International Statistical Review (1999), 67, 1, 49–62, Printed in Mexico

Social Change: Measurement and Theory*

Paolo Garonna and Umberto Triacca

Istituto Nazionale di Statistica, Via Cesare Balbo, 16, I-00164 Rome, Italy

Summary

Societal change, which takes a variety of directions and forms and in no way can be assimilated or reduced to a single dimension, is often accompanied by a perception of insufficient understanding and lack of control. There is a frustrated need for mastering complexity and instability, separating the voluntary from the involuntary, the intended from the unintended, opportunities from risks, getting to the real causes and dominating the uncertain implications of social change. Social change catches us unprepared and confused.

In this context statistics are generally considered a fundamental instrument of knowledge, but also part of the problem! In the public debate and in the specialized literature, the ability to measure social phenomena through current statistics and indicators is increasingly questioned. Data—it is claimed—are lacking, particularly longitudinal data; their quality (accuracy, relevance, timeliness, comparability, etc.) should be improved; indicators do not provide early warning signals, policy performance evaluation, and a precise indication of outcomes. Statistics cannot be used as a reliable and timely basis for decision making by individuals, organizations, governments, and for understanding these decisions. In some cases, statistics have been accused of giving a misleading and false picture of reality: do we measure the real extent of social exclusion and unemployment? Do we fully capture the quality of life and the degradation of the environment? Mismeasurement has been deemed by some commentators as being responsible for the wrong focus in inflation and stabilization policies, science and technology, unemployment and poverty. The productivity paradox, the informal economy, failure to measure welfare and the quality of urban life are instances where statistics do not seem to provide complete and satisfactory answers to the demand for information and knowledge. Our paper illustrates how, quite independently of measurement techniques and data production processes, the inadequacy of the conceptual framework may explain mismeasurement in relation to complex (multidimensional) and dynamic social phenomena. It is then to social theories, explanations and interpretations that statisticians need to turn, in order to come to grips with the new challenges in social measurement.

We will develop this thesis looking at a few cases where measurement issues can be connected to both theoretical and empirical difficulties. The statistical gap which reveals itself in the mismeasurement or difficult measurement of social phenomena is closely interconnected with the social science gap. Only close collaboration between statisticians and social scientists can bring about continuous advancement in social science and quality improvement in social statistics.

Key words: Problems of measurement; Social change; Social measurement data.

1 Signals and Symptoms

In the history of western civilization, there have been long waves of public attention and debate shifting from public institutions to civil society, and vice-versa. In the post-war period, there was first a strong commitment to building a policy environment at the national and international level conducive

*Revised version of a paper presented to the 51st Session of the International Statistical Institute, held in Istanbul, 18-26 August 1997.

to stability and growth: Keynesian demand management, welfare programs, trade liberalization and the international monetary system were the essential ingredients accompanying the reconstruction of Europe and the 'glorious thirties' of employment and growth. In the seventies, societies took their revenge by drawing attention to microeconomic instability, energy conservation, changing patterns of consumption, production adjustment and individual responsibility. The focus then shifted back to institutional reform: privatization, welfare cuts, regional economic integration, urban and family policies. The scenario in the present decade has been strongly conditioned by the two "revolutions" taking place in the transition countries and in the developing world, and by the remarkable convergence of political and economic institutions around the world towards the market economy and pluralistic democracy. Societies have, as such, then come back to the stage front, as the main source of prosperity and competitiveness. Coping with cultural or civilization clashes (Huntington, 1993), investing in trust and social capital (Fukuyama, 1995) and promoting competence and learning (Ciampi, 1996; Crozier, 1995) is the main policy thrust after the demise of social engineering and ideological confrontation, for the development of healthy and dynamic national systems.

Above all, we have become aware of a thriving, turbulent and often painful process of social change underway. We perceive more strongly shocks and crises; but it is undeniable that changes in life-styles, improvement in education, wider circulation of information and ideas are also creating new opportunities. A symptomatic analysis of society presents a compelling view of profound, accelerating and controversial transformations. Family structures and life-styles have clearly evolved, affecting patterns of family formation, marriage and cohabitation, sex and friendship, fertility and mobility, divisions of tasks and power in the household, social networks, community work and leisure, etc. The degree of 'openness' of societies has increased, matching the corresponding rise in the openness of economic systems: exposure to immigration pressures, worldwide circulation of news ideas and fashions, tourism and travel, common forms of entertainment music art and culture are stimulated by growing education levels, telecommunication technologies and better and wider knowledge of languages and civilizations.

Values, those profound beliefs making up the fabric and the sense of a community's identity, are shifting: from entitlements and social commitment to individual responsibility and self-help, from need to merit, from regulation and public support to education and competition.

