

CONFERENCE OF EUROPEAN STATISTICIANS

UNECE Work Session on Communication and Dissemination of Statistics
(13-15 May 2009, Warsaw, Poland)

Topic - Managing communication and credibility

**PRESENTING AND COMMUNICATING STATISTICS: PRINCIPLES, COMPONENTS AND THEIR
QUALITY ASSESSMENT. A PROPOSAL**

Submitted by University of Florence, Italy¹

I. INTRODUCTION

Assessing quality of statistical activity pays great attention to many aspects mainly referring to data collection and production, data analysis, modeling. International institutions, like World Bank, Unesco (Patel et al., 2003) and Eurostat (2000) identified different attributes to be considered in evaluating quality of statistics (methodological soundness, integrity, serviceability, accessibility). At the same time, less attention has been paid to communication of statistics, which represents an important aspect of statistical activity and should be considered an integral part of data production and dissemination. The need to deal with this issue is significantly increasing especially in the perspective of the role the statistics have in ICT societies. Statistical communication can not be accomplished through improvising and approximating methods and instruments. It requires a combined and joint knowledge and expertise of statistical methodology, cognitive science, and communication.

Our work aims at identifying the fundamental aspects (namely ethics, rhetoric and aesthetics) involved in communicating statistics and the components of statistical communication. With reference to components, a particular attention is paid to communication codes, identified by (i) the outline, (ii) the tools, and (iii) the cloths. Subsequently, a conceptual framework is introduced aimed at constructing a model for statistical communication assessment, by riasserting the needs of standardized codes.

II. COMMUNICATION: INTEGRAL COMPONENT OF STATISTICAL WORK

Presentation of results represents an important aspect of statistical activities and should be considered not just at the end of them but constantly from the beginning. Communicating statistics can not be accomplished through improvising and approximating methods and instruments: statistics, considered a hard methodological discipline, does not allow short cuts, approximations, dilettantism, also in communicating its results. Attempts exist in order to measure the quality of output with quantitative indicators (punctuality of releases, number of errors discovered in published information, revisions in statistical database, etc.) or user's satisfaction surveys. However, less attention is paid to the *statistical results presentation* aspect, maybe because of efforts dedicated to the previous stage of the activities (data collection and production, defining research model, data analysis, modelling, and so on). Nevertheless, communication is not just an appendix of the core business focused on data production, but a key function that can determine the success or the failure of an official data provider. (Giovannini, 2008) A formula has been defined in order to define the *value added of official statistics* (VAS) (Giovannini, 2008):

¹ Prepared by Filomena Maggino (filomena.maggino@unifi.it) and Marco Trapani.

$$VAS = N * [(QSA * MF) * RS * TS * NL]$$

where

- N size of the audience
- QSA statistical information produced
- MF role of media
- RS relevance of the statistical information
- TS trust in official statistics
- NL users' "numeracy"

This detailed formula, including many relevant aspects like the role of media and users' numeracy, can be reconsidered by including also aspects concerning "quality" e "incisiveness" of the message:

$$VAS = f(N, QSA, MF, RS, TS, NL, QIP)$$

where

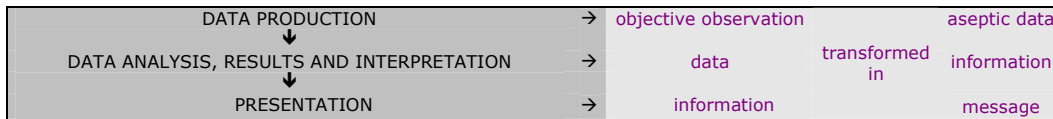
- QIP quality and incisiveness of presentation

This assumes a particular relevance if we consider (Giovannini, 2008) that 45% of Europeans has no faith in official statistics and that – at the same time – 69% of them believes in the necessity to know data concerning economics trends.

