EXPANDING CONCEPTIONS OF STATISTICAL LITERACY: AN ANALYSIS OF PRODUCTS FROM STATISTICS AGENCIES

IDDO GAL
University of Haifa, Israel
iddo@research.haifa.ac.il

SUMMARY

This paper reports the results of an exploratory study of the characteristics of key information products released by statistics agencies. Such products are central to debates and decisions in the public arena, but have received little attention in the literature on statistical literacy, statistics education, or adult numeracy. Based on a qualitative analysis of Internet-based products of six national and international statistics agencies, the paper sketches the characteristics of five product types (Indicators, Press releases, Executive summaries, Reports, and Aggregate data) and of the environment in which they are found. The paper discusses implications for the specification of the skills needed for accessing, filtering, comprehending, and critically evaluating information in these products. Directions for future research and educational practice are outlined.

Keywords: Statistics education research; Statistical literacy; Official statistics; Educational technology; Adult numeracy; Mathematics education

1. INTRODUCTION

Increasing attention has been given over the last decade by the statistics and mathematics education communities to the development of statistical literacy and numeracy skills of all citizens (Gal, 2000). This trend has most recently been exemplified in the overall theme of ICOTS-6, the 6th International Congress on Teaching Statistics (Cape Town, South Africa), "Developing a Statistically Literate Society". Multiple paper sessions and presentations in this conference directly or indirectly addressed issues related to this theme (Phillips, 2002).

The term "statistical literacy" has not yet gained an agreed-upon meaning among educators and professionals, and some use it without an explicit definition (Cerrito, 1999). The view of statistical literacy that guides this paper stems from the assumption that most adults will be consumers, rather than producers, of statistical information. Diverse but related conceptions of statistical literacy have followed from this assumption. Wallman (1993) argued that statistical literacy is the ability to understand and critically evaluate statistical results that permeate daily life, coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions. According to Gal (2002a), 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. Lehohla (2002) views statistical literacy as the ability to read and understand quantitative information such as indices and indicators.

While statistical literacy appears as a key competency for many adults and for various professionals and officials, discussions of the kinds of information to which students and

adults have to be able to apply this competency are lacking. Many educators focus on developing students' ability to be aware of or not be fooled by "misleading" statistics or "biased" reports in the *media* (Crossen, 1994; Cerrito, 1999; Moreno, 2002). Very few sources focus on the ability to read and critically interpret information from other sources (Frankenstein, 1990; Gelman, Nolan, Men, Warmerdam, & Bautista, 1998). This paper aims to add to an emerging dialogue on the promotion of statistical literacy, by examining the nature of the products made available to the public by *statistics agencies* and by reflecting on the skills demands of these products.

1.1. CLIENTS AND PRODUCTS OF STATISTICS AGENCIES

The term *"statistics agencies"* is used here to encompass three types of statistics-producing organizations operating in the public sphere and funded by governments:

1. *"National/Central agencies"*: National statistical offices and organizations responsible for conducting a nation's *census* and associated surveys and reporting their results (e.g., the United States Bureau of the Census), or for producing all national social and economic statistics, including the census (e.g., Statistics Sweden, Italian National Statistical Institute-ISTAT).

2. *"National thematic agencies":* National organizations assigned by their governments to collect and report official statistics in designated areas not covered by the agency conducting the census or by other national agencies. Examples are the U.S. National Center for Education Statistics, or the Australian Institute of Health and Welfare.

3. *"International agencies":* These are international organizations established by member nations in order to contribute to social, human, and economic progress. One of their primary missions is the reporting of comparative statistics collected by member nations, but they also initiate or conduct special comparative studies. Examples are: UNESCO Institute for Statistics, World Bank, Organization for Economic Cooperation and Development [OECD]. This category also includes ad-hoc multi-national research projects not associated with specific countries, such as the Third International Mathematics and Science Study [TIMSS].

These three types of organizations are viewed as *official* statistics agencies since they focus primarily (though not exclusively) on statistical and research work aimed at informing policy makers, and are funded by tax-based budgets. Of course, many other organizations which release statistics to the public do exist, but fall outside this definition because of their funding sources or the issue-specific nature of their statistical work. Examples are Non Governmental Organizations (NGOs) such as Amnesty International, non-profit research institutes and advocacy groups, and industrial companies or survey organizations such as Gallup.

Murray and Gal (2002) analyzed products from Statistics Canada, a central national agency responsible for the census and other official statistical work in Canada, and grouped them under five key categories: (1) *Indicators*, (2) *Press releases*, (3) *Executive summaries and Highlights*, (4) *Reports*, and (5) *Aggregate data*. (They also listed other products, such as raw data files or technical documentation that are mostly of interest to specialists rather than to the general public). Murray and Gal (2002) argued that information products from statistics agencies serve multiple clients, but are especially geared for policy-makers and politicians and designed to inform their decisions and policy setting. They suggested that such products are also of major interest to the media, whose job is to transform official statistical reports and findings into simpler messages that reach wide audiences who

otherwise may be unaware of official statistics. In addition, official statistics are of interest to NGOs, business managers, and researchers and consultants. Finally, Murray and Gal claimed that understanding of information products provided by statistics agencies is essential for citizens at large and to interest and advocacy groups who desire to participate in public debate and influence the direction of decisions on local and national issues (European Commission, 1996).

The Murray and Gal (2002) work has provided a preliminary map of key information products from official statistics agencies, but little is known about their actual features. Such products overall have received little attention from both educators and researchers, although they have an important role in the information fabric of modern societies, and are the basis for many media messages that are of interest to the general public and to various officials and interest groups (European Commission, 1996; Podehl, 2002). With this in mind, this paper reports the results of an exploratory study designed to analyze the characteristics of key products of statistics agencies and of the environment in which users have to find these products. Such an analysis is an essential step toward expanding the understanding of the statistical tasks facing citizens; it can help to clarify the skills that should be considered part of statistical literacy and to identify targets for research and education efforts.

2. METHOD

2.1. APPROACH

Given the lack of prior research regarding products of statistics agencies, the approach taken was to use a multiple-case-study method as a basis for generating qualitative descriptions of the characteristics of key products of statistics agencies. Such an exploratory approach is an accepted way to aid the formulation of research directions and questions in an uncharted area. The study focuses on analysis of information appearing in *Internet sites* of statistics agencies since they provide clients, including researchers and educators, with access to all current key products (printed as well as interactive products not available in print), and since the Internet appears to be a central distribution medium for products of statistics agencies.

