Official Statistics and Statistics Education: Bridging the Gap

Iddo Gal¹ and Irena Ograjenšek²

This article aims to challenge official statistics providers and statistics educators to ponder on how to help non-specialist adult users of statistics develop those aspects of statistical literacy that pertain to official statistics. We first document the gap in the literature in terms of the conceptual basis and educational materials needed for such an undertaking. We then review skills and competencies that may help adults to make sense of statistical information in areas of importance to society. Based on this review, we identify six elements related to official statistics about which non-specialist adult users should possess knowledge in order to be considered literate in official statistics: (1) the system of official statistics and its work principles; (2) the nature of statistics about society; (3) indicators; (4) statistical techniques and big ideas; (5) research methods and data sources; and (6) awareness and skills for citizens’ access to statistical reports. Based on this ad hoc typology, we discuss directions that official statistics providers, in cooperation with statistics educators, could take in order to (1) advance the conceptualization of skills needed to understand official statistics, and (2) expand educational activities and services, specifically by developing a collaborative digital textbook and a modular online course, to improve public capacity for understanding of official statistics.

Key words: Statistical literacy; skills and competencies; official statistics literacy; dissemination; adult education.

1. Background and Motivation

In recent years, both national and international statistical offices as well as other producers of official statistics (hereafter: official statistics providers) have been paying increasing attention to the formal training of professional statisticians who work in national and international statistical systems, and sometimes to the training of other user groups. Programs awarding either a diploma or a full academic degree related to official statistics are offered by several intergovernmental institutions or networks, such as the European Master in Official Statistics (EMOS; Zwick 2016), the Statistical Institute for Asia and the Pacific (SIAP), and the University of the South Pacific. Several national statistical offices (some via institutional collaboration) are very active in this regard as well. For example, in New Zealand, a postgraduate course in official statistics is offered that covers areas such as data visualization, confidentiality, geographic information system, demography, health

¹ University of Haifa, Department of Human Services, Abba Houshi Road 199, Haifa, 3498838, Israel. Email: iddo@research.haifa.ac.il
² University of Ljubljana, Faculty of Economics, Kardeljeva pl. 17, Ljubljana, 1000, Slovenia. Email: irena.ograjensek@ef.uni-lj.si (corresponding author)

Acknowledgments: 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.
statistics, and economic statistics (Harraway and Forbes 2013), in collaboration with New Zealand’s National Certificate of Official Statistics (Forbes and Keegan 2016). The Central Statistics Office Ireland (MacCuirc 2015) or Statistics Finland (Helenius and Mikkela 2011) have also developed training modules, or full diploma programs, for specific target groups of users such as government employees and analysts, business managers, or journalists, who are usually not statisticians but who work with official statistics in various ways.

This article focuses on a gap in the world of formal training in official statistics, pertaining to wider, non-professional audiences. These include, among other groups within the adult public at large, the many educators who may teach non-specialists about statistics (for example, lecturers in introductory statistics at the undergraduate level in many different disciplines and departments, mathematics teachers who also teach statistics at the high school level), their many students (who would soon be adults and enter the workforce), or various administrators, and managers in diverse sectors.

On the one hand, official statistics providers are interested in increasing the use of their information products through multiple user groups that include the general public. They are taking many steps to improve the quality of their information services: they have been opening up free access to their information products through digital portals, and have been continuously seeking ways to improve levels of public trust and confidence in official statistics, as well as the level of satisfaction with their information products (Biemer et al. 2014; Steenvoorden et al. 2015).

On the other hand, the provision of training or resources related to official statistics for wider, non-professional audiences, has been largely left aside. Even if official statistics websites are being made more user-friendly, comprehension of the statistical information in them is far from optimal (Schield 2011). Very few official statistics providers offer structured materials designed to enable the public, or stakeholders from the education sector (i.e., teachers and students), to better understand official statistics on their websites. Even leading national statistical offices such as Statistics Canada or the Australian Bureau of Statistics have cut down on their support to statistics education at schools over the last few years.

The gaps noted above also exist within the professional field of statistics. Official statistics providers have been operating for decades around the world, and represent an indispensable element in the information system of a democratic society (United Nations 2014). However, a dire and surprising lack of solid educational materials designed for professionals (i.e., statistics or economics majors entering careers in official statistics) has been noted by numerous scholars involved in the training of statisticians (Murphy 2002; Nathan 2007). Pfeffermann (2015) has recently reviewed curricula of statistics departments at over 20 leading universities and concluded that most departments pay little attention to formal instruction in key aspects of official statistics (such as survey sampling, seasonal adjustment, or national accounts). Given that official statistics is a prime employment area for statistics graduates, this is a very surprising finding.