Statistics cannot be presented in an aseptic and impartial way by leaving honour and onus of data interpretation to the audience. At the same time, interpretation could be represented by different and equally correct perspectives ("the bottle is half-empty" or "the bottle is half-full"), which could be completed by additional information, e.g. in dynamic terms ("the bottle is getting filled up" or "the bottle is getting empty").

Whatever approach/statement will be assigned to presentations, the message will be nonetheless transmitted and interpreted by the audience in one of the possible ways (empty or full). The audience rarely will grasp only the pure numerical aspect by leaving of consideration any evaluation.

Since it is quite impossible to present data and results objectively, impartially, and neutrally, a step-by-step model should be followed defined by the following sequential elements:



III. COMMUNICATING STATISTICS

A. Fundamental aspects

Presenting statistics involves three fundamental aspects or pillars, related to (i) contents, (ii) aesthetics, and (iii) persuasion. They reflect the own base of classic rhetoric, according to the principles of teaching ("docere"), entertaining ("delectare") and moving ("movere"). The three aspects refer to three philosophical-scientific disciplines, ethics, aesthetics, and rhetoric. The following table summarizes the main aspects of each discipline.

Communicating statistics		Corresponding discipline	Bipolar constitutive elements	Dimensions of change
aspects	goals			
Content	using correct and accurate contents	Ethics	fair ⇔ unfair right ⇔ wrong legal ⇔ illegal honest ⇔ dishonest impartial ⇔ partial good ⇔ bad moral ⇔ immoral	"Fairness" changes through time and space
Appeal	allowing the message to be easily reached by the audience.	Aesthetics	beautiful ⇔ ugly pleasant ⇔ unpleasant agreeable ⇔ disagreeable	"Beauty" changes with reference to social canons
Persuasion	using instruments of persuasion ("theory of argumentation")	Rhetoric	preferable ⇔ not preferable convenient ⇔ inconvenient best ⇔ worst wise ⇔ unwise adequate ⇔ inadequate	"Preferable" changes across individuals
		↓ Theory of presentation		

The three related disciplines could represent the foundation of a “theory of presentation.” By following this, we can identify the abilities required in order to prepare an effective presentation of statistical results, can be led to the following fields: (1) Rhetoric (theory of argumentation), (2) Eloquence (public speaking), (3) Psychology of persuasion, (4) Gestalt Psychology, (5) Neurolinguistic programming, (6) Graphics and design, (7) Ergonomics, (8) Project Management, (9) Physiology.

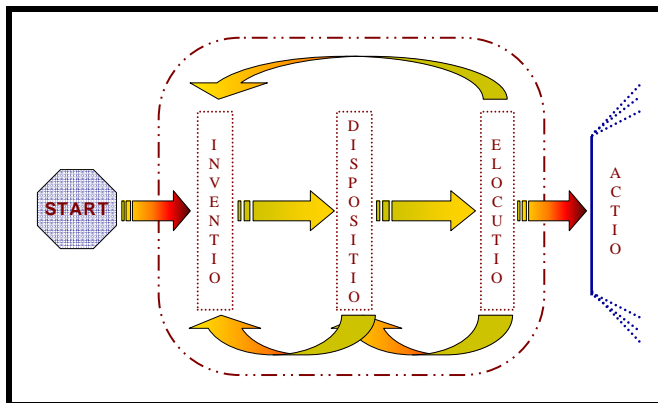
B. Main components

As known, elements composing any communication are: (i) transmitter, (ii) receiver, (iii) channel, (iv) message, (v) [transmitter’s / receiver’s] code, (vi) context, (vii) feedback, (viii) noise. The following picture shows the relationship between the elements involved in communication.