2.2. SAMPLE

This study used an intentional sample of six Internet web sites of statistics agencies, two from each of the three main categories listed above. Only agencies operating sites in English were considered, due to the need to compare both national and international agencies. The sample was designed to include agencies operating in countries with different types of governments, economies, and demographics, as well as international agencies with diverse missions. The chosen agencies created products encompassing a very wide range of themes and issues of interest to citizens and policy-makers in both developing and developed or knowledge-based nations.

The resulting sample included the agencies listed in Table 1 (acronyms appearing after the full agency name are used throughout this paper).

Table 1. Agencies Included in the Sample

Agencies and Acronyms	URL
National/Central agencies:	
a. Statistics South Africa (Statistics-SA)	www.statssa.gov.za
b. UK Statistics (Statistics-UK)	www.statistics.gov.uk
Thematic agencies:	
a. United States National Center for Education Statistics (US-NCES)	www.nces.ed.gov
b. Australian Institute of Health and Welfare (Australian-IHW)	www.aihw.gov.au
International agencies:	
a. UNESCO Institute for Statistics (UNESCO-IS)	www.unesco.org
b. Organisation for Economic Cooperation and Development (OECD)	www.oecd.org

2.3. PROCEDURE

All Internet sites were accessed in July 2002 and analyzed for the presence of the five product categories described in Table 2. A site was considered as carrying a product if at least three instances of this product type were found. An overall impression of the characteristics of the product category was reached through content analysis of at least three instances/items in each product category on each site.

Table 2. Product Descriptions

Category	Description
Indicators	Summary statistics that reflect the condition of key aspects of social, human, or economic importance. Examples: Gross National Product [GNP], Per-capita income, Pupil-teacher ratio, Infant mortality.
Press releases	Short summaries written to inform reporters and media personnel about changes in key indicators, about key findings from new or recent studies, etc.
Executive summaries and highlights	Summaries of studies, describing key findings and conclusions, without technical details. Designed primarily for policy makers.
Reports	Lengthier publications than Executive Summaries. Contain longer discussions, elaborate statistical tables and displays, and information about methodological and technical aspects of studies
Aggregate data	Detailed numerical information about indicators or other key variables that users can access directly, i.e., not in a Report.

3. RESULTS AND ANALYSIS

Overall, the product types listed in Table 2 were found on all six sites, yet ranged widely in their characteristics. Below is a sketch of the environments encountered when accessing the six sites, followed by an outline of the key features of the five product types. Subheaders in each section below sketch key findings in each area. (Note: exceptions to key findings were found in some cases, as could be expected given the breadth and complexity of the

information carried by six different official agencies; to facilitate the presentation of overall trends, such exceptions are not always noted below).

3.1. OVERALL STRUCTURE AND USABILITY OF WEB SITES

This section sketches the characteristics of the Internet web sites of the six agencies analyzed. Attention was placed on how each site is organized, how descriptive and technical information is presented to the user, and how easy or difficult it is to use the interfaces or facilities provided on each website for finding needed materials, to access different products, or to perform various operations. Given that this exploratory study was designed to analyze five types of information products of statistics agencies with an eye towards educational or institutional implications, the analysis of ergonomic characteristics and of features that may affect the overall usability of the six websites was informal and qualitative in nature.

Agencies carry a large number of items on different topics

Most sites carried many hundreds and often thousands of separate information items (i.e., specific instances of a product, such as reports, executive summaries, or stand-alone tables). To illustrate the range of items carried, Statistics South Africa, one of the smaller agencies reviewed, carries over 70 press releases, over 60 reports, over 100 documents of various types defined as "publications" or "papers", and hundreds of other types of stand alone tables and time series. Australian-IHW noted that it adds around 80 new reports a year. Most other agencies carried many more items. The OECD site lists 20 different product categories and enables access to thousands of different items.

Sites differ in their structure and organization

The homepage (the first screen the user encounters, from which various items can be located) of some sites was organized according to type of product sought by the user, while in others items were organized around areas or themes. For example, Statistics-SA lists on its homepage categories such as "Reports" or "Working Papers". When one is chosen, the user can further select from a list of sub-categories or of products sorted alphabetically within sub-categories, requiring the user to skim long listings to find if a document on a specific topic exists. In contrast, upon entering the OECD site, the user encounters a list of over *30* distinct themes (e.g., ageing society, energy, health, education, corruption, taxation, etc.). Upon choosing any of them, a list of sub-themes is displayed and associated documents shown for each. As a result, users have to be able to use different search logics when searching for information in different sites.

Help and search options are cumbersome

Most sites offer help screens with tips on how to use the site, or a site map with an overview of the parts of the site, but the sheer volume of options in all sites requires that users adopt various strategies to locate needed information. To find a document on a specific topic, users have to scan multiple and sometimes long listings, or use the site's search engine with the hope of reducing the labor involved in pinpointing potentially useful documents through list scanning. Yet, search engines follow a different logic in each site. Users have to generate alternate search terms or keywords when searches fail to provide adequate hits, or, when searches yield many hits, navigate through the resulting lists or try other keywords. Hence, users have to be able to combine different search facilities: browse product lists on

screen, use a search function which is offered by the site, or use the search function of the browser with which they view the site.

Glossaries are formal and demanding

A search for a glossary, i.e., dictionary explaining the meaning of terms or computation of various variables, is a demanding process in most sites. A clearly marked and comprehensive central glossary could not be easily found on any of the sites reviewed. Requesting a search engine on any site to locate a "glossary" usually yielded multiple hits referring to glossaries developed independently by specific authors within some documents, forcing the user to open these documents in search of the desired definition. Such embedded glossaries are not available to users who do not reach into their host reports. Explanations in glossaries, when they could be found, were often formal and verbose, and some assumed prior technical or statistical knowledge (as illustrated by the example in section 3.2 below).

Documents may be created by multiple sources

Some agencies offered products created not only internally but also by outside sources, such as other statistics agencies, government departments, or collaborating institutions in member nations. For example, some of the products carried by Statistics-UK originated from collaborating agencies in Scotland and Wales, and Australian-IHW carried products created by departments in different Australian states. The existence of multiple source organizations can help explain the diversity in style and internal organization found in some of the product categories described below.

