STATISTICAL LITERACY ASSESSMENT AND TRAINING OF GOVERNMENT PERSONNEL USING DATA FROM NATIONAL STATISTICS OFFICE: PHILIPPINE CONTEXT

Enriqueta Reston
University of San Carlos, Philippines
edreston@yahoo.com

The National Statistics Office (NSO)- Philippines is the main producer and provider of statistical data used by government agencies for various purposes. In this paper, data from NSO provided the context for various forms of statistical literacy assessments used as part of a government-sponsored statistics training program aimed to upgrade the statistical capability of government personnel. The assessment scheme includes a researcher-made Statistical Literacy Assessment Scale and performance-based assessment activities on data organization and interpretation which culminated with the writing and presentation of an evidence-based technical report. Using Watson's (2007) model for developing statistical literacy, participants' statistical literacy levels were further analyzed. The results inform both the government and the academe on the need to promote the development of statistical literacy for effective functioning in the workplace.

INTRODUCTION

The need for statistical capability building programs among adults in the workplace is now highly recognized in this information-driven society as individuals and organizations realize the importance of informed decisions based on evidence from data. Many professionals with diverse educational backgrounds with respect to statistical concepts, methods and tools are now faced with the challenge of applying statistical thinking and using statistical tools as a necessary part of work. In the government as well as in business and industry, many employees are faced with the need to gather, manage, organize, analyze, present and report data as part of their job functions. Further, national statistics offices and other government agencies provide data for use in planning, monitoring and evaluation, and in policy making for both the government and private sectors.

These challenges in the workplace as regards the management of data for various purposes provided the motivation for assessing adult statistical literacy among working professionals. Although there is no universally accepted definition of statistical literacy, statistics education research has produced an expanding view of what comprise this construct (see, for example, Wallman, 1993; Watson, 1997; Schield 2001; Gal, 2002; Watson and Callingham, 2003). In the context of adult statistical literacy, Gal (2002) defined statistical literacy as a term that refers broadly to two interrelated components, namely: (a) people's ability to *interpret and critically evaluate* statistical information, data-related arguments, or stochastic phenomena, which they may encounter in diverse contexts, and when relevant; and (b) their ability to *discuss or communicate* their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concerns regarding the acceptability of given conclusions. He further summed up statistical literacy as the "ability to interpret, critically evaluate, and communicate about statistical information and messages," and claimed that this is now recognized as an important skill for adults if they are to become informed citizens and employees.

Wallman (1993), on the other hand, contended that statistical results permeate our daily lives and traced many of the citizen's misunderstanding about sources of data, mistrust on the uses of data and misgivings about the value of statistics for guidance in public and private choices to the society's lack of statistical literacy. Thus, in any statistical capability building program for adults, the development of statistical literacy is very fundamental.

In this paper, Watson's (2007) model for developing statistical literacy as a complex hierarchical construct comprising six levels will be used as framework for analysis of adult statistical literacy. These six levels are as follows: (1) idiosyncratic-personal engagement with context using basic graph/table reading skills; (2) colloquial-informal engagement with context using basic chance, graph, and numeracy skills; (3) selective engagement with context involving qualitative interpretation of statistical ideas; (4) appropriate non-critical engagement with context using basic statistical skills; (5) critical-questioning engagement with context using appropriate

statistical terminology; and (6) critical-questioning engagement with context using sophisticated mathematical-statistical understanding.

PARTNERSHIP BETWEEN GOVERNMENT AND ACADEME

This past decade has seen a growing development of research and data-based consciousness among various sectors of the Philippine Society, particularly in the government and the academe. In response to the government's thrust for decentralized development, the need for build-up of research and statistical capability at the regional and local levels has recently been more strongly felt. The situation calls for more manpower support for generating and managing data useful for local area planning and evidence-based policy and decision-making by local government officials and policy-makers. In Region VII (Central Visayas Region), for instance, the importance of local area studies that respond to development needs has been highly recognized and highlighted in the Regional Development Research Agenda for 2009-2012. This, in turn, requires training and manpower development particularly in the area of statistical capability building for local government personnel and researchers.