Furthermore, a literature search we conducted did not find a single current textbook that describes key knowledge bases that have to be emphasized in detail when educating statistics majors about official statistics. Over 20 years ago the modular online Course on European Economic Statistics (CEES) was developed with the support of Eurostat
(in cooperation with the Institute for Training of European Statisticians – TES) within the Phare Multi-Country Co-Operation for Distance Education Programme (Bregar et al. 2000). Unfortunately, it seems to have been published too early to be adopted by traditional universities, most of which at that point in time had not begun to recognize digitalization as the future of educational systems. Lacking solid marketing support at its launch, the course therefore remained a short-lived attempt to fill in the gap that was identified decades ago and still exists today.

The only book currently available that appears to be dedicated to the role of an official statistics provider is Citro and Straf (2013), which is also in use by the EMOS programme. This US-based text focuses 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 core issues for all official statistics providers around the world, yet they are not related to a comprehension of the actual products from the contents’ point of view. Consequently, the text should be regarded as a very incomplete basis from which to define what non-specialists need to know to understand official statistics products.

The situation described above implies that educators who wish to introduce non-majors, high school students, business graduates or adults in general to the fundamentals of official statistics do not have a set of suitable resources geared for their needs, even at the beginning of the twenty-first century. If one accepts the tenet that citizens should know something useful about official statistics, many questions arise: first of all, the question of “what are the basics that citizens (or non-specialists) should know about official statistics?”.

While this question seems simple, the answer is not straightforward. It has not been discussed in detail in the professional literature on official statistics; and certainly not with regard to non-majors and adults at large. Other related questions are “whose responsibility is it to develop materials on official statistics for non-specialists?”, and “to what extent (if at all) should official statistics providers divest resources in order to increase public knowledge of official statistics?”.

Our goal in this article is to assist, but also to challenge official statistics providers to ponder the questions raised above. We focus our contribution on specific issues that official statistics providers may face if they want to help non-specialist users 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). To this end, in our article we first briefly review the general ecology of skills and competencies that adults may need in order to make sense of statistical information regarding societal matters. We then examine possible building blocks of the desired knowledge base that is specific to OSL in more detail. Based on this conceptualization, we then discuss some directions for future developments that official statistics providers could make in order to contribute to educational efforts aimed at increasing official statistics literacy, thereby enriching the course for the development of statistics education in general.
2. On Quantitative Competencies and Literacies

A discussion of the statistical capacity needed to understand and engage with official statistics requires that we first describe the larger environment within which understanding of official statistics (by non-specialists) is situated.

Over the last few decades, the academic and professional literature has identified several related but separate constructs that describe general competencies that adults should possess in order to effectively cope with the quantitative demands of the adult world, including those related to statistics and probability. Key constructs that have so far been defined and discussed at some length are (adult) numeracy and mathematical literacy (Gal et al. 2005; PIAAC Numeracy Expert Group 2009; Geiger et al. 2015; Stacey 2015; Tout and Gal 2015), quantitative literacy and quantitative reasoning (Steen 2001; Madison 2014; Karaali et al. 2016), as well as statistical literacy and probability literacy (Gal 2002, 2005; Watson 2016). Separate constructs such as health numeracy (Ancker and Kaufman 2007), scientific literacy (Rutherford 1997), financial literacy (Lusardi and Mitchell 2014; Xu and Zia 2012) or media literacy (Coddington 2015) also encompass, among other components, diverse quantitative skills which incorporate understanding of specific types of statistics and data collection methods. Examples include an understanding of long-range trends in the economy or in ageing which affect pensions or poverty levels; risk estimates associated with health conditions, pollution levels, and mortality rates; or notions of (the strength of) evidence.

The usage of ‘literacy’ when coupled with a term denoting an area of human activity (e.g., ‘statistical literacy’) may conjure an image of a minimal subset of basic skills expected of all citizens in this area, as opposed to a more advanced set of skills and knowledge that only specialists may achieve. Yet, many scholars warn against such a restrictive interpretation, and argue that “literacy’, when used to describe people’s capacity for goal-oriented behavior in a specific domain, suggests a complex cluster of skills that may range on a continuum from very low to very high; and furthermore, that such skills involve not only certain formal and informal knowledge, but also desired beliefs and attitudes, habits of mind, and a critical perspective (Gal 2002; Geiger et al. 2015). This has already been recognized in the area of mathematics education, where conceptions of mathematical literacy (Kilpatrick 2001) or quantitative literacy (Steen 2001) have extended the definitions of the mathematical knowledge desired of school graduates, in light of the complex nature of everyday situations adults have to understand and manage.