- (i) **Transmitter.** In communicating statistics, the transmitter is typically the statistician.
- (ii) **Receiver.** In communicating statistics, we could refer to receivers in terms of “audience.” In general, receivers of statistical communication can be distinguished in five categories, (a) experts, (b) politicians and policy makers, (b) students, (c) statistical data users, (d) not specialized.
- (iii) **Channel.** The channel represents the transmissive mean by which the message reaches the receiver. In statistics we can identify the auditory channel (“listening”, requiring oral explanation), the visual channel (“looking”, requiring explicative slides), and – when applicable – kinetic channel (“doing”, requiring practical exercises).
- (iv) **Message.** In communicating statistics, the message is represented by statistics (data, comments to data, and so on).
- (v) **Code.** It refers to the “technological” apparatus allowing communication. The apparatus has its grammatical, syntactical and stylist rules that, in statistical communication, refer to (i) the way statistics are reported (outline), (ii) the tools used in order to transmit statistics (tools), and (iii) the way in which statistics are dressed (cloths).
- (vi) **Context.** It refers to the situation or occasion in which the communication is accomplished. With reference to communication of statistics, we could identify different contexts, like seminars, conferences, meetings, press conferences, and so on. Each context has its own **setting**.
- (vii) **Feedback.** Feedback coming from the receiver allows the transmitter to adjust the code and to review and revise the message in order to make it clearer and more explicit.
- (viii) **Noise.** It is represented by whatever element disturbing the communication process. Noises could be identified in each of the previous elements. For example, lack of adequate lighting could represent a decisive noise in a seminar context. The goal is to reduce or eliminate its presence and effect.

C. Codes in statistical communication

I. Telling statistics: the outline. “Outline” refers to the process of telling statistics. In order to prepare it, five steps have to be carried out:



- A. ***Inventio* (invention).** The identified topics are distinguished mainly by pointing out which are relevant, positive, or pleasant and by overshadowing and leaving out others considered irrelevant, negative, or

unpleasant. In this perspective, the presentation should produce evidences supporting the selected topics. Evidences can be distinguished into the following categories:

- technical evidences, to be found by turning to rhetoric techniques,
- extra-technical evidences, shown by facts, shared rules, indisputable standards, scientific confirmations, valuable and prestigious citations supporting and strengthening the presented “story”.

In order to organize the topics to be presented, it is possible to refer to the well-known 5 W:

- o *Who* → the subject of the telling
- o *What* → the fact
- o *When* → the time location
- o *Where* → the field location
- o *Why* → the causes

We can add also the followings: “in which way” and “by which means.”

- B. **Dispositio (layout)**. This stage aims at ordering the topics, previously identified for presentation. The sequence of presentation should follow a logical order, appropriate to topics we are presenting and to the results we would obtain. The presentation sequence can follow one of the following basic criteria:

Criterion	Description	Positive aspects	Negative aspects
rising / growing	From weaker topics to more convincing ones	Audience recalls the favourable topics	The first impression could not be favourable
decreasing / declining	From more convincing topics to the weaker ones	It draws audience’s attention	Audience could remember the weaker arguments
Nestorian / Homeric	The more convincing topics are presented at the beginning and at the end; the less valid ones are left at the centre (from the IV book the Iliad: Nestor puts the weaker troops at the centre of the array).	It is the most effective	It is difficult to be organized

The argumentation plot could be seen as a woof in which each element represents a yarn. The “ideas concatenation” metaphoric figure asserts that argumentation’s validity depends on the weaker yarn.

However, in organizing the argumentation in a logical tissue it should be taken into account that the organized union of several yarns allows a woof to be obtained which is more resistant than their mere sum (Perelman, 2005). A possible plot is the following (found already in Quintilian), expressed in terms of advising:

- introduce what you want to tell → exordium
- tell that → development
- mention what you have said → summary / epilogue

Obviously, each of the different layout approaches has pros and cons and is more or less suited to different situations and audiences. The choice should take into account expected objectives, kinds of argument, audiences, and, last but not least, presenter’s preferences.