2 The crisis of intelligence

Societal change, which takes a variety of directions and forms and in no way can be assimilated or reduced to a single dimension, is often accompanied by a perception of insufficient understanding and lack of control. There is a frustrated need for mastering complexity and instability, separating the voluntary from the involuntary, the intended from the unintended, opportunities from risks, getting to the real causes and dominating the uncertain implications of social change. Social change catches us unprepared and confused. We suffer from an intelligence crisis (see Crozier, 1995). This crisis is deemed to be at the root of the erosion of the trust we placed in the working of our resources, of our democracies and our economic systems. Facing a shortage of intelligence, challenges become threats, sources of conflict and protectionism, sometimes even xenophobia and selfishness, undermining the foundations of trust upon which our communities are built and the functioning of the democratic process relies.

In this context statistics is generally considered a fundamental instrument of knowledge, but also part of the problem! In the public debate and in the specialized literature, the ability to measure social phenomena through current statistics and indicators is increasingly questioned. Data—it is claimed—are lacking, particularly longitudinal data; their quality (accuracy, relevance, timeliness, comparability, etc.) should be improved; indicators do not provide early warning signals, policy

performance evaluation, and a precise indication of outcomes. Statistics cannot be used as a reliable and timely basis for decision making by individuals, organizations and governments, and for understanding these decisions. In some cases, statistics has been accused of giving a misleading and false picture of reality: do we measure the real extent of social exclusion and unemployment? Do we fully capture the quality of life and the degradation of the environment? Mismeasurement has been deemed by some commentators as being responsible for the wrong focus in inflation and stabilization policies, science and technology, unemployment and poverty. The productivity paradox, the informal economy, failure to measure welfare and the quality of urban life are instances where statistics do not seem to provide complete and satisfactory answers to the demand for information and knowledge.

3 Is There a Gap in Social Statistics?

Undeniably considerable progress has been made at the national and international level for the development of social statistics: population censuses have become richer and more precise; administrative files (such as population registers or welfare data) are being exploited for statistical purposes; sample surveys have become widespread (e.g. multipurpose sample surveys); time budget and longitudinal surveys provide precious information; definitions and classifications are improved and standardized; data collection techniques (area sampling or CASIC) and methods for data correction, compilation, imputation, etc. have been upgraded; data presentation, analysis and dissemination are considerably more developed now.

The 'gap in social statistics', in spite of the considerable investment being made upon it, is mainly due to an unprecedented surge in expectations and demands; it is therefore the by-product of education, growth in civil society and in its consideration of the importance of statistics; this gap must therefore be addressed by a preliminary investigation of users' requirements, particularly the new features.

Five such features can be singled out as being the most relevant and intriguing, albeit not the only ones:

1. Performance assessment of social policies has become both more needed and more difficult. Social policy reform requires complex, informed, rational and forward-looking decision-making. Tight budgets push towards cost-efficiency and effectiveness. Alternative institutional arrangements are being experimented with and tested. 'Coherent' policy making (see OECD 1993), a cross cutting approach to indicators, accounting, and analysis, and the integration of different social data archives figure highly in the political agenda.
2. The increasing role in social protection of non-profit, and other private organizations, households and citizens, reliance on self-help, individual responsibility and choice determine a widespread demand for guidance, counseling and public information. Which school has better facilities and teaching? In which areas and sectors is one more likely to find a job? Which degrees give a higher probability of successful careers? How safe is a neighborhood? How do we choose a hospital or a specialized physician? Should one postpone having children for career reasons? etc. While individuals and business have, in general, easy access to basic information for their economic choices for consumption, investment, saving, etc. (e.g. short-term indicators, CPI, interest and exchange rates), data for social choices are less easily available.
3. People increasingly demand disaggregate data relating to their territory, community and environment. Information must be susceptible of being linked through geo-referencing with other information drawn from different sources.
4. Development of telecommunications, global technologies and world news stimulate demand for comparable cross-country indicators and analysis. This requires standardization of concepts, classifications and definitions at the national and international level.
5. Democracies today struggle for understanding the mood and analyzing in depth the malaise of

society, sending messages across different social classes, communicating facts and decisions, sharing views and aspirations. "Learning to listen" is the most critical, and often forgotten, art of governance and leadership (Crozier, 1995). Statistics are a fundamental instrument for giving voice to people's feelings and frustrations. People pretend to be listened to. This explains the rather high response rate to opinion polls, and their popular use in political or commercial campaigns, the growing importance of social surveys, the attention paid to questionnaire design, interviewing skills, and their impact on data quality. Disillusion with politics and disaffection with industrial relations can also be understood on the basis of the exit/voice dilemma (Hirschman 1977, Galbraith 1996): a more informed society can bring the citizens closer to, and make them more active in, the institutional arena (the policy).

If the challenges posed by the explosion of implicit and explicit demand for good quality social statistics are not met, statistics will be called upon to bear a part of the responsibility for social failures; it will even risk becoming the scapegoat of conflict and incomprehension. Blaming mismeasurement and the sluggish response of statisticians, however, undermines the public confidence in data and in their usefulness for decision-making thus further eroding the social fabric. The statistical gap and the social gap risk becoming mutually reinforcing in a vicious circle. Statisticians therefore need to engage in an in-depth analysis of measurement difficulties in social statistics, their nature, sources, implications and possible solutions.