Access to information requires familiarity with various computer programs

All sites store many texts of interest in Acrobat Reader (PDF) or in Microsoft Word or Excel file formats, or in these formats inside compressed archives (ZIP files). Documents a user chooses to view are sometimes opened automatically by a plug-in module of the browser (without user intervention). However, in many cases users are asked to open or save the file, or may only be given the option to download (and save) the file, in which case they need to know how to locate the saved file on their computer and open it on their own in separate steps.

3.2. INDICATORS

Indicators are statistics that reflect the condition of key variables and phenomena in areas such as economic activity, health, education, or the environment. Some indicators are generated and reported on a recurring basis, such as every week or month (e.g., *Consumer Price Index, Gross Domestic Product*), and are of interest to local clients who need to know of changes in these measures as a basis for their ongoing activity. Others are derived on a yearly or ad-hoc basis (e.g., *Infant Mortality, Number of teachers, Transport-related Air Emissions*). All six sites presented information about many types of indicators, including historical and time series information. The following three interrelated issues emerged from the analysis:

Indicators are reported with varying levels of explanatory support

Agencies enable direct access to recurring indicators, though with different levels of ease. For example, Statistics-SA has a "Key indicators" option on the opening homepage, while users of Statistics-UK have to follow a longer route, by choosing "latest figures" but then

browse a long list of releases to find those carrying indicators information. In both agencies, accompanying and sometimes lengthy notes explain changes in indicator values from prior months and discuss points that should be considered when interpreting the figures. Yet, the nature of the indicators themselves is not explained, probably on assumption they are standard entities with which all key clients are already familiar. In addition, in all six sites multiple lengthy reports are available that discuss the status of different indicators (e.g., changes in the status of the health system as reflected by indicators such as mortality, staffing levels, or expenditures).

Indicators vary in concreteness and transparency

Some indicators carry straightforward names (*Number of hospital beds*, *Adult literacy*) that make intuitive sense even for people without advanced levels of education or who do not know how the indicators are actually measured or calculated. Other indicators appear more abstract (*Rate of primary school enrollment, Life expectancy*) but still use common terms. Yet, some indicators carry complex names that make no intuitive sense and hence cannot be understood by users lacking specialized training. For example, *Apparent Intake Rates* (UNESCO-IS), or *Purchase Power Parities* (OECD). Possibly their meaning could be understood by consulting additional explanatory material, but as noted earlier most sites do not offer easy access to glossaries.

Understanding indicators requires familiarity with mathematical and statistical terms

A vague sense for the issue(s) addressed by an indicator may sometimes be sufficient to allow a user to understand the general meaning of indicator values. However, the names of some indicators require familiarity with mathematical or statistical terms. For example, knowing what "ratio" or "average" mean is a prerequisite for understanding *Pupil-teacher ratio* (an education indicator) or *Average length of stay* (a healthcare indicator). Full understanding of an indicator's meaning, however, requires going beyond its name and coming to grips with how it is measured or calculated (especially if the indicator is a composite of other variables, as with *Gross national product*). For example, on UNESCO-IS, when users request additional information regarding the meaning of *School-life expectancy*, one of the 16 key indicators reported, the following glossary entry appears:

> *"The total number of years of schooling which a child of a certain age can expect to receive in the future, assuming that the probability of his being enrolled in school at any particular age is equal to the current enrolment ratio for that age".*

And the calculation method is defined as follows:

> *"For a child of a certain age, school life expectancy is calculated as the sum of the age specific enrolment ratios for the reference age-range a to n, divided by 100."*

Complex glossary descriptions such as the above, which were also found on other sites, may reflect an attempt to be technically accurate and inform users of "best practices" of official statisticians. The price is that their understanding requires users to be familiar with school-based terms such as "total", "probability" or cope with more complex mathematical and statistical terms. In addition, users have to grapple with the linguistic complexity created by composite phrases (e.g., "*age specific enrolment ratios*", "*reference age-rang*e").

3.3. PRESS RELEASES

The term "press release" refers to short summaries issued by statistics agencies to inform media personnel about changes in key indicators, about key findings from recent studies, or

about new statistical data or results that may be of interest to the general public. As noted earlier, press releases are expected to be used by journalists or reporters as the basis for messages suitable for their own particular audiences.

Releases have a relatively uniform structure within agencies, less so across agencies

Press releases were found on all sites, and had a rather uniform style and format within most agencies, but a more pronounced variability in format and style across agencies. For example, press releases from OECD and US-NCES were normally 1-2 pages in length, and comprised of separate, relatively short paragraphs written in a journalistic style, i.e., in the way reporters may wish to write results and conclusions so as to pique reader's interest. Releases by Statistics-SA employed a journalistic style in discussing main points, but tended to make fewer interpretive statements and presented more raw statistics (absolute numbers and percentages) than OECD or US-NCES.

In contrast, "press releases" from Statistics-UK were very diverse in content and format. Some were only a short notice about the release of a new report, without any summary of results. Some also contained a link through which the user could access a table of contents of a large report and there review a "summary" or "highlights" section. Others present data tables (e.g., total numbers and causes of death last month in the UK), followed by brief background comments about the data collection process, but without a discussion of the implications of the data. Only some Statistics-UK releases were similar to those from OECD or US-NCES, but with a different internal structure and a more formal, technical style; these start with a single opening page containing a lead paragraph, then a few separate sentences presenting key points, and 1-2 simple graphs or a table. Additional pages either had text paragraphs discussing other findings, or multiple tables with supporting data but with little interpretive text.

Press releases vary in the amount of interpretation they offer the user

Press releases vary somewhat within agencies but more so across agencies in the extent to which they list raw statistical findings and "let the numbers speak for themselves", or on the other hand surround the findings with interpretations and commentary about their meaning, implications, or limitations. The diversity found in the amount of interpretive text may stem from different factors, such as differences across countries or agencies in expectations regarding the media's role in society, in assumptions regarding the ability of journalists to understand statistical information, or in the philosophy of agency administrators regarding what is proper to include in a press release.