The literacies pertaining to the area of statistics usually fall under the umbrella term statistical literacy, though there are several related constructs, such as probability literacy (Gal 2005), data literacy, or risk literacy. According to Gal (2002), statistical literacy refers to people’s ability to interpret, critically evaluate, and (when relevant) express their opinions regarding statistical information, data-related arguments, or stochastic phenomena. He further argues that statistically literate behavior requires the joint activation of dispositions (supporting motivation, positive attitudes, and a critical stance), coupled with five cognitive knowledge bases: literacy skills, statistical knowledge (also including some knowledge of probability, albeit informal), mathematical knowledge, contextual or world knowledge, and knowledge of critical
questions that have to be asked. Watson (2002), as well as Watson and Callingham (2003), have described three levels that reflect increasing degrees of sophistication in statistical literacy and can be viewed as a developmental trajectory through which learners may progress: (1) basic understanding of probabilistic and statistical terminology; (2) understanding of statistical language and concepts when they are embedded in the context of wider social discussion; (3) and the ability to apply a questioning attitude to statistical claims and arguments.

Several of the constructs described above have also been defined and evaluated as part of large-scale international comparative assessments and, in some countries, as part of national assessments. Consequently, we do have some information about proficiency distributions. For instance, results from the Programme for International Assessment of Adult Competencies (PIAAC, also referred to as the OECD Survey of Adult Skills) for 33 countries that participated in the first two waves of this comparative assessment show that in the area of adult numeracy, a large percentage of adults in most countries, usually between 20–40%, has low or very low numeracy skills. In most countries few adults (less than eight percent) reach the highest proficiency levels possible in the assessment, though there is considerable variation around these general patterns at the country level (OECD 2013b).

International comparative data that shed some light on knowledge and skills of adults in the specific area of statistical knowledge comes mainly from the OECD’s Programme for International Student Assessment (PISA). While PISA assesses proficiencies of students aged 15–16 years, it shares many similarities with the PIAAC assessment of adult proficiencies (Tout and Gal 2015), both in terms of its conceptual framework and its use of assessment items that purport to simulate real-world demands facing future adults. Specifically, the PISA 2003 and PISA 2012 assessment cycles have reported separate findings in four subareas of mathematical literacy, one of which is ‘uncertainty & data’ (i.e., statistics & probability). In PISA 2003 (OECD 2004), whose test-takers now approach 30 years of age and thus classify as adults, results were reported for six levels of proficiency, from 1 (lowest) to 6 (highest), including a seventh group of ‘below level 1’. At the risk of oversimplifying the complex pattern of reported results, the findings suggest that, across all 25 participating countries, on average, 46% of the respondents did not reach level 3, showing poor ability to read and interpret statistical displays and statistical messages that involve more than a few straightforward data elements. A similar pattern was reported in PISA 2012 (OECD 2013a), whose participants are now aged around 18 years.

Results concerning numeracy (in PIAAC) and mathematical literacy (in PISA) proficiencies thus suggest that in many countries the adult population is very diverse in terms of its ability to comprehend quantitative and statistical messages. Further, PIAAC also shows similar patterns regarding other skills that are involved in finding, understanding and engaging official statistics, in particular reading literacy and the ability to solve problems in [information] technology-rich environments. It is, of course, possible that quantitative and statistical competencies at the individual level change (even evolve positively) over time. Nevertheless, when viewed together, findings and gaps documented by PISA and PIAAC, motivate and inform further dialogue about ways to conceptualize, and in turn improve, official statistics literacy.
3. Towards a Definition of Official Statistics Literacy


In this section we reflect on what are the unique or specific knowledge bases and skills that citizens at large and non-specialists need in order to make sense of official statistics in addition to having the knowledge bases and skills subsumed under the more generalized constructs reviewed in the previous section. 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).

The traditional content of introductory courses for non-specialists is reflected in the table of contents of basic statistics textbooks. There is no single structure for an introductory course across textbooks and disciplines, and even well-established series (such as Freedman et al. 2007; or Moore 2012) change some of their internal elements over time. That said, a typical introductory course for non-majors may cover a mix of ideas and techniques related to topics such as:

- the purpose of statistics,
- descriptive statistics (e.g., measures of center and spread), normal curve and distributions such as z and t,
- some graphing,
- notions of association and correlation as well as some regression,
- sampling and sampling error,
- basic ideas concerning probability and binomial distribution,
- basics of statistical inference (including expected values, confidence intervals and simple statistical significance tests),
- and possibly other subtopics such as data collection methods (surveys and experiments), measurement and questionnaire design.