4 Measurement and Theory in Social Science

Blalock (1968, p. 12) observed that "social theorists often use concepts that are formulated at rather high levels of abstraction. These are quite different from the variables that are the stock-in-trade of empirical sociologists. The problem of bridging the gap between theory and research is then seen as one of measurement error". Measurement in the social sciences may be conceived as a process linking abstract concepts to empirical indicators. It transforms concepts into accounting indicators or schemes. Lazarsfeld (1965) outlined this transformation, singling out four phases:

1. The abstract definition of the phenomenon or concept that is the object of the survey;
2. The breakdown of the original concept into "constituent aspects" or "dimensions". This phase is necessary because, as a rule, the original concept corresponds to a complex set of phenomena rather than to a simple, directly observable phenomenon;
3. The choice of indicators. In this phase an indicator is assigned to each dimension;
4. The construction of aggregate indicators. This phase may be skipped since the characteristics of the phenomenon that is the object of the survey may be such as to advise against constructing a synthetic indicator. In other cases, the aggregate indicator entails construction of an accounting scheme, as for instance a social accounting matrix, or accounts of employment and health.

If we analyze the measurement process just described, it becomes quite clear that developing it correctly necessitates a theoretical frame of reference. Theory is necessary at every level of exploration and in every phase of the process. In particular, theory plays a fundamental role in the abstract specification of the concept and in its breakdown into constituent aspects. Moreover, and again from a theoretical point of view, the intervening relations between the various constituent aspects, between the indicators chosen and the constituent aspects, and between the indicators themselves must be rendered explicit. For this reason, even phase three is pervaded by theory: judgment of the conceptual validity of an indicator (relevance), in other words its capacity to represent effectively a given dimension, cannot be separated from theoretical considerations.

On the other hand, the empirical aspects also play an extremely important role. To illustrate this point, let us imagine we have reached, through a measurement process of the type described above, construction of n indicators $I_1, I_2, \ldots, I_n$ (n variables). Considering the correlation among

these variables is worthwhile. In particular, a high correlation between two or more indicators can be interpreted as a symptom of an erroneous breakdown of the phenomenon and/or an erroneous choice of the indicators themselves. If, for example, the correlation between I_p and I_q is elevated, the corresponding dimensions A_p and A_q are probably not distinct constituent aspects; or the two indicators may measure the same dimension, A_p or A_q. In any case, empirical evidence of this type should spur additional theoretical efforts in order to redefine the measurement process and to improve the quality of the measurement or measurements.

The measurement of social phenomena is, therefore, a process involving aspects of both a theoretical and empirical character. This circularity between theories and data is inherent in the scientific method and characterizes the accumulation of knowledge (Garonna & Geretto, 1996). Data are needed to construct and validate theories; but at the same time theories are needed to construct and validate data. If in the process something goes wrong, as shown by the increasing difficulties of capturing social phenomena, we need to look at the whole measurement process and its different stages; in particular we must consider the interplay between conceptual and empirical aspects, between measurement and theory. Mismeasurement or measurement difficulties signal that something has gone wrong in this interplay: they call into question, therefore, not only techniques and methods of data collection compilation and processing, but above all the conceptual frameworks underlying the design and operation of statistical information systems. We claim that measurement gaps in social sciences are to be attributed to both theoretical and empirical shortcomings, and above all to the insufficient integration of the two dimensions. It can be shown that not taking into consideration the multidimensional nature of social phenomena in the conceptual framework may lead to mismeasurement.

Consider the following example, which follows the scheme of the measurement process developed above. A represents a given social phenomenon, assumed to be not directly observable. Previously we saw that the possibility of passing from an unobservable social phenomenon (or concept) to an indirect measurement of it entails breaking down the original phenomenon into its measurable constituent aspects (or dimensions). By *measurable dimension* we mean a constituent aspect to which an indicator can be assigned. In formal terms, this breakdown could take the form of a partition, $A_1, A_2, \ldots, A_n$, of the phenomenon A, in other words,

$$A_1 \cup A_2 \cup \cdots \cup A_n = A,$$
$$A_i \cap A_j = \emptyset, \quad i \neq j.$$

It is evident that multiple breakdowns are possible. But not all of them are useful in detecting the phenomenon A breakdown that identifies a single constituent aspect, which we will indicate as A_1, for example, is manifestly erroneous although convenient in practical terms. In this case, $A_1 = A$, but this contradicts our assumption that A is not observable. So, in this type of situation, although we can make use of an indicator that will give us a "perfect" measurement of the constituent aspect this indicator would "indicate" itself, rather than the initial phenomenon, in other words, A_1; and A_1 is only a part of A. So, if social change regards exclusively its complement this change goes undetected.

The logical contradiction pointed out previously can be overcome by breaking down the phenomenon into at least two measurable constituent aspects or dimensions: A_1 and A_2. However, such a breakdown might not be "fine" enough. In this case, once again, social change may not emerge, by virtue of a sort of "aggregation effect". One way to surmount this hurdle is to adopt a more sophisticated breakdown: $A_1, A_2, \ldots, A_n$, of the phenomenon considered by $n > 2$.