When raw data or findings are presented to users without much interpretation or explanatory text, it may be argued that users benefit because they gain access to "objective" data or to "untainted" reporting. On the other hand, the lack of interpretations and explanations about the import of the findings or the overall meaning of separate findings puts more burden on the users. In such cases they need to interpret on their own factual statements and read messages carefully to notice patterns or links between findings presented separately. Thus, the amount of interpretation or commentary that agencies offer in press releases has implications for the type and level of background knowledge and skills that users need to possess.

Press releases use a wide range of statistical and mathematical terms

Along with diversity in the amount of interpretive commentary, press releases also varied in the range of statistical and mathematical terms and ideas they contained. A partial quote

from the opening paragraphs of an US-NCES release about the National Assessment of Educational Progress' Geography 2001 survey illustrates the use of a journalistic style but also the use of multiple statistical and mathematical concepts and ideas:

A new report released today by ...National Center for Education Statistics shows that average scores of the nation's fourth and eighth graders, while low, have improved from 1994. Lower-performing students at grades four and eight showed an increase in average scale scores, whereas no overall changes were seen for 12th graders… the improvements for fourth-and eighth-graders were seen among students scoring in the tenth and 25th percentiles of performance.

Clients seeking concise, non-technical information in press releases have to be prepared for the range of approaches to writing of press releases in different agencies. Clients may need to be able to read, interpret, and critically evaluate not only highly abstracted text, but also understand non-technical but also more advanced statistical or mathematical terms, and be able to elicit additional information from tables or graphical displays provided as part of the press release.

3.4. EXECUTIVE SUMMARIES AND HIGHLIGHTS

According to Murray and Gal (2002), policy makers and officials are the primary clients of statistics agencies. Given that these individuals do not have the time or need to deal with many details, they seek concise documents that provide an overall picture of key findings, or a summary of the main conclusions from a given study or activity of a statistics agency. Short summary publications that can satisfy the information needs of such key users were thus expected to exist as a distinct product category in all agencies examined in this study.

Executive summaries vary in length and location

The analysis provided mixed support for the presence of short publications aimed at policy makers or officials as a separate category from longer publications. In all sites it was possible to identify some brief stand-alone summary documents, often called Executive summary or Highlights, usually as brief as two pages but sometimes ten pages or more in length. Executive summaries or Highlights were more commonly found, however, as an integral opening section in a long Report.

Searching for executive summaries is complex and affected by site organization

As noted in 3.1, some agencies organize their homepages in terms of key products, others by topics. Users who are interested in reading a summary on a topic, not a full report, face a challenge: it is not obvious how to find an executive summary on a given topic where agencies place summaries inside reports, not as stand alone documents. Search engines do not enable identification of documents that contain an executive summary or highlights section, leaving users no choice but to download or open lengthy reports in order to locate executive summaries on topics of interest. An interesting example for the flexibility required of users was found in Australian-IHW, which presents to users two separate lists titled "Welfare" and "Health", each of which contains dozens of items. Almost none of the documents listed under "Welfare" were stand-alone Executive Summary or Highlights (though some reports opened with an Executive summary), whereas many documents listed under "Health" were marked as "Highlights". Thus, products in the two key theme areas of this agency were being shaped by different principles, requiring different reading and search strategies in the same site.

Executive summaries can demand substantial statistical and mathematical knowledge

As indicated above, Executive summaries varied in length. The longer ones included extended text passages, as well as graphs, charts, and tables to organize key results or show important trends discussed in detail in an accompanying Report. The range of statistical terms and concepts addressed in these summaries is very wide and covers the full spectrum of topics included in introductory as well as more advanced statistics textbooks or courses. Table 3 includes excerpts from executive summaries from three of the six agencies, chosen to illustrate the range of statistical concepts and ideas that users can expect to encounter in texts.

Table 3: Excerpts from Executive Summaries from three Agencies

#1: On a worldwide scale, the total number of primary school pupils of any age rose from just under 600 million to over 680 million between 1990 and 1998. The percentage of children of primary age in school edged upwards, from 80 per cent to 84 per cent. Since the beginning of the decade, primary enrolments have increased by an average of 10 million each year, almost twice that recorded in the 1980s.

#2: Latin America and the Caribbean and East Asia and the Pacific appear to be the only less developed regions with the capacities to provide education for all primary school-age children. Although their net enrolment ratios are still below 100 per cent, their gross enrolment ratios were 100 per cent and above throughout the decade.

#3: The disparities within regions are immense, however, with a difference of over 70 percentage points between the highest and lowest enrolment rates... The greatest variations can be seen in sub-Saharan Africa with gross enrolment ratios lower than 1 per cent in Congo to 111 per cent in the Seychelles.

#4: The number of pupils per teacher varies enormously on a global scale, from a low of 9:1 to a high of 72:1. In 1998, 75 per cent of the countries reporting had pupil/teacher ratios below 37:1...The highest ratios are found in Central and Western Africa, where the average (median) pupil/teacher ratio rose from an already high 50:1 in 1990 to 52:1 in 1998.

#5: For explanatory aspects of poverty, for example educational attainment and access to services, the IES data were merged with data from the 1995... household survey (OHS)... A series of regression analyses was carried out, using annual household expenditure as the dependent variable, and the poverty-related variables common to the OHS and the census as the explanatory variables, to impute expenditure values for each household....Two...indices – the household infrastructure index and the household circumstances index – were constructed to measure the extent of under-development in different parts of [the country], using both the data from Census '96, and the imputed expenditure values described above.

#6: (Table subtitle) Relationship between average performance across combined reading, mathematical and scientific literacy scales and cumulative expenditure on educational institutions up to age 15.

#7: (Explanation on how to read a complex table with over 30 country names on both the X and Y axis): The figure below summarises the performance of countries on the reading literacy scale. It also indicates whether countries perform significantly higher or lower than the comparison countries as well as the estimated rank order position of each country. For example, Finland, with all triangles pointed up performed significantly better than all other countries while Canada performed significantly lower than Finland shown by a triangle pointed down, similarly to New Zealand, Australia, Ireland and Japan shown by a circle and significantly better than all other countries shown by a triangle pointed up.

#8: Because data are based on samples, it is not possible to report exact rank order positions for countries. However, it is possible to report the range of rank order positions within which the country mean lies with 95 per cent likelihood.