Not surprisingly, the contents of an introductory course and related teaching approaches, have been the subject of expert analysis over several decades, in the United States in particular. Numerous scholars have debated the sequencing as well as relative importance and weight of some components (Moore and Cobb 2000; Chance and Rossman 2001; Cobb 2007; Malone et al. 2012). There are calls to change the balance between conceptual understanding and computations or the use of technology, along with the need to deepen understanding of big ideas in statistics via the use of randomizations or simulations (e.g., Tintle et al. 2015), for examining alternative approaches to teaching (Vehkalahti 2016), or for expanding the attention to qualitative ideas in statistics (Ograjenšek and Gal 2016).

There is a plethora of introductory textbooks and scholarly interest in, and debates on, the content of introductory statistics courses for university and high-school students. However, there are virtually no scholarly debates or sources that provide an integrative view of basic knowledge elements regarding official statistics expected of the same students, and adults at large. In this context, we note the work by the United Nations Economic Commission to Europe (UNECE), which, as part of its efforts to improve good practices for communicating and using official statistics, has also aimed to define general
knowledge elements in statistics required of decision-makers and citizens. The UNECE (2012) proposed four primary areas:

1. data awareness,
2. ability to understand statistical concepts,
3. ability to analyse, interpret and evaluate statistical information, and
4. ability to communicate statistical information and understandings.

We believe that these areas are generally important, but not sufficiently specific to the area of official statistics.

To contribute to further thinking in this regard, we have reviewed and integrated information from the references mentioned in this section so far, along with references from the following three types of sources:

- Syllabi of established programs that impart either graduate degrees or diplomas related to official statistics (e.g., by Central Statistics Office Ireland or Statistics Finland); selected key publications on the websites of national official statistics providers active in statistics education (e.g., Australian Bureau of Statistics, Statistics New Zealand, Statistics Canada); and texts from Eurostat’s Statistics Explained website.

- Preliminary insights from ProCivicStat, a new collaboration effort by six universities in five countries (Germany, Hungary, Israel, Portugal, and the United Kingdom) funded by the European Commission’s ERASMUS+ program. The project (see http://community.dur.ac.uk/procivic.stat) aims to promote civic engagement and understanding among young adults regarding ‘civic statistics’ about key societal phenomena (Engel et al. 2016). Among other things, the consortium of partners has analyzed the cognitive demands of texts and displays in publications of official statistics providers, news media, and other stakeholders, and is developing a new framework regarding skills and attitudes needed to understand civic statistics and related teaching resources.

---

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

While a full analysis of information from these diverse sources is still in progress, at this interim stage we can propose a new model, depicted in Figure 1 and explained later in this section in more detail.

The model encompasses six elements about which non-specialists and adults in general should possess knowledge to be considered literate in official statistics:

(1) the system of official statistics and its work principles,
(2) the nature of statistics about society,
(3) indicators,
(4) statistical techniques and big ideas,
(5) research methods and data sources, and
(6) awareness and skills for citizens’ access to statistical reports.

3.2. The System of Official Statistics and Its Work Principles

Adults can be expected to know that their country has a system of statistics producers or official statistics providers that work cooperatively on the basis of fundamental principles (United Nations 2014). These official statistics providers aim to make data and diverse information products available to keep policy-makers, various user groups, and the general public apprized of the current economic and social situation. Their aim is also to facilitate description of changes over time (historical analysis) and to create predictions (e.g., population projections) in order to anticipate future trends for a wide range of topics relevant to society.

Towards these goals, official statistics providers employ scientific principles, accepted procedures and standards, as well as quality criteria for data collection, analysis, reporting, release, and dissemination (Biemer et al. 2014). Official statistics providers aim to collect, analyze data and report findings in an impartial and ethically sound way, and work in ways that create and retain public trust and confidence in the national statistical system (Holt 2008).