Up to this point we have conducted the discussion postulating a static context. However, the world is not static, so it is wiser to view the social phenomenon used as an example as a process evolving over time. We indicate this fact utilizing the following notation: A^t, $t = 1, 2, \ldots$. Introduction

of the dynamic aspect further complicates the picture. Suppose we have identified, at time t, an "optimum" or "ideal" breakdown: $A_1^t, A_2^t, \ldots, A_n^t$, which allows us to capture the changes affecting the phenomenon considered in their entirety. This is not to say that at time $t+k$, this breakdown will still be optimum. For example, at time t, $A_1^t \cup A_2^t \cup \cdots \cup A_n^t = A^t$, at time $t+k$, $A_1^t \cup A_2^t \cup \cdots \cup A_n^t \subset A^{t+k}$. In other words, social change will have rendered the optimum breakdown relative to time t obsolete, in the sense that this breakdown, corresponding to time $t+k$, no longer "covers" the phenomenon in its entirety. It is clear that, once again, social change cannot be completely detected.

Our argument illustrates how, quite independently of measurement techniques and data production processes, the inadequacy of the conceptual framework may explain mismeasurement in relation to complex (multidimensional) and dynamic social phenomena. It is then to social theories, explanations and interpretations that statisticians need to turn, in order to come to grips with the new challenges in social measurement.

We will develop this argument looking at a few cases where measurement issues can be connected to both theoretical and empirical difficulties.

5 Changing Family Arrangements

This is an area where changes in statistical tools and data have fed new theoretical perspectives, opening up issues for monitoring and understanding the reshaping of the family and its role in society. For a long period, official statistics had adopted a monodimensional approach focusing on family structures. This approach was basically induced by reliance upon census data, as the main source of information on households. But the family is a process; it goes through different stages and cycles; it develops links and patterns inside and outside; and the population census cannot capture the whole picture.

The standard indicators of household size, composition and age structure reflect basically what the census configuration can offer. This is an unfortunate but typical pattern in the relationship between theory and measurement: the analysis focuses only on those aspects for which data are available and indicators can be constructed; therefore measuring only what *can* be measured, rather than what *should* be measured. For instance, the number of households with no children increases; but we do not know whether this reflects patterns of fertility, age, or family planning.

There has been a long-lasting discussion on the lack of data concerning the family and its operation, and on the limitations of the statistical concepts underlying the production of demographic information. Several authors (see Barbagli, 1984) have discussed the many dimensions of the family. In particular three different aspects were singled out: 1) family structures; 2) the relationships within the family; 3) interfamily relations. The first is the most conventional dimension: it involves size and composition, patterns of family formation, transformation and breaking-up. The second and third aspects were the less well understood, concerning modalities of interaction, roles, division of responsibilities, frequency of contacts, support, ties, etc; requiring new statistical tools, sample surveys (ad-hoc or multipurpose), exploitation of registers (where available), extension and adjustment of census definitions.

In many countries demographers and statisticians have a well established tradition of working together. This partnership has prompted cross-fertilization and innovation in the field, affecting methods and measurement techniques (imputation, interviewing, but also analysis and dissemination).

Many issues remain outstanding however, requiring a greater degree of interaction between statisticians, demographers and other social scientists. For instance, the boundaries of households become increasingly mobile and permeable (Saraceno, 1994). Do children of divorced couples belong to a single family? Do weekend partners constitute one or two households? What about college-students or retired people living in different contexts for parts of the year? Cohabitation is still considered an essential element of the statistical definition of a household; but relations, networks and relationships

at the boundary among different households need to be further analyzed. In some countries, the development of informal networks of self-help, advice and assistance centered around an extended concept of family or kinship (see Sabbadini, 1996) is assuming growing importance, vis-à-vis the prevailing evolution of welfare reform.

Longitudinal patterns, sharing of roles between partners, people in institutional care (dependent elderly, the disabled, etc.), and other issues on the frontier of the transformation of the family are stimulating a new wave of statistical activities based on the integration of data from various sources (time-budget surveys, ad-hoc and multipurpose surveys, home production accounts, 'social economy' or voluntary labor accounts, etc.), more in-depth analysis and a multidisciplinary approach (e.g. the economics of the family).

Statistical institutes are venturing ever further in the explanation and interpretation of the changing role of the family, producing and disseminating analytical monographs and papers, cooperating and sometimes competing with research centers and universities.

6 Measurement Without Theory: The Case of the Labor Market

Social sciences suffer from a sort of paradigmatic poverty with respect to the other sciences. Often, we find before us a plurality of *quasi-paradigms* (or potential paradigms), none of which is able to emerge as truly dominant: able, in other words, to consolidate the results of concurrent paradigms within a more solid and comprehensive theory. Socio-economic paradigms have a tendency to become atrophied because they are isolated from the outside world—from the empirical evidence, the practice of measurement and, above all, from the stimuli the outside world produces. In our opinion, it is precisely this separation between the theory and practice of measurement which prejudices the possibility for a new paradigm to emerge and conquer the others. This makes it very difficult to formulate and specify a sufficiently robust theoretical scheme within which to align and organize the process of measuring social phenomena. The end result is the determination and legitimization of a situation of "measurement without theory".