In line with these developments, the National Statistical Coordination Board (NSCB), the policy-making and coordinating agency on statistical matters in the Philippines, has created the Regional Statistical Coordination Committee (RSSC) for each of the 17 regions of the country to strengthen statistical capability and coordination at the subnational level. Further, its Statistical Research and Training Center (SRTC) has forged Memoranda of Agreement with regional offices of the National Economic Development Authority (NEDA) and selected higher education institutions (HEIs) in the regions to "strengthen statistical training opportunities for local government units and regional government agencies with data analysis and planning responsibilities" (Central Visayas Regional Development Research Agenda, 2009).

As a statistics educator in one of the universities in the region and one of two selected private sector representatives of RSSC for Central Visayas Region, I have been directly involved in the conduct of statistical capability building training programs for government personnel, particularly those involved in local area planning for development. This paper presents the assessment of statistical literacy among the training participants as an initial step towards building continuing statistical capability training programs for government personnel.

OBJECTIVES

This paper presents and analyzes statistical literacy skills among government personnel who participated in two batches of a training program on Basic Statistics with Microsoft Excel during the period June-July 2009. In particular, it sought to assess specific statistical literacy skills that the training participants demonstrate as they deal with various types of data and assess their capability to reason with data. In addition, the participants' responses to the various forms of statistical literacy assessments were used to analyze and classify their statistical literacy level using Watson's (2007) hierarchical model for developing statistical literacy.

TRAINING PROGRAM ON BASIC STATISTICS FOR GOVERNMENT PERSONNEL

Training Needs Assessment

Prior to the training, a Training Needs Assessment (TNA) among government personnel in the region was done in order to serve as basis in planning and implementing training activities that are more responsive to the statistical capacity-building needs of government offices in the region. Two sets of TNA forms were made: one for government staff in selected government offices who are either data producers or data users, and the other for their respective supervisors. The results of the TNA showed the need to start with the basic tools of statistics and in developing competencies for the use of a computer package for data management. In this case, Microsoft Excel with PH Stat *add-in* was chosen due to its widespread availability in the government workplace.

The Participants

The training participants comprise 56 government personnel involved in data management and analysis for local and regional development planning, as well as for policy review and formulation. There were 29 and 28 participants in the first and second batches, respectively. Of these participants, 20 (35.7%) came from regional line agencies while 36 (64.3%) came from local

government units, particularly working in planning and development offices in the four provinces and cities of Central Visayas Region. None of the participants finished a statistics or related degree; most of their educational qualification is in the field of commerce, management and engineering.

Training Overview and Assessments

The training provided the participants a conceptual overview of statistics as a field of knowledge and focused on data management, organization, reduction, presentation and interpretation using descriptive statistics measures with the aid of *MS Excel* tools. It followed the presentation-workshop activity-assessment format. A ten-item pre-and post test prepared by SRTC was administered to the participants before and at the end of the training. In addition, statistical literacy assessment activities were conducted including a 15-item researcher-developed Statistical Literacy Assessment Scale (SLAS) and a performance-based group activity on constructing tables and graphs using MS Excel and interpreting data in tables and graphs.

STATISTICAL LITERACY ASSESSMENT RESULTS

Statistical Literacy Assessment Scale

The Statistical Literacy Assessment Scale (SLAS) is a 15-item scale designed to provide a measure of adult statistical literacy. Developed by Reston (2005), it has been used to assess adult statistical literacy among graduate students. While the data used in SLAS were taken from situations within Philippine context, the various skills encompassing statistical literacy which served as basis for the development of SLAS were drawn from the literature (Schield, 2000; Watson, 1997; Gal and Garfield, 1997; Rumsey, 2002; cited in Reston, 2005). All items were answerable in "yes-no-cannot tell" format; moreover, the participants were required to justify/explain their answers or raise a question/comment on the data presented. For the scoring, a three-point holistic rubric is used to assess accuracy of students' answers and statistical reasoning.