#9: The fact that [the survey] measures proficiency not at the aggregate country level, but at the level of individual students, makes it possible to also examine variation in student performance within countries. Such variation may result from the socio-economic backgrounds of students and schools, from the human and financial resources available to schools, from curricular differences, from selection policies and practices and from the way in which teaching is organised and delivered.

The excerpts in Table 3 show that Executive summaries make reference to all key types of standard summary descriptive and inferential statistics, such as totals (#1), percents (#1-#3), percents below 1% or larger than 100% (#3), range or high-low differences (#3), ratios (#4), rank ordering (#8), average (mean and median; #1, #4, #6, #8), correlation and regression (#5, #6), or confidence intervals (#8). In addition, phrases also relate, directly or indirectly, to core "big ideas" in statistics, such as variation and causes of variation (#5, #9), sampling (#5, #8), data reduction and aggregation (#5, #6), significance of difference (#7), and prediction (#5, #6). The many types of tables, charts and graphs that appear in summary documents cannot be illustrated here due to space limitations, but excerpt #7 illustrates that such displays may take non-standard or complex forms that force their creators to add rather dense explanations to make sure users can read them.

Often, several types of statistical ideas are combined in single text passages in ways that increase their complexity. For example, the single sentence in #4: "The highest ratios are found in Central and Western Africa, where the average (median) pupil/teacher ratio rose from an already high 50:1 in 1990 to 52:1 in 1998" requires that the reader can make sense of composite phrases that are seldom found in standard textbooks, such as "highest ratio" or "Average (median)". Texts may also refer to data management or data transformations, such as in the case of discussion of merging of datasets or imputation (#5). In addition, indices and composite variables of various degrees of complexity (#4, #5, #6) may be central to the text, requiring that users can make sense of second-order variables, i.e., that represent relationships derived from other relationships and ratios.

3.5. REPORTS

In addition to brief executive summaries described earlier, statistical agencies carry an assortment of long documents termed "reports". These publications include a detailed narrative text that discusses the background, methodology, findings, conclusions, and implications of a specific study or project. Some reports summarize an ongoing statistical activity, such as an analysis of trends or changes in certain indicators. In addition, reports also present numerous tables with detailed statistical information and graphical displays, and provide technical details that enable users to understand the limitations of the work or findings. Murray and Gal (2002) claimed that such publications are not expected to be consumed by decision makers and officials, given that these clients seek encapsulated information. Rather, reports are intended to inform other user groups, such as the support staffs who advise policy makers, or analysts and researchers who need detailed information.

Naming conventions and content vary

Multiple labels are used to describe reports across agencies. For example, Statistics-SA lists on its homepage, in addition to "Reports", "Publications", and "Working papers", users also subcategories such as "Discussion papers", "Occasional papers", and "Working papers", all of which may qualify as a report in some way. OECD lists over 20 documents types, such as "Case studies", "Country surveys", "Reviews", "Reports", "Policy Briefs", "Working papers", and others. A similar if somewhat narrower assortment can be found in other sites. Yet, no general explanation is provided on any of the sites as to the distinguishing features of overlapping entities. Document lists that appear on screen often present only document names, without additional explanations about the nature of the document. Users thus need to open documents and browse their introductory sections to figure out their actual content. With long reports, users may also have to be able to activate the "search" function of the program (e.g., Word, Acrobat Reader) they used to view the particular document in order to locate text of interest.

Reports summarizing statistical studies vary in scope, depth and length

Some reports reached upwards of 200 or 300 pages, and on occasion could not be viewed on-screen or downloaded and could only be purchased in a hardcopy or book format. A unique type of a report that appeared in all sites was an "annual statistical yearbook", a compendium of large amounts of statistical information covering the full scope of issues or areas under the jurisdiction of the agency. The information in such yearbooks (sometime carrying a different name, such as "The condition of education") often appeared with little interpretation of the data, though with some cautionary comments about data quality (see excerpt #8 in Table 3), and sometimes with a glossary or other texts explaining the meaning of some of the entities discussed in tables or graphs.

Report demand diverse and substantial statistical and mathematical knowledge

Since reports are normally longer and more comprehensive than executive summaries that summarize them (see 3.5 above and the excerpts in Table 3), reports make reference to a very wide range of ideas and concepts from descriptive and inferential statistics, combined with descriptions of data management techniques and computational procedures, such as those related to the creation of indices. However, reports do not necessarily explain the meaning or derivation of indicators and variables discussed or of various technical issues. For example, basic concepts in official statistics, such as *incidence* and *prevalence*, which are at the heart of understanding phenomena and trends related to health and well-being, were often used freely in reports. Likewise, technical comments in some reports refer to issues that are not addressed in introductory statistics courses, such as weighting of cases, dealing with missing data, standardization of values, and complex sampling schemes. Further, reports extensively use tables of varying degrees of density, including multi-page tables or tables with "nested" subdivisions, i.e., representing multiple variables by subgroupings on both the horizontal and vertical axes; see Mosenthal & Kirsch (1998).

3.6. AGGREGATE DATA

Murray and Gal (2002) argue that clients such as managers in the public and private sectors, or members of advocacy groups, need to have information on "local" issues they care about, and hence see data about subgroups, not the total population. All agencies indeed enable users to retrieve numerical information organized according to certain breakdowns, but data could be accessed in varied ways.

Users can view static tables that show statistical data for fixed subgroupings

In some agencies, users can choose tables from a long list of available titles; the row-and column summary data (e.g., counts, percents) that appear on screen seem to be a page from a published report or a yearbook. In other cases users do not see a long list of titles of tables, but rather encounter an interactive interface through which they define the table to be retrieved. For example, UNESCO-IS users can see data about specific geographical areas by picking an indicator (e.g., *Adult literacy*) and then choosing from lists of "countries" (or "regions") and "years". This form of retrieval just simplifies the interface – the tables that appear are pre-fabricated, just like above. In some cases users choose the data source, not only the type of aggregation. Users of Statistics-UK, for example, can access over 4000 different datasets pooled from various statistical agencies in the British Commonwealth via a facility called StatBase. Individual tables retrieved in this way allow users to reach information that otherwise they will be challenged to find in printed reports, though most data tables again seem to be pre-fabricated, i.e., as if they are tables from a former report.