Let us consider, for example, the problems of measurement regarding the labor market. In this field it is sufficiently clear that, in many respects, measurement difficulties have to do with inadequacies and delays of a theoretical-conceptual nature. Prevailing concepts and definitions are, in fact, taken from the traditional models of the labor market, based on the standard Keynesian approach. In these models, labor was assumed to be relatively homogeneous and fully flexible, while salaries were considered relatively rigid and, in any case, having little impact on unemployment (which is, therefore, *involuntary*). The economy was seen as being closed. Instability and uncertainties were essentially controlled by public infrastructures and interventions, and, on the whole, the social structure was deemed stable and traditional.

Today's labor market situation is completely different. Under the thrust of the changes taking place, economic analysis has for some time qualified, clarified and, in part, abandoned this reference model, creating new interpretative schemes in their place: efficiency wages, the theory of implicit contracts, insiders/outsiders, game theory applications, models of adverse selection and of the internal labor market, etc.

However, these models have not yet produced sufficiently solid conceptual frameworks to enjoy a consensus that would allow them to become operational on the statistical level. Attempts in this direction have been made in reference to the concept of "natural" unemployment, NAIRU (non accelerating inflation rate of unemployment), or the reservation wage. But these attempts still have not met with success because of the lack of clarity and consensus surrounding these indicators, which makes their popularization and diffusion problematic.

Greater success is enjoyed by the search for indicators complementary to the unemployment rate and the activity rate. Growing importance is being attributed to the disaggregated indicators of

unemployment by age, sex and duration, the distinction between job losers, job leavers and persons seeking their first job, the disaggregations by educational and skill levels, by territorial area, by household typology and position within the household. One group of indicators has linked the job search with the conditions of supply: minimum acceptable salary, geographic availability, the search for temporary or atypical jobs, the intensity of the job search (number and types of search actions). The labor force survey provides numerous relevant indicators in this respect, albeit sometimes with limited statistical significance. The sensitivity of the unemployment rate to variations in these supply conditions has proven to be significant and growing. This means that the labor market functions as such, and the circumstances of purely Keynesian (or involuntary) unemployment, depending, in other words, exclusively on effective demand, are limited.

The data could allow indicators which highlight structural types of relationships to be constructed; for example, persons willing to work on certain specific conditions depending on a shortage of information, on the house rental market, or search costs; or labor demand indicators conditioned by the rigidity of hiring/firing practices, by the difficulties of finding labor, by the mismatch between requested and available qualifications. Important discussions and formulations are underway to define "discouragement" indicators, involving persons who do not seek a job because they doubt they will find one, involuntary part-time labor, and offers of a second or an informal job.

There is a great request for data on duration and flows, which permit us to understand the persistence (hysteresis) of unemployment, inter-company and inter-sectorial movements and their implications for salaries and careers (the Lilien index, for example).

Labor statisticians have sought to separate out the inactive, i.e. those who are not part of the labor force, and the persons who are in the educational or professional training system or whose status alternates (cf. OECD, 1992). Operators in the so-called "third sector", those engaged in "socially useful", unpaid community or voluntary activities are hard to classify.

Registration of vacancies, with the indispensable accompaniment of information on duration contractual and salary conditions, has become particularly difficult, primarily because of the increasingly marginal role played by public placement offices and the lack of statistical surveys on the subject.

Given this variety of indicators, the problem becomes one of choosing the correct one or group of correct ones. At this point the lack of theoretical and conceptual frames of reference able to guide these choices and furnish assistance in the decision-making processes by public and private operators and individual citizens becomes clear. The prevailing trend has been to substitute the unemployment rate with a set of indicators, based on the consideration that a single index is unable to satisfy the need for information and to adequately represent a complex and multi-faceted situation. This policy, assumed by the pioneering American Bureau of Labor Statistics, has been followed by numerous statistical institutes and by Eurostat, although sometimes with a dose of abstract empiricism.

The true obstacle is the still wide divergence existing between analysts and labor scholars and official statisticians. The former often tend to neglect conceptual questions connected to statistical production, only to then lament the scarcity and imprecision of data. The latter feel sometimes they are able to act alone, taking charge of survey techniques and limiting themselves to a conventional harmonization of definitions and classifications.

Reassessing unemployment statistics may, in this context, represent an important occasion to bridge this gap and work together. Under pressure from the European Parliament, public opinion and the governments of the member-states, the European Union has initiated an ambitious program to work out harmonized indicators, to define a common structure on labor force surveys, and to develop the European panel survey. An equally ambitious program is underway to formulate conceptual and methodological definitions in preparation for the 1998 meeting of the Conference of Labor Statisticians. Without the active involvement of the scientific community, labor scholars, statisticians, economists, sociologists and institutional scholars, etc., we will be unable to close the gap between unemployment theory and measurement practice. This will create heavy risks for the

credibility of the indicators in this crucial field of economic policy.

