Presenting and communicating statistics: principles, components and their quality assessment. A proposal

eurostat

NTTS - Conferences on New Techniques and Technologies for Statistics

The meeting place for Research in Official Statistics

POSTER SESSIONS
FULL PAPERS

Brussels, 18-20 February 2009

eurostat
EUROPEAN COMMISSION

Presenting and communicating statistics: principles, components, and their quality assessment. A proposal

Filomena Maggino[1], Marco Trapani[2]
[1] Università degli Studi di Firenze (Italy), email: filomena.maggino@unifi.it
[2] Università degli Studi di Firenze (Italy), email: trapani.marco@gmail.com

Abstract

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.

Keywords: quality of statistical presentation, statistical presentation tools.

1. 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):

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

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:

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

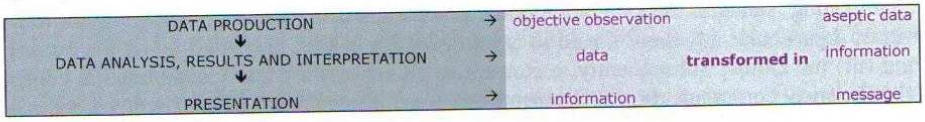
where

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.

Information, in itself, is not knowledge[1]: moving information up to knowledge needs to activate complex knowledge processes. *"The virtues of a good statistician, therefore, involve not only the skills of a good detective, but also the skills of a good storyteller. As a good storyteller, it is essential to argue flexibly and in detail for a particular case; data analysis should not be pointlessly formal. Rather, it should make an interesting claim by telling a tale that an informed audience will care about, doing so through an intelligent interpretation of data."* (Abelson, 1995)

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:



2. Communicating statistics

2.1 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)

[1] Albert Einstein, cited by Giovannini (2008).

Graphics and design, (7) Ergonomics, (8) Project Management, (9) Physiology.

2.2 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. Obviously, wider and larger the channel, higher the probability for the message to be reached in a complete way. For this reason, using all the possible and foreseeable channels has become the main goal of communication. 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).[2]

(iv) **Message.** In communicating statistics, the message is represented by statistics (data, comments to data, and so on). Assessing quality of statistical activity pays great attention to many aspects referring to mainly data collection, production, data analysis, modelling, and so on. Many international institutions, like World Bank & Unesco (Patel et al., 2003) and Eurostat (2000) have identified different attributes to be considered in evaluating quality of statistics, such as, methodological soundness, integrity, serviceability, accessibility.

(v) **Code.** It refers to the whole "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). Code could represent a double barrier, at the moment of transmission (the message is codified by the transmitter) and at the moment of reception (the message is decoded by the receiver). Consequently, in our perspective, we should make sure that the "right" code is utilised. The "right" code is that that minimizes misunderstandings and maximises understandings of the message.

(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**. For example, with reference to seminars, setting concerns rooms, tables, and so on.

(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.

2.3 Codes in statistical communication

2.3.1 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), (b) *dispositio* (layout), (c) *elocutio* (expression), (d) *actio* (execution). 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.

[2] Attempts could be done to involve also other channels, e.g. olfactory, in order to underline pleasant or unpleasant aspects of the message.

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:

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

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

In order to identify the argumentation model the following items should be defined

1. what is to be demonstrated, consistently with the message
2. which are the evidences in favour, to be sustained and reinforced
3. which are the evidences against, to be dismantled and refuted

The "treatise of argumentation" provides with six "loci" to be considered in telling "stories" (Ellero, 1997):

(i) locus of **quantity** (something is more important than another because of quantitative reasons, e.g. many people's benefit is preferable to few people's benefit)
(ii) locus of **quality** (something is more important because of it is unique, irreplaceable or fragile)
(iii) locus of **order/sequence** (preceding event is more important than the subsequent event, the cause than the effect, the end than the means)
(iv) locus of **existent** (concrete and real thing is more important than probable things; *line of low-risk* falls within this argumentation)
(v) locus of **essence** (higher value are assigned to subjects representing typologies or functions)
(vi) locus of **person** (preferring individual values, like merit or autonomy)

Further, different techniques exist in order to support getting "good ideas", like brainstorming, chiasmus or association of words.

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. The following table shows the main layout approaches by listing their positive and negative features.

Approaches	Right for ...	Risks
Deductive	Transmitting definitions Investigating thoroughly (for experts) Giving information in short time Transmitting sense of presenter's competence	Starting parts could turn out to be incomprehensible Boredom Language and approach could turn out to be too much theoretical
Inductive	Explaining complex concepts Novice audience	Trivializing It may take long time
Time progression	Presenting cycle phenomena Presenting projects	Too much analytical Not calibrated and suited to all kinds of audience
From different point of view	Enlarging views Involving intersectorially Analysing deeply	It may take long time Partiality It may lack a general view
Advantages and disadvantages	Analysing in detail Enlarging views Catching hidden advantage/disadvantage	Trivializing Too much analytical It may take long time
Problems	Showing efficacies Convincing and persuading Showing in a easy way	It may fail in reaching all the arguments It is not appreciated by theorists It is difficult to be planed
Top-Down	Training novice and students Presenting complex arguments and subjects	It could turn out to be too much easy or too much difficult It may take long time It makes difficult to reach the core of the matter

In particular cases a mixed approach can be adopted, especially when presentations concern extensive and complex subjects, in order to avoid the *boredom-effect* risk, even if a mixed approach could appear confused in its logic.

Expression (elocutio)

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
Figures of	**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.

This taxonomy production seems aimed at distinguishing language deformations (rhetoric) from the pure and simple language – "*an ingenuous and without-artifices discourse*" (Groupe μ, 1970). The latter is considered the aim of the scientific language. Nevertheless, should statistical language shun rhetoric? Even if rhetoric figures should be used *"cum grano salis,"* we should take into account that rhetoric is integral part of language in which almost all is metaphor (Lakoff & Johnson, 1980), it seems quite impossible to reduce language to an aseptic form, without a reference code. The language adopted in communicating statistics should be (a) appropriate to the audience and (b) consistent with the message to be transmitted. Besides, in telling statistics, special attention should be paid to (a) wording (choice of proper words to be used), (b) languages (use of specialist terminologies), and (c) tongues (use of languages in international contexts).

Execution (actio)

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.

2.3.2 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:

a. invention of new techniques suitable to complex data structures;
b. new research results concerning human psycho-physiological perception suggesting correct strategies aimed at presenting quantitative information;
c. 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.

Communicating statistics through graphics

A. Functions

Graphics have relevant role in transmitting any kind of information. Clark e Lyson (2004) tried to classify graphics functions:

Function	Definition
Supporting attention	Graphics that draw attention to important elements and that minimize divided attention.
Activating and building prior knowledge	Graphics that engage existing mental models or provide high-level content overview to support acquisition of new information.
Minimizing cognitive load	Graphics that minimize extraneous mental work imposed on working memory during learning.
Building mental models	Graphics that held learners construct new memories in long-term memory that support deeper understanding of content.
Supporting transfer of learning	Graphics that incorporate key features of the shared work environment; graphics that promote deeper understanding.
Supporting motivation	Graphics that make material interesting and at the same time do not depress learning.

By using a typical rhetoric figure (anaphora), Tufte (2001) summarises criteria that should guide in designing a statistical graphical representation in the following way: *"graphical excellence is the well-designed presentation of interesting data – a matter of substance, of statistics, and of design. Graphical excellence consists of complex ideas communicated with clarity, precision and efficiency. Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space. Graphical excellence is nearly always multivariate. And graphical excellence requires telling the truth about the data."*

B. Principles in graphing statistics

According to Kosslyn (2006; 2007), eight basic principles should be taken into account in constructing effective graphs. The principles can be distinguished in three macro categories:

Categories	Message should	Principles
(i) Connect with the audience	connect with the goals and interests of the audience	1. relevance 2. appropriate knowledge
(ii) Direct and hold attention	lead the audience to pay attention to what is important	3. salience 4. discriminability 5. perceptual organization
(iii) Promote understanding and memory	be easy to follow, digest, and remember	6. compatibility 7. information changes 8. capacity limitations

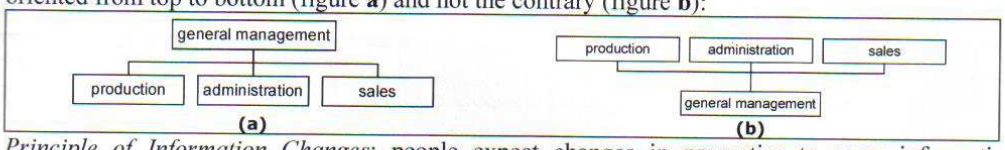
1. *Principle of Relevance*: communication is most effective when neither too much nor too little information is presented. (a) A surplus of information creates lack of attention, runs the risk of "not-reaching the point" and of drawing hurried conclusions. Examples: graphics with too many elements, frames, too long captions. (b) Being too much concise runs the risk of appearing superficial and useless. Examples: graphics with too few captions, values, numbers.
2. *Principle of Appropriate Knowledge*: communication requires prior knowledge of pertinent concepts, jargon, and symbols. The informative effect of graphics will be lost if graphical elements (not only contents) need to be explained.
3. *Principle of Salience*: attention is drawn to large clearly perceptible differences. A graphical element, distinctly different from all the others, will draw more attention. Examples: a colour spot, an element represented in bright colour in comparison with all the others in pastel colours, a

graphical element in motion, a word dimensioned bigger than the others; images recreating human features, a graphical element reproducing and simulating a movement.

4. *Principle of Discriminability*: two properties must differ by a large enough proportion or they will not be distinguished. This could depend also on environmental conditions or individual capacities/incapacities.
5. *Principle of Perceptual Organization*: people automatically group elements into units, which they then attend to and remember. This principle derives directly from Gestalt psychology. This phenomenon is related to the human capacity (*pareidolia*) to perceive as significant also vague and random stimuli (often an image or sound). An example, which illustrates this principle, is using colour in order to group elements. Elements represented through the same colour will be seen as a group. Using the same colour for all titles and a different colour for all text entries will clearly group the material into these two categories. Using the same colour allows corresponding elements to be paired. Bars concerning the same variable, the same group should show the same colour.
6. *Principle of Compatibility*: a message is more easily understandable if its form is compatible with its meaning. Example in the following picture, the first group of words is much more quickly readable than the second:

GREEN RED BLUE YELLOW
YELLOW BLUE GREEN RED

This effect is also connected to metaphors of orientation: an organization chart is expected to be oriented from top to bottom (figure **a**) and not the contrary (figure **b**):



7. *Principle of Information Changes*: people expect changes in properties to carry information. Consequently, representation should be "homogeneous" and "consistent" and each graphical element (colours, bullets, backgrounds, and so on) should keep the same reference meaning and should be used consistently along the whole presentation. This criterion allows each element to carry on, underline, and emphasize meanings, changes, and so on. An example is using the same terminology in labels, as well as in the surrounding text and spoken words. Using different terms in a display, in text, and in what is said may lead the audience to wonder if the presenter means different things. Attempting to distinguish these differences will definitely tax the cognitive capacities of the audience.
8. ***Principle of Capacity Limitations***: people have a limited capacity to retain and to process information, and so will not understand a message if too much information must be retained or processed. In constructing graphs, the leading ideas should be sobriety and minimalism; decoration and "baroquisme" should be avoided.

C. Perception of statistical graphics

Graphical approach to statistics involves a (a) graphical component (*coding* numerical information in imagine) and (b) perceptual component, *decoding* information included in graph.

If coding has been performed in a correct, effective, and efficient way, subsequently the decoding process would be accomplished in a proper way. The decoding channel is extremely delicate and requires coding process to be carefully accomplished in order to avoid any distortion of information to be communicated. Consequently, it is important that graphs preparation would take into account not only statistical matters but also issues connected to perceptual and cognitive psychology. A correct statistical graph should allow the receiver to (a) recognize the code, (b) recognize regularities, and (c) carry out comparisons and identify differences.

D. Adopting graphs

The role of graphical representations in the ambit of statistical analysis is mainly to:

- summarize in a clear, understandable and synthetic way also complex data
- make interaction between researcher and data easier
- provide important directions for future analytical developments or for programming and intervention.