Some agencies enable users to interactively construct new tables and displays

In addition to the ability to retrieve pre-fabricated tables with aggregate statistics for subgroups, several agencies allow users to customize displays to their specific needs, and create tables at different levels of sophistication. US-NCES users, for example, can build tables with school-related statistics by specifying table *rows* (e.g., the level of aggregation: state, county, district, school), *data category* (i.e., variables to be analyzed in the table cells, such as "enrollment by ethnicity"), and *column* (e.g., grade levels). Australian-IHW enables users to use a product called Data Cubes to compare multiple subgroups both through tables, graphs and charts; resulting displays can span a range of levels of complexity that sometimes exceed the capabilities of basic statistical packages. The implication is that agencies differ in how much flexibility they allow users, and increasing levels of interactivity seem to require more statistical and technical sophistication from users.

Agencies allow users to download data files for independent analysis

All agencies enable users to download complete datasets of aggregate statistics or time series that can be analyzed either with a spreadsheet program or various statistical packages. Given that these sub-products are meant to enable users to actively perform a statistical analysis of datasets, and that some of them can be accessed only for a fee, a review of these products goes beyond the scope of the present paper.

4. DISCUSSION

Statistics agencies are a key source for official statistical information about a broad spectrum of issues, and their products are valuable to politicians and decision makers, managers in the public and private sectors, advocacy groups, and citizens in general. Such products, however, have received little attention from both educators and researchers, although they have an important role in the information fabric of modern societies (European Commission, 1996), and enable the media to generate messages to the public (Podehl, 2002).

Given the lack of prior research in this area, this exploratory study was designed to describe the characteristics of five key products: Indicators, Press releases, Executive summaries, Reports, and Aggregate data, created by agencies of three types: National/central, Thematic, and International. All five product categories were indeed found on all six sites examined, usually with dozens and sometimes hundreds of instances in each category. This finding confirms Murray and Gal's (2002) initial observation that statistical agencies publish product variants based on the same broad analyses, and incorporate multiple rhetorical levels to make sure that publications respond to the questions and concerns user groups may have.

As illustrated in the Results section, there is some overlap between the five product categories, for example when executive summaries are part of longer reports, when aggregate tables pertain to indicator data, or when press releases discuss details of an executive summary. Nonetheless, despite the existence of hybrid cases and variability within and across agencies, the five product categories appear distinct and each has specific characteristics. Each product category hence merits separate attention from researchers and educators.

The results show that Internet sites of statistics agencies vary in their organization and ease of use, and that the products created by such agencies range widely in their characteristics and content. The results of this study should be viewed with caution, however, as the purposeful sample used here did not include non-English speaking agencies, nor did it cover all existing agencies within some countries. It was not possible, due to lack of prior knowledge of the organization of products on each site, to construct a probability sample of items (i.e., product instances) within each agency. That said, agencies addressing a wide

range of themes were selected, and multiple items from different topics in all product categories were chosen within each agency to increase representativeness. With these caveats in mind, conclusions are discussed below with regard to the skills needed for statistical literacy and to future research and educational practices.

4.1. SKILLS DEMANDS

A model proposed by Gal (2002a) assumed that statistically literate behavior requires the joint activation of five knowledge bases (literacy, statistical, mathematical, world, and critical) and of a dispositional component (involving critical stance, and beliefs and attitudes). This study highlights the importance and multi-faceted nature of the "literacy" aspect of statistical literacy, both in terms of general literacy as well as in terms of "information technology and computer literacy" (Dede, 2000). Equally importantly, the breadth and depth of issues addressed in the five product categories implies that clients of statistics agencies need to possess and activate diverse statistical and mathematical knowledge.

Literacy-related knowledge and skills

Statistical literacy was broadly defined as the ability to interpret, critically evaluate, and express one's opinions about statistical information and data-based messages. This study shows that statistical literacy can also involve the ability to *access, define, locate, extract, and filter* needed information in a complex array of information products. Clients must be familiar with the expected content of typical products, and recognize the differences between press releases, executive summaries, full reports, or indicator and aggregate data. They have to be able to plan and combine flexible search methods and adjust search strategies and keywords to sometimes inconsistent product organization and naming conventions. When reaching potentially useful documents, users have to be aware of the mediating role and basic features of computer programs used to view or open relevant files. When viewing a document, users have to employ different reading strategies given the diversity found within and across sites in document content, style, and terminology. Users may have to fit their reading strategy to the nature of the document examined, such as by skimming large reports to identify needed segments, or by activating critical reading skills when reading interpretive statements in press releases (Thistlewaite, 1990). Users also need to possess adequate *document literacy* (Mosenthal & Kirsch, 1998) and *graph interpretation skills* (Bright & Friel, 1998) to be able to analyze complex tables, graphs and charts found in all product types.

Even if users have adequate levels of *both* general literacy and information technology skills, they could be helped by possessing complementary knowledge of the specific characteristics and organization of the products of statistics agencies and the environment within which they can be found. For example, clients need to be aware of the varied terminology used by statistics agencies to name product categories and subtypes, or specific indicators and variables and of differences between agencies in this regard, but at the same time be aware of *common* terms or concepts used across agencies. Clients also need to possess knowledge and skills that enable them to effectively navigate, search, use help systems and glossaries, and browse diverse types of lists and documents. Finally, clients need to be able to comfortably switch between interfaces and naming conventions both within and across agencies, and activate site-specific programs or tools for retrieving tables or constructing customized displays.

Statistics-related knowledge and skills

Documents such as press releases, executive summaries, and reports were found to cover a broad spectrum of statistical topics, and carry a wide range of tables, graphs and charts at different levels of complexity. Press releases presented some surprises: they could be expected to require lower level of statistical sophistication from readers, as they are supposed to consist of non-technical narratives that are within the grasp of politicians and the general public (Podehl, 2002). Yet, some press releases employ statistical concepts and quantitative statements and arguments that appear quite demanding or present interpretation challenges, even though they are written in a relatively informal language.