While the SLAS was originally used to assess the statistical literacy of graduate students, in this paper, the SLAS was administered to adult professionals working in the government as part of statistical literacy assessments in a statistical training program. Of 15 items in SLAS, four items using data generated by the NSO are presented in this paper to reflect different contexts for assessing statistical literacy. These data sets include the following: (1) Employment Status by Gender of the Philippine Labor Force (PLF) Survey in 2004, (2) Quarterly Unemployment and Underemployment Rates for the Fiscal Year 2003 from the 2004 PLF Survey, (3) data from the 2000 National Demographic and Health Survey, and (4) Inflation Rate by Commodity Group in the Philippines from 1999 to 2003.

Sample Items: Interpreting Counts, Percentages and Rates in Tables and Graphs

The sample item from the SLAS as shown in Figure 1 below assessed the participant's understanding of the concept of rates and percentages.

The table below represents the employment status by gender of the Philippines Labor Force (2004).

	Male	Female	Total
Employed (in thousands)	19,829	11904	31,733
Unemployed (in thousands)	2,311	1,574	3,885
Total (in thousands)	22,140	13,478	35,618

Source: Labor Force Survey (October 2004), National Statistics Office
    Is "the percentage of male who are employed" equal to "the percentage of male among those employed"?

Figure 1. Item No. 1 in the Statistical Literacy Assessment Scale

The results showed that among the 56 participants in the two batches, only 70% of the answered this item correctly in the yes-no-cannot tell response format. Moreover of these participants, 64% did not attempt to justify or provide reasons for their answer. For those who reasoned out, only 10% got a full credit of two points. Some of the justifications provided by the participants revealed some misconceptions, as shown in the following responses:

- *They are the same; it's just a matter of semantics.*
- *The % of male who are employed refers to the total male employed.*
- *They are not equal because of the different data reference on male employment vs. both gender employment.*

These sample responses revealed that even for a simple concept as *Percent*, many adults still lack the understanding on the idea for a reference base for computation and interpretation of the percent rate and percentages. Employment and unemployment rates are measured in percent; thus, it is important to understand and interpret figures in percent accurately.

In relation to Item 1, another item (no. 13) depicted trends in unemployment and underemployment rates in a graph as shown in Figure 2 below.

The graph below shows the trend in unemployment and underemployment rates in the Philippines for the month of January, April and July 2003 based on the preliminary result of the Philippine Labor Force Survey. *Based on the trend, is it expected that both unemployment and underemployment rates will steadily increase?*

Source: Labor Force Survey (October 2004), National Statistics Office
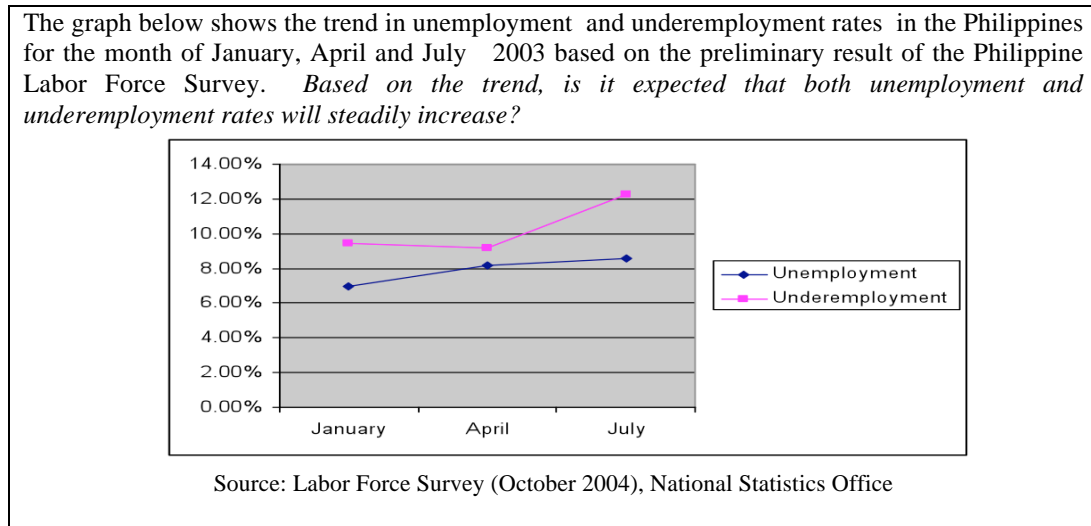
Figure 2. Item No. 13 in the Statistical Literacy Assessment Scale