We argue that citizens may also need to know about seemingly more technical aspects of the broad statistical system that affect how, and what types of, statistics are reported to the public. For example, the fact that official statistics providers release certain statistics (e.g., regarding economic indicators such as the CPI, the GDP, or population statistics) using prescribed release schedules; that they may revise and correct already published findings due to methodological or other considerations; or that they have to use international standards for collecting and reporting key statistics in order to enable comparability across societies. These, and related details about the statistical system, are normally not included in introductory statistics courses, yet are essential for adults to understand, in general terms, where official statistics come from, how they are produced and reported, and why they are produced and reported in specific ways.
3.3. The Nature of Statistics about Society

Based on work by the ProCivicStat project noted earlier, Engel et al. (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, crime, poverty, access to services, energy, education, human rights, and other domains. The ProCivicStat analysis points to five general characteristics of civic statistics and the ways in which they are reported to the public:

- **Multivariate phenomena.** Data about social variables of interest usually do not stand in isolation; their description and understanding involves other variables that are correlated, interact with each other, or have non-linear relationships.

- **Aggregated data.** Statistics about society are often reported not with regard to continuous raw variables per se, but involve data that are grouped in diverse ways, sometimes using qualitative variables (Ograjenšek and Gal 2016). Thus, data may be combined into indicators, or reported for multiple subgroups.

- **Dynamic data.** Civic statistics are often not the result of a one-time data collection effort (e.g., unlike a single survey discussed in an introductory statistics course) but based on data collected periodically (e.g., each month, quarter, year) or on a comparative basis (e.g., in multiple countries). Consequently, data are often reported as a trend over time, and may be updated when new data become available or old data are re-evaluated, leading to the creation of an information space and displays that are more complex and dense compared to the simplified data used in teaching introductory statistics.

- **The use of rich texts.** Statistical information about society is brought to the public mainly via texts published by statistics producers (e.g., press releases or brief reports) or via articles in the media. Thus, text is a primary medium for communicating statistics (Gal 2002), and the public needs to be capable of comprehending and critically interpreting different genres of writing, such as formal language used in official reports, journalistic writing, and more.

- **Diverse visualizations.** Since data and findings about social phenomena are multivariate, dynamic, and aggregated, their description across time or comparison units requires the use of diverse types of representations. Hence, today users encounter a range of static, dynamic, and interactive visualizations (Ridgway 2016) that are much broader and more sophisticated compared with the limited range of graphs and histograms used in introductory classes.

The five broad characteristics of ‘civic statistics’ outlined by the ProCivicStat project influence the nature of the data and statistical messages from statistics providers that reach the general public and non-specialist user groups, albeit in different ways. Information products describing statistics with the above characteristics are made available to the public via multiple channels, including traditional (printed and visual) media, social media, private entities such as NGOs, advocacy groups, independent research centers, and other information or data intermediaries (e.g., bloggers). These ‘secondary players’ usually present only selected aspects of the original publications or findings, and may
sometimes re-analyze or present them in ways that aim to explore specific topics of social or political significance or advance specific points of view. Some findings may also be redistributed via social or digital networks and discussed by private citizens, NGOs, or academic instructors, outside the purview of the original producers.

3.4. Indicators

What kind of official statistics are conveyed to the public (and to policy-makers) via media channels? The answer is complex, of course, as many types of findings and insights are shared, and their flavor may change across topics or countries. Yet all official statistics providers create messages regarding levels or changes in dozens of indicators, such as unemployment level, child mortality, gross domestic product, or income inequality (e.g., Gini coefficient), that reflect the state of some aspect of our society, economy, or well-being.

These and many other indicators in use by official statistics providers 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. They may be computed or derived, from simple rates to complex aggregates of weighted elements. They may be based either on objective (e.g., consumer spending) or subjective data (e.g., consumer confidence), and their definitions may develop and change over time to reflect society’s needs for information about itself. However they are defined, indicators are widely used by official statistics providers to report on a wide range of issues, and their understanding is essential for all citizens.

Although they are seemingly included in the broad description of the prior aspect regarding the nature of statistics about society, we highlight indicators as a separate aspect of official statistics because of their privileged role in public discourse and as a key product category that may influence policy-makers. Yet, surprisingly, despite their centrality in society and their prevalence in public and political discourse, indicators are hardly ever described or analyzed in textbooks and statistics curricula for non-specialists. (That said, see Haack’s 1979 textbook for non-statisticians for an early, yet quite comprehensive, treatment of indicators.)

3.5. Statistical Techniques and Big Ideas

There is a vast range of techniques used by official statistics providers. The basics of descriptive statistics and statistical inference may be encountered by the subgroup of those who learn statistics at an introductory level at the high school or college level.