As the various examples in the Results section illustrated, statistical terms and findings in executive summaries and full reports are embedded in rich and sometimes complex text passages. Readers have to consider explanatory variables and contextual factors, and grasp comparisons and trends across time and subgroups, taking into account the context of the study. Such products also present users with "qualifying" information, i.e., regarding the limitations imposed by data-gathering processes on the quality of the data and conclusions, or the relative confidence that can be placed in inferences or conclusions, for example due to characteristics of the samples or instruments. This means that users have to recognize the significance of such qualifying information and use it to inform their interpretation of what is being read (Utts, 1996). Qualifying information or details about a study's context and methods, however, are not always available or easy to find, especially when viewing stand-alone tables of aggregate data or displays of indicators, which often appear without much supporting narrative text. Indeed, many publications do not fully explain the meaning or computation of standard indicators or variables, probably on the assumption that the primary target audiences of statistics agencies, policy makers and officials, are familiar with the standard tools and concepts of the trade.

Active aspects of statistical literacy

A view of adults in general as "passive" consumers of statistical messages (Gal, 2002a) may be insufficient once information products from statistics agencies are considered as a legitimate part of the information landscape that all citizens may have to deal with. This study shows that "consumption" of such products requires clients to take a more active role than traditionally implied with regard to the consumption of media-based statistical messages. Of course, users have had to be able to define their information needs and the level of detail they desire even with traditional printed products, and have had to activate critical reading skills when reading any type of printed publication. Yet, the possibilities afforded by the Internet for quick and instant interaction with multiple types of products and items present new opportunities but also demand additional, more active involvement than before from clients of statistics agencies.

After defining their information needs and retrieving and reading documents deemed relevant (e.g., a press release, an executive summary) users may realize that more detailed information is needed. This may result in cycles of search, retrieval and analysis of additional items, or a decision to generate views of data for specific subgroups of the user's choice, through interaction with web-based facilities for accessing indicators or aggregate data. These stages require that users can re-examine their information needs and adjust or extend their actions, in light of their evaluation of the quality of the information they read so far. The notion of critical evaluation, though, implies a form of action, not just passive interpretation or understanding. As Gal (2002a; 2002b) argues, for any action to occur (internally or overtly), certain dispositions need to be in place. A statistically literate or numerate person has to possess positive attitudes and beliefs, such as a belief in the legitimacy of questioning

information from official sources, before he or she will be willing to invest energy in obtaining the information needed or feel comfortable to activate critical questions to evaluate this information.

Future skill demands

Products of statistics agencies and the information space within which users have to find such products (i.e., the website with all its elements and hierarchy of webpages, screens, and items that can be viewed or used) are not static entities that will remain unchanged for years to come. It can be assumed that cumulative wisdom is developing among website designers in general and within statistics agencies as well, in light of feedback and suggestions from various users and professionals. Such knowledge, together with improvements in Internet technology, may enable statistics agencies to improve upon some of the design problems noted in this study, such as by enhancing the features of online help systems and search engines, by simplifying and expanding glossaries, or by redesigning the structure of websites and reorganizing document groupings. However, core skills demands associated with comprehension of information in products of statistics agencies cannot be expected to lessen, given the ever-present need to read and comprehend texts that combine statistical terms and tabular and graphical displays in diverse and sometimes complex documents.

While some aspects of Internet sites of statistics agencies may become simpler to use, it can be conjectured that such sites will gradually evolve into more complex entities. Agencies will have to continue and make accessible existing "legacy" products, some of which were shown in this study to be problematic, but will continue to add each year many new items in all products categories, given the accumulation of new data and findings. Also, to better serve the broadening information needs of a spectrum of users whose average educational level is rising, agencies may offer users more services or product subtypes, including the ability to interactively tailor data views or aggregations to users' specific needs. Thus, even if agencies improve retrieval processes and re-organize sites and products, future users may need to possess broader and more advanced skills in order to extract maximum value from products of statistics agencies.

Summary

The picture emerging from the above analysis is that the statistical (and mathematical) knowledge base needed to comprehend and critically evaluate various products can be quite substantial. Clients of statistics agencies will often need relatively *formal* knowledge of the meaning of concepts and underlying statistical ideas. Some of the needed knowledge falls outside the material normally included in introductory-level statistics instruction, and will not be familiar to all those students or adults whose exposure to statistical topics is based only on seeing statistics in use in the media, or on exposure to procedural or computational aspects of statistics (Cobb, 1992). It follows that in conceptualizing the skills needed for statistically literate behavior there may be a need to think of subtypes, depending on the nature of the real-world "target stimuli" or "target tasks" that adults face. The present study points to the need to consider "official-statistics literacy" as a subtype that places unique demands on strong integration of relatively formal statistical and mathematical knowledge with both general literacy and information technology skills, coupled with specific knowledge regarding the types of products, document structures, terminology, and interfaces that comprise the information space presented by statistic agencies to their clients.

5. IMPLICATIONS FOR RESEARCH AND EDUCATION

The use of a qualitative approach is well-accepted as a first step in exploratory research. The analysis of the six case studies selected in this study, while of limited generalizability, can inform further research and contribute to clarification of educational practices that can promote the statistical literacy of students and adults from all walks of life.

Product characteristics

This study sketched in broad strokes the characteristics of key categories of product from statistics agencies, but further inquiries are needed to specify and quantify the characteristics of such products in a broader range of agencies. For example, it would be useful to examine the relative frequency of specific mathematical and statistical entities (e.g., percents, average) or of different types of tables, graphs and charts that appear in specific product categories, and compare to the frequency of such entities in regular media articles that refer to statistics issues (Joram, Resnick, & Gabriele, 1995). Similarly, it would be important to identify and describe linguistic structures and writing styles that more commonly appear in different product categories, to better understand the shift in skills required to comprehend each one (e.g., press releases as opposed to more formal executive summaries). Research along these lines can clarify the differences between the separate but related target stimuli on which educational efforts should focus when attempting to develop different subtypes of statistics literacy.

Knowledge and skill levels

Little information exists about the extent to which adults in general, members of specific occupational groups, or students in academic institutions or in high-schools, possess the required skills (in literacy, information technology, statistical, and mathematical areas) described above as necessary for accessing, filtering, comprehending and critically evaluating the information in products from statistics agencies. National and international comparative studies aimed at school students (e.g., TIMSS and PISA, see www.oecd.org) and adults (e.g., IALS and ALL, see www.ets.org/all) assess only a subset of the essential skills described earlier, and thus can provide only partial information in this regard. Studies are needed that will examine students' and adults' actual ability to effectively engage the interactive environment found on Internet sites of statistics agencies. Through a combination of qualitative and quantitative techniques it should be possible to analyze interpretation and thinking processes of clients who possess different skill levels, and identify gaps in comprehension, performance, or ability to critically analyze actual texts or displays in different product categories. Such studies have to be designed with the understanding that knowledge bases underlying statistical literacy, and hence manifested performance on functional assessment tasks, can exist on multiple levels, from informal and rudimentary to more formal and advanced (Gal, 2002a; Watson, 1997; Watson & Moritz, 2000). Research on the above areas can contribute to the specification and understanding of the skills and abilities that underlie statistically literate behavior, and inform educational activities.