Analysis of results for this item showed that only 15% of the participants answered this item correctly. Most of them answered "yes" and even for those who answered "no", the justifications for their answers revealed some misconceptions on the uses and interpretation of graphical displays. A sample of their justifications is shown below:

- *Yes, unemployment curve is already beginning to go down.*
- *No, because the trend for unemployment is fluctuating.*
- *There is no clear correlation of variables.*

These answers reflect the need for the participants to realize that since a graphical display is simply a descriptive tool, it cannot be used solely to make inferences about future trends of the data. The third answer that mentions "correlation" also reflects lack of understanding on the use of this term. These answers also indicated the lack of statistical thinking on variability and the role of chance when making statements about trends given limited sample information. These concepts were further clarified and illustrated during the training activities which led to further discussion on data collection issues and how to deal with sample information.

Sample Item: Interpreting Averages

The concept of the average pervades in many descriptions of quantitative data and the mean is the most commonly used measure of the average in many statistical reports. The sample item in figure 3 assessed the participants' understanding of the mean as a measure of the average.

In this item, 63% of the participants answered the item correctly as a "Yes." However, only 52% of them provided reasons or justifications for their answers, and of these, only four participants got a full credit for their reasoning. This result is relatively better compared to the previous two items, indicating a better understanding of the concept of the average compared to

percent and rates. As to the participants' reasoning, their responses reveal different levels of understanding and statistical literacy, as shown below:

- *Yes, since in the table, late 20's have an average of 1.24, late 30's have an average of 3.75 and late 40's 4.74.*
- *Yes, if we round off 1.75, 3.75, 4.74.*
- *No, the description does not correspond to the figures presented.*

The table below shows the mean number of children ever born by age of women in the Philippines in the 2000 National Demographic and Health Survey. According to the survey report, "*the mean number of children ever born shows that on the average, women gave birth to almost two children by their late 20's, almost four children by their late 30's and almost five by their late 40's.*" *Does this narrative interpretation agree with the figures shown in the table?*

Age, yrs	Mean No. of Children Ever Born
15 -19	0.07
20 – 24	0.68
25 – 29	1.75
30 – 34	2.86
35 – 39	3.75
40 – 44	4.16
45 – 49	4.74
Total 15-59	2.16

Source: National Statistics Office and Department of Health (2000)

Figure 3. Item No. 9 in the Statistical Literacy Assessment Scale

Sample Item: Interpreting Data Trends within Context

Finally, since most data encountered in the workplace are presented in table format showing variation in different components and time periods, ability to interpret trends from the data is also an important aspect of adult statistical literacy. This is illustrated in the sample item in figure 4.

The inflation rate is a measure of the increase in the prices of basic goods and services. An article in the business section of a national daily claims that *"Inflation rate is high in the Philippines but there are some recent signs of improvement."* The table below, showing the inflation rate (in per cent) by commodity group in the Philippines from 1999 to 2003, was presented to support the claim. **Do you agree with the author's claim?**

Commodity	1999	2000	2001	2002	2003
Food, Beverage and Tobacco	64.4	24.5	49.0	23.4	19.0
Clothing	75.5	30.8	44.9	31.2	21.9
Housing and Repairs	111.5	64.3	81.4	59.3	28.7
Services	125.0	139.6	139.6	59.3	61.1