We are convinced that considerations relative to measurement problems vis-à-vis the labor market can also be extended to other fields of social analysis. In particular, overcoming the gap between theory and measurement practice is a *sine qua non* for establishing a correct approach to any problem of measuring social change. Finally, we believe that theoretical paradigms must be open to stimuli and solicitations growing out of empirical evidence and measurement practice, interacting with them through a scheme which relies on systematic feedback. In other words, a flexible, dynamic approach is needed to offset a rigid, static paradigm that is destined to suddenly collapse the moment undetected signals of change become too strong and no longer refer back to the paradigm utilized.

7 Perceptions and Expectations: Measuring Subjectivity?

We have seen above that one of the phases of the logical-conceptual process that allows us to translate the abstract image of social phenomenon (that cannot be observed directly) into (observable) indicators, or into a synthesis of indicators, involves defining the constituent aspects or dimensions of the phenomenon considered. We have also stressed that it is possible for social change to generate a greater articulation of the social phenomenon that is the object of the survey, in terms of both dimensions and dynamics. In this regard we must emphasize that, in many cases, social change produces a growing importance of the subjective character inherent in the constituent aspects of the phenomenon under consideration. In these types of situations, if we continue to consider only the objective dimensions, the phenomenon may lose clarity. For this reason, more space must be given to the "expectations", "opinions", and "perceptions" which individuals and households have of reality. In other words, a greater use of so-called subjective-perceptive indicators is needed.

Let us consider, for example, religious phenomenon. Quite probably, the decline of the public function of the churches, the tendency to "privatize" religious feelings, the manifestation through minimal acts, their nesting in the intimate sphere of the subject's conscience have reduced the importance of the *behavioral* dimension of the phenomenon (a dimension typically of an objective nature) measurable through structural indicators, such as frequency of attending religious functions, participation in religious rites (baptisms, weddings) or commitment to religious organizations and activities, to the advantage of dimensions that are of a subjective nature, such as the *cognitive* dimension, which Glock (1965) termed "ideological", regarding the beliefs and representations of the subjects, or the *experiential* dimension, which concerns, as does the cognitive dimension, the subjective plane but, unlike the cognitive dimension, is based on feelings and emotions.

Another social phenomenon influenced by significant subjective transformations is health, and especially the concept of health. In this case, as well, the subjective dimensions have assumed growing importance. The problem of health is increasingly set in a context that spills over strict medical and clinical boundaries (which may be considered objective). An individual's health is a comprehensive consequence that grows out of the interaction of physiological and pathological phenomena, the capacity of the organism to deal with "attacks" which compromise overall "normalcy", not only in terms of physical efficiency, but also in terms of psychic and relational equilibrium. The same disease foreshadows a variety of situations determined by the different alterations and by the consequences these alterations have on the individual himself. Furthermore, the implications at the psychological and social level of the disease contribute heavily to the definition of the overall pathology. In essence, the medical aspect cannot be the sole dimension of the health phenomenon. Growing importance is being attributed to the subject as the pertinent and credible source to evaluate his own state of health. Once again, in this framework the perceptive approach assumes a key significance.

These examples confirm the utility of integrating the measurement of social phenomena, also by rendering explicit the subjective aspects characterizing them. Still, we must note that the construction of subjective indicators involves significant methodological problems connected to the difficulty

of expressing the data in statistical-mathematical and comparable forms that can be generalized. Feelings, attitudes and perceptions may vary from subject to subject, sometimes even in the same person depending on the circumstances, since the mode of expression of individuals depends on age, cultural level and socio-economic conditions. In part, these problems are overcome by constructing attitude scales (consisting of a series of statements or items on which individuals express their agreement or disagreement) which allow us to measure the intensity of the attitude by graduating the statistical units of the collectivity considered along a continuum. But the fact remains that subjective aspects are not easily measurable. Broadly speaking, the level at which subjective aspects can be measured is inferior compared to the level at which objective aspects can be measured. In fact, the majority of techniques formulated to measure attitudes up to this time have produced scales which, at most, can be considered ordinal.

In conclusion, we can affirm that the more social change augments the importance of the subjective dimensions, the more the level of accuracy in measuring social phenomena declines and, consequently, the more the possibilities of capturing their effects precisely are reduced. But the quality of the measurement is not reduced thereby. The relevance of the indicators actually grows, as does their capacity to capture the complexity of social change.

8 The Challenge of Measuring 'Outcomes' and Effectiveness

Social policy debate and reform has focused attention on the outcomes and the effectiveness of the various measures. Conventional input indicators and cost-efficiency considerations are not considered sufficient to evaluate performance and guide public choice in the difficult decision making that fiscal consolidation and demand pressures impose.