In this section, however, we refer to an array of additional techniques and ideas that are frequently used in official statistics, such as moving averages, seasonal adjustment, data smoothing, case weighting, and the like. Specific areas of official statistics may have additional important approaches, such as the use of models and assumptions for population projections, or national accounts and purchasing power parities in economic statistics (Pfeffermann 2015).

Understanding of these and related techniques may not be essential for the understanding of statistics reported in the media, as technical terminology related to the methods listed above is quite often not used in the regular media, except in the business section of
newspapers. However, knowing about their existence, even if they are treated as a ‘black box’ and their actual computational nature is not learned, may be important if an adult wants to adopt a questioning stance or desires to understand more deeply how certain conclusions are derived, or how credible the underlying data are. For instance, how is it possible to conduct comparisons across different economic, financial and social systems that have monetary systems with different characteristics, or if social or economic conditions (e.g., inflation) have changed the base against which comparisons are being made?

Furthermore, critical interpretation of the statistical findings released by official statistics providers also requires an understanding of notions pertaining to confounding variables or conditioning of probabilities (Schield 2011) and related statistical ideas and techniques that are usually not afforded much attention in introductory-level classes.

3.6. Research Methods and Data Sources

Knowledge bases related to methodological issues are often spread between the discipline of statistics and the domain loosely called ‘research methods’ (Murtonen 2015). There is an overlap between them (Gal 2007; Meng 2009), and 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 (Gal and Ograjenšek 2010; Ograjenšek and Gal 2016).

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 and Lucas 2006). Regardless of the existing diversity, however, the logic of the statistical enquiry process (Wild and Pfannkuch 1999) or the PPDAC (problem, plan, data, analysis, conclusion) cycle (MacKay and Oldford 2000) is likely to be encountered.

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. Official statistics providers, however, 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, nonresponse, or respondent bias that determine data quality or credibility receive much attention in official statistics.

3.7. Awareness and Skills for Citizen Access to Statistical Reports

As already explained, citizens need to know that much of the statistical information or statistics-based messages that appear in the media, in fact, originate in a release or report
prepared by a statistics provider (Gal 2003b). This is a source differentiated from reports generated by journalists based on ‘open data’ sources (Coddington 2015), even though such open data sources themselves may, in fact, have been created by an official statistics provider.

Hence, as UNECE (2012) also recognizes, engaged citizens need to be aware of the fact that they can access the website of an official statistics provider and often get free and easy access to the same data products or publications used by the media (i.e., a press release or a technical report). This means that adults can verify or cross-check claims they have encountered in the media, and learn about a certain topic beyond the selective information in a media article.

However, the website of a typical official statistics provider presents a complex environment, certainly to newcomers and often even to more experienced users (Gal 2003a, 2005; Bregar et al. 2006; Ridgway 2015). Citizens have to search for information without necessarily knowing how to search for it, or how to use glossaries or help systems that are often written for professionals and not for the general public (Gal and Bosely 2005). Further, they need to be aware of the fact that on the provider’s website, they may find prior versions of the same information products (e.g., press releases and reports from the same survey which was conducted earlier). In addition, official statistics providers also publish technical information, ‘metadata’, about how the data were gathered or a survey was implemented, how variables were defined and measurements performed, including access to the actual phrasing of survey questions. Finally, some official statistics providers enable citizens to use data visualizations in order to view certain data from multiple viewpoints, and in some cases even provide online analytic tools that enable citizens to conduct their own analysis on aggregated data.

The upshot is that the scope of the information presented on an official statistics provider’s website about a topic is much broader and deeper compared with the simplified information or data that students encounter in a statistics class, and may require more sophistication and mental flexibility on the part of the users.

4. Discussion: Achieving Official Statistics Literacy

4.1. Critical Examination of Past and Present Efforts to Promote Official Statistics Literacy

To date, discussions of the connections between official statistics providers and statistics educators 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. Some official statistics providers have also developed specialized sections on their websites that are geared towards teachers and students, or support the international CensusatSchool project and its various derivatives (Davies 2011). The richness and importance of such and related activities have been noted and appreciated around the world (see e.g., Sanchez 2008; Townsend 2011; Helenius and Mikkeliä 2011; or MacCuire 2015).
As valuable as these efforts to increase general statistical literacy are, it needs to be pointed out that they did little to systematically promote understanding of issues pertaining specifically to official statistics.

In this article we sketched a new model of six interconnected knowledge elements of the world of official statistics, about which non-specialists and adults at large should possess knowledge to be considered literate in official statistics. In the prior sections, we analyzed how such knowledge elements go above and beyond what is usually associated with learning introductory statistics, or how statistical literacy related to official statistics is understood by bodies such as the UNECE (2012). All our findings imply that unique efforts are needed to promote official statistics literacy.