- C. **Elocutio (expression)**. While *inventio* allowed arguments to be argued and *dispositio* allowed them to be put in order, *elocutio* allows each piece of the presentation to be prepared by selecting words and constructing sentences. This task is strictly related to rhetoric. One of the most important choices to be taken concerns the title of the presentation: except for the cases in which rules and procedures exist, the title should be prepared according to an explicative idea. Sometimes it could be shaped like a slogan to be used again along the presentation. A classical classification *rhetoric figures* (found already in Quintiliano) is the following:

Definition	
thinking	change in words’ or propositions’ invention and imaginative shape
meaning (or tropes)	change in words’ meaning
diction	change in words’ shape
elocution	choice of the most suitable or convenient words
construction	change in words’ order inside a sentence
rhythm	phonic effects

This classical classification has been considered unsatisfying and arbitrary and many other classifications have been defined. Among them, the following seems to be in our opinion more useful, suitable, and appropriate in telling statistics: a) image figures, b) repetition figures, c) technical figures.

- D. **Actio (execution)**. It concerns the way in which the telling, in terms of (i) introduction, (ii) development, (iii) comments, (iv) time/space use, (v) ending, and the receivers’ feedback (*questions & answers* stage) are managed.

The outline can not be developed through a linear progression but through a cyclic process allowing previous steps to be run through again in order to check, improve, correct, integrate, and review before reaching the “action” stage.

II. Depicting statistics: the tools. “Tools” refer to all available instruments aimed at depicting statistics, by constructing and using graphs, tables, and pictographic supports. Graphical representations are useful and advantageous instruments in order to better communicate statistics. Graphical representations may have a double function, presenting and describing results and allowing a quick and synthetic interpretation of the observed phenomenon and its trends. In this perspective, statistical graphics should be considered as a good combination of text, tables and charts (Statistics Canada, 2003). The evolution of statistical graphics has found a great boost thanks to three main factors: (i) invention of new techniques suitable to complex data structures; (ii) new research results concerning human psycho-physiological perception suggesting correct strategies aimed at presenting quantitative information; (iii) availability of computer (*software* and *hardware*) instruments allowing complex graphical applications to be managed.

Even if a clear limit between advantages and disadvantage in using graphs does not exist, general guidelines can be identified helping in determining the best strategies in depicting statistical information. The goal is to make sure and preserve graph’s capacity to autonomously communicate the message.

III. Dressing statistics: the cloths. “Cloths” refer to the process of dressing statistics. Communicating statistics should be supported also by other elements:

- text arrangement, related the disposition of the text on the used mean (e.g. slide or page)
- characters and fonts: the choice of the character font should be consistent with the spirit and character of the presentation
- colours, which use should take into account their perception, their possible cultural meanings and the used means. Particular attention should be paid in using colours in graphs where they represent a further code element
- other graphical aspects and effect. In this perspective, it could be functional using photos and clipart to give the audience time to “come up for air.” This would allow the audience to reflect and digest. This is especially the case if the photo or clipart is humorous. (Kosslyn, 2006; 2007)

IV. ASSESSING STATISTICAL COMMUNICATION

A. Conceptual framework

Evaluating errors represents one of the most important components of the statistical works and is aimed at assessing a statistical result. In this work, we assume that statistics represent the most important part of the statistical message and have been produced by respecting all *best practice* codes. Similarly, the assessment of communicating statistics should be accomplished by evaluating the level of **general noise**. Any assessment is intimately connected to the possibility to verify and control what has to be assessed. In the ambit of communicating statistics, defining and identifying the following aspects allow the assessing task to be carried out:

- **The dimensions to evaluate.** In our perspective, the assessment (and the consequent adjustment and/or adaptation) concerns the transmitters and their codes. In this context, we will concentrate our attention on the transmitter’s code, specified in terms of (i) outline, (ii) tools, and (iii) cloths.
- **The evaluating criteria.** Criteria are related to the (A) *suitability / consistency*, (B) *correctness*, and (C) *clarity* of the code according to the components of the transmission process. The criteria refer to the capacity of the transmitter in using use the code.
- **The components of the transmission process.** The identified dimensions will be evaluated through the defined criteria with references to the transmission process: (i) the receiver/audience (and its receiving code), (ii) the available channel, and (iii) the available context and setting, and, in some way, (iv) the contents message.