In issues concerning health, education or employability, the effects of a certain policy programs may occur along a period of time, after some j-curve effect; they may determine interactions and feedbacks after the first impact; they induce often dead weight and substitution; they may be accompanied counteracted or even neutralized by expectations and anticipations of users and producers; they may be indirect or diffuse to the extent of generating the well-known properties of externalities or public goods, which affect measurement.

Two aspects have, more recently, further complicated the picture: first, the public demands increasingly 'qualitative' assessments, meaning, rather loosely, more accuracy in the identification of the long-term and comprehensive effects of social expenditure or cuts thereof; second, the coherence—or rather incoherence—of the policy approach in separated specific fields is generating concern and remedial initiatives: How does education affect employment? Does crime have an impact on business development? To what extent are family formation and decisions affected by economic prospects and employment security? What are the implications of a forward-looking prevention policy in health, long-term care, the elderly.

In the literature, emphasis has been placed on the concept of "outcome", meaning the long-term impact of a certain measure on the state of welfare of the beneficiaries (Opit, 1993), rather than the more conventional concept of output. The effectiveness of a health program refers to the 'outcome' obtained in relation to the objective assumed. The outcome is an indicator, the most significant one, of the quality of health care and, as a consequence, the measurement of the 'product' of a measure or programs is part of the more general process of evaluation of the quality of health care (di Orio, 1988, p. 89). The analysis of the quality of public services therefore, particularly consumers' or personal services, must shift from the perspective of efficiency (resources-output relationship) to that of effectiveness (initial state-outcome relationship), (see Vittadini, 1977).

This shift, however, is more easily preached than made. How do we measure, for instance, the state of health of a specific population target group? The standard practice is to make use of conventional diagnostic parameters of a clinical or biologic type. These indicators, however, do not capture the

multidimensional nature of phenomena such as the state of health, and often do not allow a full understanding of structural change. It is well known for instance that a chronic state implies that clinical parameters do not change significantly; but this does not mean that the benefit of alleviating a chronic disease cannot be measured (see Levorato, Rozzini & Trabucchi, 1990). The scarce sensitivity of clinical parameters has stimulated the search for new health indicators, more accurately reflecting the health conditions of the elderly. In particular, functional measures have been developed, aimed at directly detecting the ability to lead daily activities and respond to stimuli and conditions (Williams, 1989, Guralnik, 1989). However when the health state does not depend only on physical efficiency, even these indicators may be inadequate.

Other relevant examples may be drawn from the field of education and training. In this case, the objectives of policy are often defined in terms of knowledge and competence or in terms of employability and labor market leverage, going beyond specific professional skills and operational abilities. Therefore, measuring the effectiveness of training or education requires a complex process of conceptual refinement and testing. A further complication is the need to consider the impact of the environment and other interviewing factors; isolating the incidence of a specific policy measure and controlling for side or unintended effects may be quite an impossible task.

Difficulties may become insurmountable if the conceptual work is carried out only at the empirical level or worse as a matter of conventional accords among nomenclaturists and statisticians. As a matter of fact, the distinction between outcome and output, which is proving expedient and stimulating in specific fields of statistical analysis, might lead to an excessive simplification of the concept of output, and endanger a separation of social policy analysis from the work carried out in the economics of services production and exchange. When, for instance, Vittadini states that "output is the standardized product of individual services, measurable at a given time in terms of quantity and quality (e.g. the number of blood tests, the number of students being examined, the efficiency of specific machinery in taking medical tests, etc.)", he seems to adopt a narrow and limited definition of output, and risks confusing the output of a given service with the production process of the service itself. In the national accounts literature (see Hills, 1977 in particular), a "service" is defined as a change in the state of a person or a good belonging to an economic unit, induced by the activity of another economic unit; therefore the output must be measured by identifying the extent of such a change, and not by looking at the activity of the provider. "The process of producing a service is the activity which affects the person or good belonging to some economic unit, whereas the output itself is the change in the condition of the person or good affected" (Hill, 1977, p. 318).

To a large extent output and outcome overlap and interact; they both identify a change of state and condition of the benefit-holder; they differ however in their drawing attention to a different time span for the assessment of the impact of a measure. Output is evaluated by comparing the initial state with the one immediately after the service was provided. "Services are consumed as they are produced in the sense that the change in the condition of the consumer unit must occur simultaneously with the production of the change by the producer: they are one and the same change" (Hill, 1977, p. 337). Therefore, output is measured at the time in which the activity ends. The outcome, instead, is the long-term results; it can only be determined after some time following the completion of providing the service.

But here is where theory must come to the rescue if progress is to be made. When and where do we collect the relevant information? And what is most relevant? Without a well-established theoretical foundation, the search for outcomes, and effectiveness, become unsustainable. When the longitudinal dimension is introduced into statistical information systems, and when those systems need to dialogue and integrate, only a solid theoretical perspective enables a viable and effective investment program in measurement and data collection.