We believe that improvements that may affect knowledge levels of current (primary and secondary) school pupils or tertiary students, as valuable as they are, do not directly impact the skill set of the current adult population, which is outside of formal education systems’ range, yet comprises the main audience that statistics providers try to reach. Given the relatively slow rate at which the adult population is replaced by younger cohorts, even if knowledge among school and university graduates about official statistics vastly improved overnight, it would still take two to three decades for new knowledge gained at school level to be shared among (the younger) half of the adult population. Consequently, many adults will still lack such knowledge for decades to come.

For these reasons, it is important to continue existing specially targeted collaborations between official statistics providers and school-level educators, as noted by sources discussing the development of statistical literacy at school level (e.g., Gal 2002; Sanchez 2008; Watson 2013). Townsend (2011), Helenius and Mikkilä (2011), UNECE (2012), MacCuirc (2015), de Smedt (2016), and others, describe numerous relevant initiatives and services aiming to promote official statistics literacy that have been implemented over the years by statistics providers at national and sometimes international level.

Examples include:

- the provision of workshops, brief online courses and supportive training materials about official statistics designed for specific non-specialist user groups with known characteristics such as journalists, business leaders, or government workers,
- the provision of short leaflets about key indicators that affect the general public, such as the consumer price index,
- the preparation and posting of answers to frequently asked questions about finding or interpreting selected key official statistics on the provider’s website,
- the provision of simplified explanations about official statistics in selected key areas (e.g., Eurostat’s Statistics Explained mini-website on migration statistics),
- the preparation and posting of answers and non-technical explanations about selected basic statistical terms, statistical glossaries, and more.

Such initiatives and activities are essential and have the potential to contribute to the mission of official statistics providers and to the ability of users to comprehend specific information products in several important ways. Yet, we believe the vision of systematically promoting official statistics literacy within the general adult population (including actions in countries with characteristics that differ from the few that have
spearheaded educational services and activities) requires an examination of additional directions –from a long-range future collaborative perspective.

4.2. Proposed Directions for Future Collaborative Actions to Promote Official Statistics Literacy

Taken together, the six elements of official statistics proposed in this article and depicted in Figure 1 imply that if citizens aim to understand official statistics about society (i.e., civic statistics) to which they are exposed through the media, or if citizens attempt to find, read, and critically comprehend actual products (e.g., press releases, highlights, annotated visualizations) on the website of an official statistics provider, they need a knowledge base that is above and beyond what is taught in regular introductory statistics classes for non-specialists.

Figure 2 presents an illustrative example for how several of the six elements or knowledge areas in our proposed model co-exist in a seemingly simple product from an official statistics provider. The text in Figure 2 is an excerpt from a one-page regular press release by a national statistics provider (Statistics Portugal) that the public may hear about via a news website or a newspaper item. The example is taken (with permission) from Gal et al. (2016) who developed it for a workshop on understanding ‘civic statistics’ that is part of ongoing work by the aforementioned ProCivicStat project.

Despite its brevity, this excerpt can be used to show how multiple areas in our proposed model are all called upon to comprehend the given text. The text refers to:

- the production of statistics as part of a system of official statistics that relies on general international standards, and generates modifiable or provisional data (area 1),
- the nature of statistics about society, that is, use of rich text to convey a statistical finding, or the dynamic and aggregated nature of statistics (area 2),

Press release, Statistics Portugal

At risk of poverty rate, in 2014–15

The 2015 EU Statistics on Income and Living Conditions survey provisional data on previous year incomes indicates that 19.5% of people were at risk of poverty in 2014, keeping the value of the previous year.

The risk of poverty for the elderly population has increased for the second consecutive year.

The presence of children in a household is associated to a higher risk of poverty, reaching 22.2% for households with dependent children vis-à-vis 16.7% for households without dependent children.

Fig. 2. News about poverty – press release from an official statistics provider.
the use of an indicator, that is, risk of poverty (area 3),
big ideas in statistics, for example, risk (area 4),
the use of specific research methods (area 5).

As Gal et al. (2016) explain, the example in Figure 2 also illustrates the need for adults to
be able to critically reflect on the origin and quality of data, and on how variables or social
phenomena are defined and measured.

With the above in mind, we outline two possible initiatives at the international level,
and some additional ideas that specific official statistics providers can implement at a
local level.

Firstly, 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. 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.