Educational activities

An interest in the ability of all segments of the population to act as informed and empowered citizens requires attention not only to people's ability to interpret and critically evaluate statistics in the media, but also to their ability to access, locate, filter, comprehend, and evaluate information products from statistics agencies. The skills involved in the latter, however, seem broader, more formal, and at times more advanced than those required for

interpreting and evaluating statistical messages in newspapers or advertisements. Clients of statistics agencies have to *activate in integration* a network of skills and knowledge bases, but these are often *separated* in traditional modes of instruction which are organized around distinct subject areas, such as statistics, mathematics, language arts, or computer skills. To develop "official-statistics literacy", the facet of statistical literacy that pertains to official statistical information, it will be necessary to create, in both teaching and assessment, opportunities for students to apply skills in the context of realistic, socially meaningful, and motivating tasks. This requires that educators employ authentic products from statistic agencies, and implement teaching methods that focus on the actual skill demands imposed by products of statistics agencies.

6. REFERENCES

Bright, G. W. & Friel, S. N. (1998). Graphical representations: Helping students interpret data. In S. P. Lajoie (Ed.), *Reflections on Statistics: Learning, Teaching, and Assessment in Grades K-12* (pp. 63-88). Mahwah, NJ: Lawrence Erlbaum.

Cerrito, P. B. (1999). Teaching statistical literacy. *College Teaching,* 47(1), 1-7.

Cobb, G. W. (1992). Teaching statistics. In L. A. Steen (Ed.), *Heeding the Call for Change: Suggestions for Curricular Action* (pp. 3-43). Washington, D.C.: Mathematical Association of America.

Cohn, V. & Cope, L. (2001). *News and Numbers : A Guide to Reporting Statistical Claims and Controversies in Health and Other Fields*. Iowa City: Iowa State University Press.

Crossen, C. (1994). *Tainted Truth: The Manipulation of Fact in America*. New York: Simon & Schuster.

Dede, C. (2000). Emerging influences of information technology on school curriculum. *Journal of Curriculum Studies,* 32(2), 281-303.

European Commission. (1996). *White paper on education and training: Teaching and learning-towards the learning society*. Luxembourg: Office for official publications of the European Commission.

Frankenstein, M. (1990). Incorporating race, gender, and class issues into a critical mathematical literacy curriculum. *Journal of Negro Education,* 59(3), 336-347.

Franklin, B. (Ed.) (1999). *Social Policy, the Media and Misrepresentation*. London: Routledge.

Gal, I. (2000). The numeracy challenge. In I. Gal (Ed.), *Adult Numeracy Development: Theory, Research, Practice* (pp. 1-25). Cresskill, NJ: Hampton Press.

Gal, I. (2002a). Adult statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.

Gal, I. (2002b). Dispositional aspects of coping with interpretive numeracy tasks. *Literacy and Numeracy Studies,* 12(2), 47-61.

Gelman, A., Nolan, D., Men, A., Warmerdam, S. & Bautista, M. (1998). Student projects on statistical literacy and the media. *The American Statistician,* 52(2), 160-166.

Joram, E., Resnick, L. & Gabriele, A. J. (1995). Numeracy as a cultural practice: An examination of numbers in magazines for children, teenagers, and adults. *Journal for Research in Mathematics Education*, 26(4), 346-361.

Lehohla, P. (2002). Promoting statistical literacy: A South African perspective. In B. Phillips, (Ed.), *Proceedings of the Sixth International Conferences on Teaching Statistics*. Voorburg, the Netherlands: International Statistical Institute. CD ROM.

Moreno, J. L. (2002). Toward a statistically literate citizenry: What statistics everyone should know. In B. Phillips, (Ed). *Proceedings of the Sixth International Conference on Teaching Statistics*. Voorburg, the Netherlands: International Statistical Institute. CD ROM.

Mosenthal, P. B. & Kirsch, I. S. (1998). A new measure for assessing document complexity: The PMOSE/IKIRSCH document readability formula. *Journal of Adolescent and Adult Literacy, 41(8)*, 638-657.

Murray, S. & Gal, I. (2002). Preparing for diversity in statistics literacy: Institutional and educational implications. In B. Phillips, (Ed). *Proceedings of the Sixth International Conference on Teaching Statistics*. Voorburg, the Netherlands: International Statistical Institute. CD ROM.

Phillips, B. (2002) (Ed.), *Proceedings of the Sixth International Congress on Teaching Statistics*. Voorburg, the Netherlands: International Statistical Institute. CD ROM.

Podehl, W. M. (2002). Statistical literacy and the media. In B. Phillips, (Ed). *Proceedings of the Sixth International Congress on Teaching Statistics*. Voorburg, the Netherlands: International Statistical Institute. CD ROM.

Statistics Canada and Organisation for Economic Cooperation and Development. (1996). *Literacy, economy, and society: First results from the International Adult Literacy Survey*. Ottawa: Author.

Thistlewaite, L. L. (1990). Critical reading for at-risk students. *Journal of Reading, 33(8)*, 586-593.

Utts, J. M. (1996). *Seeing through statistics*. Belmont, CA: Wadsworth.

Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association, 88*, 1-8.

Watson, J. (1997). Assessing statistical literacy through the use of media surveys. In I. Gal & J. Garfield, (Eds.), *The Assessment Challenge in Statistics Education* (pp. 107-121). Amsterdam, Netherlands: International Statistical Institute and IOS Press.

Watson, J. M., & Moritz, J. B. (2000). Development of understanding of sampling for statistical literacy. *Journal of Mathematical Behavior, 19*, 109-136.

IDDO GAL

Department of Human Services

University of Haifa

Eshkol Tower, Room 718

Haifa 31905, Israel