Figure 4. Item No. 4 in the Statistical Literacy Assessment Scale

In this item, 47.4% of the participants answered the item correctly with 52.5% of them provided a reason or justification for their answer. Among their answers and reasons are as follows:

- *Yes, there is a downward trend in general.*
- *Yes, from 94% in 1999, inflation rate decreased to 32% in 2003*
- *Yes, the table shows a decreasing rate of inflation rate from 1999 to 2003.*
- *Cannot tell, there are also other commodity groups /or factors that influence increase/decrease in inflation*

These responses indicate varied interpretations when data are now presented in two levels, by commodity groups and time period in years. Attempts to generalize the trend shown in the data are evident but there is need for interpretation of the data in the context of the given inflation rates by commodity group. In the 2nd answer, the respondent attempted to summarize the inflation rates by commodity groups by taking an average measure which is not the proper way to deal with rates.

Statistical Literacy Levels Based on Watson's (2007) Framework

The participants' responses to these statistical literacy assessment indicate that most of them are either in Level 1 or Level 2 of the scale at the beginning of the training program. This means that either they demonstrate idiosyncratic-personal engagement with the given data sets in context using table/graph reading skills or of colloquial-informal engagement with context using graph and numeracy skills.

The SLAS was no longer given as a post test at the end of the training since answers to the items were discussed as part of the activities. Moreover, post assessment was done in the form of performance-based activities where participants were given six data sets in table format from NSO and were required to either interpret the table or construct graphical displays from the given data in the tables. As a final assessment, participants, in groups, came up with a coherent technical report about some chosen aspects of their work, demonstrating appropriate modes of data organization, presentation and interpretation. Aside from a written report, they prepared a Powerpoint presentation of their report and defended before a panel of evaluators comprising the resource persons and the head of SRTC. These performance-based assessment activities provided opportunities for evaluation of their statistical literacy at higher levels. Some have now demonstrated the third level characterized by selective engagement and a few have demonstrated critical-questioning engagement with context using appropriate statistical terminology.

CONCLUSION AND RECOMMENDATIONS

This paper presents a case of collaboration among various government agencies and the academe in a statistical capability training program aimed to improve statistical literacy among government personnel. In particular, it shows the important role of the National Statistics Office in providing data in various contexts as basis for statistical literacy assessments among government personnel. The findings reveal the need for government personnel to improve their statistical literacy level using data from various contexts, including those used in their workplace. Efforts to strengthen ties and collaboration among data providers and data users and consumers, as well as statistics educators, is extremely needed as a valuable step towards promoting statistical literacy among the citizens. It is recommended that statistical capability building programs among these government personnel and other professionals be developed and sustained, and that university programs in statistics for non-majors should address the development of statistical literacy needed for professionals in a wide variety of work settings.

REFERENCES

Gal (2002). Adults' Statistical Literacy: Meanings, Components and Responsibilities. *International Statistical Review, 70*(1), 1-51.

National Statistics Office (2009). Republic of the Philippines. http://www.census.gov.ph/.

Schield, M. (2001). *Three Kinds of Statistical Literacy: What Should We Teach?* Online: www. Augsburg.edu/~schield/MiloPapers/2002ICOTS6web.pdf.

Reston, E. (2005). Assessing Statistical Literacy in Graduate Level Statistics Education. *Paper presented at the 55th Session of the International Statistical Institute*, Sydney, Australia.

Wallman, K. (1993). Enhancing Statistical Literacy: Enriching our Society. Presidential Address. *Journal of the American Statistical Association*, *88*(421). Online:
http://client.norc.org/asapresidentialpapers.info/documents/Wallman%201992.pdf.

Watson, J., & Callingham, R. (2003). Statistical Literacy: A Complex Hierarchical Construct. *Statistics Education Research Journal*, *2*(2), 3-46.

Watson, J. (2007). A Model for Development of Statistical Literacy. Educating for Statistical Literacy.
Online: http://www.educ.utas.edu.au/users/watsonjm/prof%20workmodel.html.