015). You don't teach, students learn: lessons learned in statistical literacy and statistical education in Ireland. *Austrian Journal of Statistics*, 44(2): 73-83. Retrieved from <http://www.ajs.or.at/index.php/ajs/article/view/62>.
- MacKay, R.J., & Oldford, R.W. (2000). Scientific method, statistical method and the speed of light. *Statistical Science*, 15(3), 254–278.
- Meng, X. (2009). Desired and feared—what do we do now and over the next 50 years? *The American Statistician*, 63(3), 202–210.
- Moore, D.S. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association*, 93(444), 1253–1259.
- Ograjenšek, I., & Gal, I. (2016). Enhancing statistics education by including qualitative research. *International Statistical Review*, 84(2), 165–178. Doi: <http://dx.doi.org/10.1111/insr.12158>.
- Pfeffermann, D. (2015). Methodological issues and challenges in the production of official statistics: 24th annual Morris Hansen Lecture. *Journal of Survey Statistics and Methodology*, 3(4), 425-483.
- Sanchez, J. (ed.) (2008). *Government Statistical Offices and Statistical Literacy*. International Statistical Literacy Project and ISI, <http://iase-web.org/islp/Publications.php?p=Books>
- Schild, M. (2016). GAISE 2016 promotes statistical literacy. *Statistics Education Research Journal*, 16(1), 46-50, <http://iase-web.org/Publications.php?p=SERJ>
- Horn, R. V. (1993). *Statistical Indicators: For the Economic and Social Sciences*. Cambridge University Press.
- Wild, C.J., & M. Pfannkuch. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–248.