Data inadequacies, particularly longitudinal data and lack of integration of survey and administrative files, are often pointed out as being a major cause of difficulty for social theory advancement

and the assessment of social policy performance. However without social policy research and investment in social analysis it would be practically impossible to fill the statistical gaps. Paraphrasing Thomas Jefferson we might say that we may measure a lot of aspects across different fields at a given time, or a limited number of focused (and theory-driven) aspects though time; but we cannot measure everything everywhere always! Social theory must guide a bold and realistic statistical policy reengineering in the social field.

9 Partnership Between Official Statistics and Social Science

Other examples can be given to illustrate how, and why, measurement difficulties are linked to theoretical inadequacies in social science; consider for instance the measurement of competence, skills, human capital and the outcome of investment in education and training; the measurement of cultural identities, ethnic or minority status, social classes or groups, and the degree of cohesion or integration in a community; the measurement of values, beliefs and confidence, and of the shifts in values; the measurement of poverty, social exclusion, the quality of life and the urban environment.

Many, if not most, of the widely discussed and controversial concepts concerning society, social policy and people's welfare are not directly nor easily observable; they require a revamped investment in measurement and theory.

Much has been said of the knowledge-based economy and its vast requirements of investment in science, technology and know-how. Much less diffused, however, is an awareness that the present underinvestment in measuring social phenomena is undermining the response capacity and self-confidence of societies themselves as they are confronted with the paramount challenges of the new world of global information and social change. This paper has intended to show that within a comprehensive effort to give social science its proper place and a leading role, a profound adjustment is also required in the strategies of national statistical agencies (NSIs) and their relationship with social research. First, NSIs are investing more in improving measurement tools and arrangements for censuses and social surveys, integrating survey and administrative data, applying new data capturing, processing and dissemination techniques. Second, NSIs are devoting more resources to the analysis of social questions, using data as input to test models (e.g. microsimulation) and conceptual frameworks, presenting them in a comprehensible form to the general public. NSIs have become the first users of their own data, and are strengthening their skills and role of social analysis. Third, this approach leads NSIs to develop much closer ties with the research community, social science centers and universities. These institutions are consulted for the definition of programs of statistical activities; they are asked for transmission of new results, frameworks and ideas which may be encapsulated in the design of information systems or in their restructuring; they are provided with much more data, particularly disaggregated data, for their empirical work. Partnership in this context is a necessary implication because only joint research programs allow the constraints imposed by the stringency of confidentiality rules to be overcome while safeguarding the concerns of respondents. Finally, NSIs are opening up their work on statistical concepts, definitions and classifications in order to involve social science specialists, draw on best available theories and keep up to date with innovation in social science research and achievements.

Statistics are therefore becoming active partners in the world of social research, promoting and supporting, using and questioning ongoing theoretical work. It is well known how much statistics owes to social science, using it widely in the data production process; it is less well understood how much social science owes to statistics. Statistics not only provides valuable empirical evidence against which theories can be tested, but also validates theoretical frameworks putting them to the test of the measurement process. Theories, in fact, are the main ingredients for the construction of the conceptual frameworks underlying the measurement of social phenomena. Their viability and effectiveness in coping with the dynamism and comprehensive nature of social change represents a

crucial test of their validity. Theories are validated by empirical data; but at the same time the quality of data reflect the quality of the underlying theoretical framework.

The statistical gap, therefore, which reveals itself in the mismeasurement or difficult measurement of social phenomena is closely interconnected with the social science gap. Only close collaboration between statisticians and social scientists can bring about continuous advancements in social science and quality improvements in social statistics.

References

Barbagli, M. (1984). *Sotto lo stesso tetto*, Bologna Il Mulino.

Blalock, H. M. (1968). The Measurement problem. In *Methodology in Social Research*, Eds. H.M. Blalock & A. B. Blalock pp. 5–27, New York: McGraw-Hill.

Ciampi, C. A. (1996). *Sfida alla disoccupazione*. Bari: Editori Laterza.

Crozier, M. (1995). *La crise de l'intelligence. Essai sur l'impuissance des élites à se réformer*. Paris: InterEditions.

di Orio, F. (1988). *Statistica medica*. Roma: La Nuova Italia Scientifica.

Fukuyama, F. (1995). *Trust. The Social Virtues and the Creation of Prosperity*. New York: A Free Press Paperbacks Book.

Galbraith, J. K. (1996). *The Good Society. The Human Agenda*. Boston: Houghton Mifflin Company.

Garonna, P. (1994). Statistics Facing the Concerns of a Changing Society. *Statistical Journal of the United Nations ECE*. **11**.

Garonna, P (1997). *Misure e teorie del mercato del lavoro: il ruolo della statistica ufficiale*. Convegno SIS '97, La Statistica per le Imprese, Torino, Italia, April 1997.

Garonna, P. & Geretto, P. (1996). *Official Statistics and Scientific Research: an Evolving Relationship*. Conference in Honour of the Bi-centennial of Stefano Franscini's Birth, Ascona, Switzerland, November 1996.

Glock, C. Y. (1965). Y a-t-il un réveil religieux aux Etats-Unis?. In *Méthodes de la sociologie*, Eds. R. Boudon & P. F. Lazarsfeld. Vol **I**. