Working statistician versus new graduates in survey statistics

Giulio Ghellini
University of Siena Dept. Quantitative Methods
P.za S. Francesco 8
53100 Siena, Italy
E-mail: ghellini@unisi.it

Vijay Verma
University of Siena Dept. Quantitative Methods
P. za S. Francesco 8
53100 Siena, Italy
E-mail: verma@unisi.it

Abstract

Parameters of short-term training in sampling and survey statistics are naturally very diverse. In this paper we consider a very particular type of training arrangement: a relatively small, short-term workshop for participants with heterogeneous backgrounds, with the objective of moving the participants closer to what may be called "good survey statisticians". Heterogeneity - an intriguing challenge on one side, but also a potential opportunity on the other - is the *fil rouge* of the training situation we describe. In order to bring into focus the training paradigm we describe, we view our trainee 'input' and 'output' as ideal-types in the Weberian sense: the input is a mixture of practitioners and theoreticians. As to heterogeneity being an opportunity, the central concept is that of *cross-fertilization*: how to best utilize incoming heterogeneity in background and skills. We argue that *the quality issue must serve as the framework*: it determines the type of issues and materials to be included in the course, and the whole approach to training. This approach helps in a balanced treatment of theoretical and practical aspects. We find that real surveys provide the most effective instrument for organising the training, and conclude by noting a few additional practical aspects of the training workshop on survey statistics.

1 Introduction

"La notte scorsa ho pensato a che cosa renda un uomo capace o incapace di fare scoperte, ma è un problema che mi lascia molto perplesso. Vi sono molti uomini dotati di grande intelligenza , notevolmente superiore a quella degli scopritori, i quali tuttavia non creano mai nulla. Penso che la condizione necessaria sia la ricerca continua delle cause e del significato, di tutti gli avvenimenti naturali. Ciò implica un'acuta

osservazione e richiede la massima conoscenza possibile degli argomenti studiati" (C. Darwin, letter for his son Horace 15/12/1871)[1].

Parameters of short-term training in sampling and survey statistics are naturally very diverse, depending on the context, objectives, resources, and many other practical considerations; for examples, see Morganstain and Marker's conversation with J. Waksberg, (2000), Kish (1978, 1996), Verma (1996, 2001). In this paper we consider a very particular type of training arrangement: a relatively small, short-term workshop – perhaps of 3-5 weeks duration, 15-25 participants with heterogeneous backgrounds (in statistics, and also in other respects) – with the objective of moving the participants closer to what we term (and explain below) as good "survey statisticians". Our primary motivation is a general one: to bring out important considerations in the design and conduct of non-academic training for survey statisticians, with the specific concern of promoting *good practices* for training programmes which are international in setting and in the context of official statistics.

Heterogeneity - an intriguing challenge on one side, but also a potential opportunity on the other - is the *fil rouge* of the training situation we describe. As implied by the title of this paper, our emphasis is on the contraposition between two very heterogeneous groups of participants: *working statisticians* closely involved in the practical aspects of survey, but typically with a limited knowledge of statistical theory; and *fresh graduate statistician* with (hopefully) a good theoretical background, but with little experience in the actual design and implementation of real surveys. Consequently, the context and methods of training also need to be heterogeneous: a judicious mix of theory and practice; of general principles and specific examples; of classroom instruction and discussion, balanced against group projects and assignments. To achieve this, the providers (teachers, instructors) must also present a mixture of skills and experience; indeed the two authors of this paper themselves a good example of this – the first one a survey sampling teacher at an Italian University, the second one a researcher in international survey design and teaching.

The ultimate scope of such training course will be at least to inoculate in the partecipants some principle that have to be familiar for a survey statistician: (i) learning to think in terms of the overall quality of survey data; (ii) appreciating flexibility in considering what is acceptable and how best to meet the needs for the survey; (iii) understanding that it is not possible to be a good practical survey statistician without having a good grasp of the underlying theory which have to guide all the survey operational choices.

The paper is organized as follow. In Section 2 we introduce the type of training situation considered in this paper. Section 3 argues that quality assurance of survey data serves as the framework for training, and Section 4 discusses how reference to a particular survey , or a group of surveys, can serve as an organizing tool for the training workshop. As noted in Section 5, heterogeneity is the essential condition of the type of training courses we are discussing, which represents a challenge and an opportunity at the same time. We conclude in Section 6 by noting a few additional practical aspects of the training workshop on survey statistics.

2 The training paradigm: input and the output "ideal types"

In order to bring into focus the training paradigm we describe, we view our trainee 'input' and 'output' as ideal-types in the Weberian sense (Figure 1). The input is a mixture of practitioners and theoreticians.

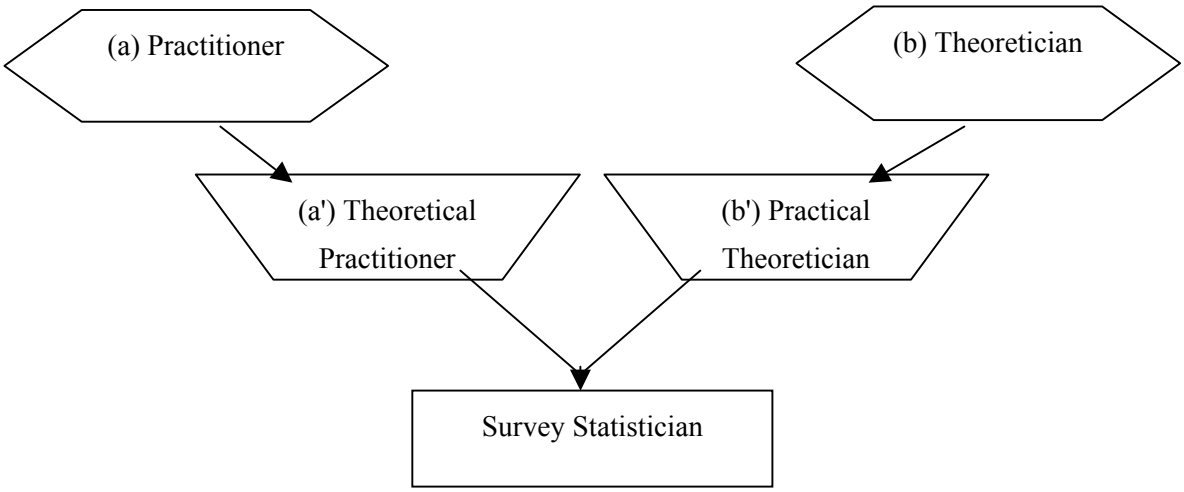
---

[1] We have found this quotation only in an Italian translation of the Charles Darwin's autobiography. It will be appreciated if some reader can point to us the original source. In any case, we will provide a rough translation of these beautiful words at the presentation.

A *practitioner* as our ideal-type (a) is someone who is familiar - through experience (typically as an official statistician) - with practical conditions of survey implementation and with the consequences of particular choices in survey design and procedures on the management and quality-control as well as cost aspects of the operations; plenty of common-sense is the hallmark of a good practitioner. On the other hand, however, our practitioner is typically lacking in statistical skills required for choosing among appropriate classes of designs (e.g. the survey structure), and in working out specifics of the design chosen (e.g. its sample size and allocation) – something for which a good grasp of sampling theory is essential.

The ideal-typical *theoretician* (b) has skills - through academic learning and research - in sampling theory which can guide in the evaluation of designs in terms of their statistical efficiency, and thus in choosing between design options (albeit in this limited sense) and then working out technical details of a chosen design. On the negative side, however, the theoretician lacks awareness of practical conditions and constraints of survey implementation, and of all the consequences of a particular choice on the total quality (and cost) of what the survey would yield. The typical weakness of a theoretician, as envisaged here, is a lack of intuitive understanding and appreciation of what is important (central) in practice and what is trivial (peripheral) – something for which pure theory, with its unavoidable simplifying assumptions, is often a poor guide.

Figure 1. The training paradigm: input and output "ideal-types"



Imperfect theoretical knowledge without practical underpinning can actually be quite dangerous. Here is an example - we actually found 'theoretical experts' arguing thus: interviewer variance goes down with the number of interviewers, so the ideal solution is to have one per interview. But a practitioner will not even consider such a 'solution', knowing that with more interviewer to train and manage, the cost goes up and the quality down. On the other side practical choices without theoretical knowledge could also have a very negative effect. An example of that can be found in the already cited interview of Waksberg (2000) where he speaks about large oversampling for rare population (a common practice among practitioners) in a well known survey of the National Centre for Health Statistics: "the results was a large increase in the variances for the total population and only a marginal gains for the specific population they were oversampling. That's usually an important issue, how to reconcile the oversampling for the specific rare population with the effect on other parts of the population".

The aim of the training workshop is to convert the first ideal type into a 'theoretical practitioner (a'), and the second type into a 'practical theoretician' (b'), both moving nearer to our ideal-typical *survey statistician* (SS). An SS is someone capable of determining survey design and procedures with an awareness of their consequences for the quality of the data in all its diverse aspects, and of the relationship of these aspects to survey objectives and costs. A most important quality of the SS is the ability to *work out good*

compromises and to *seek out and evaluate approximations* - hence to appreciate what is and is not important in survey design, what is critical and what is dispensable.

3 Quality assurance of survey data as the framework for training

As a consequence of the training objectives sketched above, a *data quality perspective* becomes the focus of the training. In fact, *the quality issue serves as the framework*: it determines the type of issues and materials to be included in the course, and the whole approach to training. A quality-oriented approach has the added value of combining theoretical and practical aspects; it forces integration and balance between theory and practice.

A comprehensive assessment of data quality requires its diverse dimension to be taken into account. Various organization have developed their own specific lists of the "quality dimensions", but they all have a great deal in common. Essentially, they all share the view that "improving statistical quality" means "increasing the utility of statistical products and services for the community of their users". Commonly, the concept of data quality includes aspects such as:

(i) Relevance - capacity of data to meet users' needs;
(ii) Accuracy - control of the magnitude of errors of measurement and errors of estimation;
(iii) Timeliness - assessing how fresh are the data and their adherence with users needs;
(iv) Comparability - to which extent data collected can be compared in multi-country surveys or between similar surveys;
(v) Accessibility and Clarity - these condition accessibility, form of dissemination, technical documentation, information services provided;
(vi) Coherence - with other statistics and over time;
(vii) Completeness - covering all the main variable connected with the subject.

The content of the training workshop has to be organised around these dimensions, and in discussing any aspect of survey design or procedures, each of these dimensions has to be considered. The approach to quality issues in training has, therefore, to be comprehensive, wide-ranging and integrated, without ignoring its complexity. In organizing the course around this framework, we must answer at least two basic points:

- How does the quality approach, and each of its dimensions, influence the choice of the training method (academic versus laboratory lessons; the whole class seminars versus small working groups; theory versus practice)?
- What kind of course content and materials has to be provided to cover the various quality dimensions involved in survey design?

Here are some examples to clarify the point.

(i) question wording and cognitive laboratory - among other dimensions, it concerns *relevance* (are the questions and response categories effectively targeting the user needs?), *accuracy* (which question forms are more reliable and less influenced by measurement errors?), *completeness* (is all the information produced in the required detail?).

(ii) weighting or imputing procedures: *accuracy* (is it necessary to weight/impute the data? what is the impact on mean-square error of extreme weights or of imputed values?), *comparability* (do different procedures affect comparability, and what common standards must be followed?), *timeliness* (is there a proper balance between the accuracy of more complex procedures, and the additional time they require?).

There is a special 'tension' in survey statistics training concerning a balanced mixture of technical (mathematical and statistical) and practical (substantive and operational) issues. "The challenge facing those of us who try to teach sampling or devise training programs in sampling, is to combine an understanding of the subject matter objectives and problems with an adequate level of technical expertise to appreciate the implications of design decision" (O'Muirchertaigh, 2005).

Starting from the total quality point of view means seeking a balance between theory and practice, showing on the one hand common elements in apparently diverse situations, and on the other, consequences in term of final data quality of specific design choices.

Unfortunately, for many (most?) teachers of survey statistics, the paradigmatic reference for evaluating survey design is simply the sampling variance - a very important aspect of course, but not the only one. Instead, in our vision *total quality assurance* of sampling data has to become a guiding principle in training. Surely, theoretical presentation of sampling theory has to be an occasion for raising the wide range of issues and consequences that every choice we make during survey design and implementation impacts on data quality.

Consequently, the training involves a continuous interaction between theoretical and pratical aspects. After any theoretical presentation, time has to be left for discussion on the implication of the choices for the design, not only from a sampling point of view but also from that of the whole survey operation. A constant reminder is needed of the fact that the results of a survey are not to be measured in terms of only of one target, such sampling variance or interviewer variance, but in terms of data quality as a whole. This means that all the related issues (sampling, interview training, interviewing, supervision and control, respondent burden, weighting procedures, etc.) have to be carefully considered in defining the best design and operational solution for the survey. Theoretical presentation of sampling theory has to become the occasion for feeding up curiosity of scholar and for stimulating their criticism on evaluating the wide range of consequences that every choice could have on the whole data quality.

4. Reference to a survey (group of surveys) as an organizing tool

We find that real surveys provide the most effective instrument for organising the training. *They can serve as an integrating tools for aiding the comprehension of survey design principles, as a spring-board to bring into focus all the complexity and interdependency of the diverse aspects of survey taking.* Big and complex, well-established and documented surveys are probably the best for this purpose, especially if the participants already have some familiarity with them (e.g., an important survey from their own country or organisation). What does the survey involve in all its aspects – objectives, design, procedures, implementation, evaluation, and use? How do the choices concerning its various aspects impinge on the quality (also in all its dimensions) and the cost of the output? What is the rationale – both theoretical and practical – behind the choices? And what is their consequence? What alternatives could have been chosen? How do they compare? How can the design be adapted if some parameters (objectives, cost and time constraints, etc.) change? The following for instance could be a reasonable sequence of issues to be considered in organization of the discussion.

a) Describe the reference context and the objectives of the survey.
b) Understand its complexity, considering all the aspects and their relationships.
c) Analyse and criticize specific aspects of the design and implementation.
d) Suggest and develop alternatives, with the given survey conditions and objectives.

e) Discuss adaptation or replication in changed environment (country, statistical system, organization, technological equipments), and changed subject-matter (poverty, living conditions, household expenditure, employment etc.).

A good example of it could be the Labour Force Survey, usually characterized by a complex sampling design, generally with panel components, both household and individual estimates, weighting procedures to be implemented for both design-determined (e.g. unequal sample selection probabilities) and survey - determined (e.g. non-response) adjustment, just to name some peculiar aspects.

Of course, a proper balance between considering the theoretical and the practical aspects is far from automatically achieved. Just discussing simplified, abstract aspects simply because they are mathematically elegant or tractable, without attention and reference to what is actually important for data quality, is not very useful. But nor is the recourse simply to 'practical rules of the thumb'. One should not, for example, propose logit models for non-response treatment just because they are commonly used. Instead we have to present such models as one possible solution (also considered in contrast with other models or other different techniques), and discuss issues such as (i) the formal prerequisites of non-linear model specification and their constraints and limitations; (ii) the problems arising in response probability estimate when an excessive number of regressors has been used (i.e. negative/non-robust response probability estimate); (iii) how to design surveys tools to maximize information on non-respondents; (iv) the substantive meaning of the specific variables inserted into the model; and (v) how much does this all help in improving accuracy of the estimates. The challenge here is to find a good and flexible compromise between sampling theory and field work, looking at the real survey choices to find examples useful to understand both sides (theoretical and practical) of the problems arising when you have to build up a well balanced survey design.

5 Heterogeneity as a challenge and also as an opportunity

As noted in the Introduction, heterogeneity is the essential condition of the type of training courses we are discussing, which represents a challenge and an opportunity at the same time. The challenge is obvious enough: it must cater both for working statisticians and fresh graduates. Those who know some theory have to be made more aware of the practical *meaning and significance* of that theory; those who are engaged in survey practice need to be made more aware of the *theoretical basis* of that practice. It is necessary to capture the interest and attention of both these groups, but in the same unified workshop. Actually, often the situation is even more complex. We are considering courses have to be international in scope, in particular with trainees coming from developing or transition countries. There are at least three types of individuals involved among them, distinguished by background, and often also by age. There are the younger, more recent graduates, typically well-versed of statistical and computer skills, and eager to learn of new developments in sampling and survey theory. (We have heard complaints from this category of persons about the *low level* of training they received even in some 'famous' developed country training centres). The bulk consists of middle ranking, usually somewhat older, official statisticians with a considerable practical experience but limited theoretical background; they fit well into the ideal-type we described in Section 2. The third, and perhaps the most difficult (but fortunately, usually quite small) group are some more senior statistical officers who have the opportunity to travel abroad to participate in the training workshop more by virtue of their position than of competence or interest in statistical matters. Fourthly, it can be efficient and convenient to mix with these groups at least a sprinkling of the training providing host's own fresh graduates – good survey statisticians need to be produced there as well.

As to heterogeneity being an opportunity, the central concept to be followed in organizing such training is related to the idea of contamination, or to use a more positive expression, of *cross-fertilization*: how to

best utilize incoming heterogeneity in background and skills to move towards the target objective of the ideal "SS". A very good counterexample of non-contamination can be seen from the following quotation from the famous statistician G. E. P Box, reported by Kish (1978) writing on some paradoxes of statistical disciplines. " One had the curious situation where the highest objective of the teacher of statistics was to produce a student who would be another teacher of statistics. It was thus possible for successive generations of teachers to be produced with no practical knowledge of the subject whatever". By the way, this admonition seems to be well cut-out for describing some very contemporary controversies between *sampling theory* and *survey sampling teaching* in the Italian statistical academy.

By contrast, a good example of cross-fertilization is provided by implementing work groups on specific subjects during the workshop. The type of questions useful for that purpose have to met at least these two criteria: (i) questions that imply competence both in theoretical and on practical aspects; (ii) questions that do not require too much time for writing and/or for calculation, but are more centred on the basic foundation of survey statistics concepts. Here is an example: each group has to discuss pros and cons from both theoretical and practical points of view and present a brief report on the subject to be discussed by the whole class.

Of course achieving cross-fertilisation requires one to find good criteria for working group organization. Good results can be reached by organizing heterogeneous groups with mixed ability and experience. To be effective, however, such groups cannot be made too heterogeneous. They have therefore to be supplemented by a system of rotational membership for different topics so as to maximise cross-fertilisation.

Role-playing is another useful tool. It allows participants to become better aware of different stake-holders in a survey process: data users, survey designers, fieldwork organisers, interviewers, etc.

6 Some additional practical aspects

In conclusion we note a few other points of practical relevance in organising training in survey statistics.

- Firstly, concerning the role of computing and numerical exercises during training workshops. We have noted with dismay how intensive introduction of computer work (e.g., in sample design and selection) is sometimes allowed to distract from what is really the objective of such training: *to enhance the participants' understanding about survey design principles and practice*. Too much time is spent on teaching or refreshing general computer skills such as SAS programming, and actually performing intensive computer-based tasks such as selecting large scale samples. Many of these numerical exercises turn out to be mechanical tasks involving application of a specific design, and not a reflection on implication of (and alternatives to) that design. Numerical exercises are more useful when used for analysing consequences of different designs.
- Despite the development of the web, we still consider it very important to prepare a comprehensive selection of written materials (notes on existing papers and articles, selected reading text that serve as reference tool, etc.).
- The question-and-answer technique is an effective tool for bringing out important and relevant issues in survey design and practice. In fact, one of the challenges we like to throw at the trainees is for them to *ask* good questions. In fact this, rather than answering questions, is a more pertinent tool for evaluating the participants. We learned from Leslie Kish that it is more important (and difficult) to ask good questions than to provide an answer.

- Means should be developed to retain contact and exchange with, and also between, past participants – such as through a blog or newsletter, or a system of mentorship.
- What kind of teachers or trainers we need? Firstly people with mixed skills, combining theory and practical experience; then their availability and willingness to be mentors after the training is over; and also their ability to work in small groups.
- Finally, here is our view on how many regular (not necessarily full-time, of course) instructors are needed for a small group of 15-25 trainees: around 2*√(days of duration of the workshop). Such a core of the trainers usually need to be supplemented by experts on special topics.

This relevant, but of course not definitive, list of practical aspect to be taken into account in training organization, shows clearly the complexity of the educational process we have tried to describe in the these few pages. As Kish (1996) suggested at the end of his fundamental paper on training course for survey samplers, "... we have come a long way [*in training courses for survey statisticians*] in one generation, but we have left a lot more work for the next one". We hope that our efforts could represent an adding step in the never ended work of organizing better tailored courses in survey statistics for enhancing survey statistician peculiar skills.

REFERENCES

Kish, L. (1978) Chance, statistics and statisticians, *JASA*, **73**, 361, 1-6.

Kish, L. (1996) Developing samplers for developing countries: *ISR*, **64**, 2, 143-156.

Morganstein, D. and D. Marker (2000), A Conversation with Joseph Waksberg, *Statistical Science*, **15**, 3, 299-312.

O'Muirchertaigh, C. (2005) Balancing statistical theory, sampling concepts and practicality in the teaching of survey sampling, *ISI Bulletin*, **55**.

Verma, V. (1996) Discussion of Developing samplers for developing countries by Leslie Kish: *ISR*, **64**, 2 156-162.

Verma, V. (2001), Leslie Kish: Development of Statistics Internationally, *Proceedings of Statistics Canada Symposium 2001*.