Consequently, the assessing conceptual model previously defined can be consistently develop an assessment table:

Assessment table			Evaluating criteria									
			(A) suitability			(B) correctness			(C) clarity			
			Audience	Channel	Context	Audience	Channel	Context	Audience	Channel	Context	
Dimensions to evaluate	Transmitter's code	Outline	Invention									
			Layout									
			Expression									
			Execution									
	Tools	Tables										
		Graphs										
		Pictograms										
	Cloths	Text arrangement										
		Characters/fonts										
		Colours										
		Other effects										

The simplest assessing approach is to evaluate the presence/absence of each combination and could concern the whole presentation or individual sections of it.

B. Need of standardized codes

As previously said, in order to verify that codes' requirements are met (assessing process), the code needs to be standardized. Standards should provide explicit, definite, unambiguous, and shared rules. The code is standardized if (i) it meets and respects the underlying logic of the message's content (consistency),² (ii) the rules are clearly and easily applicable,³ (iii) different individuals can use the same code by obtaining comparable results. The lack of standards in codes, communicating statistics will depend upon subjective judgments and evaluations. Developing and adopting standardized codes allow transmitters to warrant:

1. *objectivity* of data presentation, by avoiding introduction of any subjective components,
2. *comparability* between different presentations and along time,
3. *economicity* and *efficiency* in preparing presentation,
4. *generalization*, by avoiding any kind of "adaptability" of codes to "subjective" messages,
5. *understanding* of data structure.

In communicating statistics, the standards can be extrapolated from good practices interpreted in terms of evaluating criteria. Defining and adopting standards should not impede any experimental investigation and proposal of new codes (e.g. new graphs).

V. THE WAY FORWARD

Emphasizing communicative aspects does not aim at exasperating the importance of communications with respect to the rest of statistical activities. ICT societies increasingly require statistical information. In order to attribute more value to statistics and to increase statistics' impact on reality, two processes need to be assessed: (i) data production and analysis and (ii) communication. While rightly a great attention has been and is paid to many aspects related to the former process, more work needs to be done in order to improve the latter. In this perspective, we should take into account that one's goodness does not substitute the other's badness. Assessing statistical communication takes for granted that all the previous stages of data production have been accomplished by respecting all *best practice* codes. In order to improve the communication approach, the assessment needs standardized methods and techniques. The presented work summarizes the first stage of our study, aimed at defining the framework for assessing the quality of statistics communication. The goal is to define the assessment technology by discussing with all the willing colleagues.

References

² For example, in representing a specific content, the applied code should respect the intrinsic meaning of the content. E.g., a mean, whose meaning is strictly connected to the concept of "centre of gravity", should be always represented into a graph by a point (not by a bar!)

³ E.g., graphs should be constructed by taking into account the correct definition of scales, dimensions, relationship, and orientation between X-axis and Y-axis.

Eurostat (2000) *Definition of quality in statistics and Standard quality report*, Eurostat.

Gioannini E. (2008) *The role of communication in transforming statistics into knowledge*, OECD, paper presented at conference “Innovative Approaches to Turning Statistics into Knowledge”, Stockholm, 26-27 May 2008.

Kosslyn S.M. (2006) *Graph design for the eye and mind*, Oxford University Press, New York.

Kosslyn S.M. (2007) *Clear and to the point*, Oxford University Press, Oxford.

Patel S., M. Hiraga, and L. Wang (World Bank) D. Drew and D. Lynd (Unesco) (2003) *A framework for assessing the quality of education statistics*, World Bank – Development Data Group and UNESCO – Institute for Statistics.

Perelman C. (2005) *Teoria e pratica dell’argomentazione* (a cura di G. Fornari Luvarà), Rubettino, Soveria Mannelli.

Statistics Canada (2003) *Quality guidelines*, Fourth Edition, Catalogue No 12-539-XIE, Statistics Canada.