In using graphs, two steps should be considered: choosing and constructing. In order to correctly accomplish both tasks, the following issues should be considered

(A) **Choosing a graph**	
by taking into account	by preferring
1. number of involved variables	a simple graph with reference to the audience
2. level of measurement of variables	a clear graph instead of an attractive one
3. kind of statistical information to be represented	a correct graph with reference to data
(B) **Preparing a graph**	
By	
correctly defining and showing scale/s	⇨ scale definition
reducing dimensionality as much as possible by showing few variables for each graph using no meaningless dimensions	⇨ dimensionality
using colours consistently with statistical information	⇨ colours as statistical codes
rounding up and down through standard criteria	⇨ rounding off values
dynamic perspective should reflect a dynamical phenomenon	⇨ dynamics presentation
few elements as possible. Wise use of legends and captions	⇨ legibility

The choice of statistical graphs should be taken according to different sequential criteria:

- number of variables involved in the analysis:
  - one variable (→ univariate technique)
  - two variables (→ bivariate techniques)
  - more (→ multivariate techniques)
- level (nominal, ordinal, or metric scale) and process (discrete or continuous) of measurement
- kind of statistical information to be represented: (i) summary statistics (frequencies or statistical indexes), (ii) density information (quantile or probability), (iii) individual values, (iv) functions.

The following table allows the proper graph to be identified according to the three criteria.

How many variables are involved?	What is nature of data?	What kind of statistical information has to be represented?	What kind of graph?
One (*)	categorical or ordinal categories	counts / frequencies (concentration of data points in each category)	Bar chart (**)
			Dot plot
			Circle plot / Pie chart
		magnitude value for each case	Bar chart
			Polar chart
	quantitative	counts / frequencies (concentration of data points in each interval across the range of the distribution)	Histogram (**)
			Circle plot / Pie chart
			Box-plot
			Dot plot
Two (*)	categorical or ordinal categories	counts / frequencies (concentration of data points in each cross category)	Line chart
			Area chart
			Compositional bar chart
			Ring chart
	mixed data (category/ordinal and quantitative)		Mosaic diagram
	mixed data (ordinal and quantitative)	statistical index (e.g. mean) recorded in each category	Dot plot (Hilo chart)
	quantitative		Line chart
		magnitude values jointly recorded by each case	Scatterplots
More	quantitative		Compositional bar chart
			Icon plots
		magnitude value for each case	Coordinate display
			Function plots
			Analytical approaches

(*) univariate and bivariate techniques can include additional variables, performing as grouping variables (e.g. sex, marital status). In this case only relative frequencies can be represented.

(**) each bar should represent a quantity presenting a cumulative feature, like frequencies or counts. Non-cumulative values, like means, cannot be represented by bars or histograms. In particular, means – whose statistical meaning is related to the concept of "center of gravity" – should be shown as points.

Others can be identified according to specific needs and to graphic capacity of the authors, even if using new graphical representation should be carefully considered. In some case, the statistical result is shaped directly through particular analytical graphic techniques (e.g. tree diagram), showing out information that can be communicated through numbers with difficulty. Besides the traditional statistical graphs, we can identify other graphical instruments allowing some results to be presented (traffic lights, dashboards, pictograms, maps, project management charts, hierarchy charts, process flow diagrams, star diagrams, Deming's diagram).

2.3.3 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)

3. Assessing statistical communication

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**. Generally speaking, 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.

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),[3] (ii) the rules are clearly and easily applicable, [4] (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).

4. 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

Abelson, R.P. (1995) *Statistics as principled argument*, Paperback, USA.

Clark R.E. – Lyons C. (2004) *Graphics for learning*, Pfeiffer, San Francisco, USA.

Ellero M.P. (1997) *Introduzione alla retorica*, Sansoni Editore, Milano.

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

Giovannini 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.

Groupe μ (1970) "*Rhétorique générale*", Larousse, Paris.

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.

Lakoff G. & Johnson M. (1980) *Metaphors we live by*, University of Chicago Press.

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.

Tufte E.R. (2001) *The visual display of quantitative Information*, Graphics press, Cheshire, Connecticut.

[3] 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!)

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