Secondly, 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. 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 already mentioned
Phare project, which resulted in the modular online Course on the European Economic
Statistics (Bregar et al. 2000).

Several organizations, of which some have been referred to earlier in this article, appear
to have both the infrastructure, resources, and interest necessary to promote both a
textbook and a MOOC as outlined above. These include, among others, the EMOS
community, which presently includes over 20 universities and cooperating national
statistics offices, with support from Eurostat or UNECE, as well as SIAM and networks of
official statistics providers in Asia and Oceania. Furthermore, PARIS21 and the UNESCO
Institute of Statistics, and other organizations involved in statistical capacity-building in
developing countries are well positioned to further clarify the knowledge needs of non-
specialists who engage official statistics in such countries.

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.

The need for a comprehensive knowledge base related to official statistics may raise questions about the relative importance of the six knowledge elements outlined in our proposed model, as well as about preferred learning trajectories. Such questions may seem useful, given the need to prioritize development efforts when writing a new textbook or developing a new MOOC. We believe that all six areas are important in the long term, and there is no known consensus yet as to what may be considered ‘basic’ or more ‘advanced’ levels of knowledge in this regard, or a best learning sequence. One needs to take into account possible learning trajectories for learners with different starting points (e.g., in terms of basic knowledge in statistics or other parameters), and the need to motivate learners diverse in their background, learning styles, and so on, along the way. Arguably, at an initial stage of development, it may be advisable to select a few high-visibility, some simple and some more advanced indicators or findings of interest to the general public, and discuss some basic methodologies and working principles related to them. At a later stage, it is possible to expand the coverage of these and all other areas in our model.

The above preliminary ideas notwithstanding, we believe that the design of a textbook and a MOOC can, and should, benefit from current technological flexibilities, and be conceived from the outset as an integrated collection of digital learning resources that will be developed in parallel by multiple partners. This may reduce the need for topic prioritization. Many potential development partners that were mentioned above can build on the already existing partial resources (shareable materials from existing diploma and degree programs aimed at non-specialists) as well as ongoing work by other stakeholders (e.g., the already mentioned ProCivicStat, or individual instructors around the world) who can be called upon to share their teaching materials.

The envisioned collaborative digital resource enables the development of multiple variants of textbook chapters or MOOC units, distributed across multiple partners who work in parallel; with common as well as nation-specific modules. Subsequent review and revision processes can also move in parallel, with new resources added and hyperlinked in iterative stages. Such an approach can help to shorten the development timeline to a degree that enables the coverage of all six areas in our proposed model, initially in English, given its position in the international statistical system, with translation to other languages and localized adaptations moving ahead as materials in English become available.

Finally, apart from the two initiatives envisioned above, at the local level official statistics providers can take additional steps in order to help educate providers who work with adult learners and college or school-level populations. Educators can be equipped with collections of examples of how the media reports about press releases or other official publications, since virtually all official statistics providers nowadays use clipping services or media analysis companies that monitor all media channels. Hence, official statistics providers could develop focused packages, organized around specific issues of social significance, including the original press release and several real-life examples of how data and findings were reported in diverse media channels, selected to illustrate proper, as well as distorted, or one-sided use of statistics. Such packages could be accompanied by suggestions for in-class discussion and take-home assignments.
In summary, it is important to state that the conceptualization of the building blocks of official statistics literacy presented in this article is preliminary and open to debate, since we live in a dynamic world. Discussions emerge in professional channels on the era of ‘open data’ and its implications for both producers and users of statistics, and on the need for public understanding of statistics (von Roten 2006; von Roten and de Roten 2013). Ridgway (2016) points out that significant developments such as open data, big data, data visualisation and the rise of data-driven journalism, provoke new sorts of questions, make possible new sorts of answers and are changing the nature of available evidence, the ways in which it is presented and used to influence policy, public opinion and business practices, and the skills needed to interpret it.

The six elements we identified combine both abstract ideas and a general understanding of a complex working system in its social ecology, as well as knowledge bases of a more technical nature. Details of these elements and their operationalization have to be further examined and developed in more detail, both because official statistics itself is practiced in somewhat different ways in different contexts or by official statistics providers with different missions, and because it is evolving over time, as outlined above. We hope that the ideas proposed in this article will initiate a productive dialogue and ultimately lead to further pragmatic development-friendly decisions among statistics providers and other stakeholders interested in active promotion of official statistics literacy.

5. References


Gal and Ograjenšek: Official Statistics and Statistics Education


Received October 2016
Accepted January 2017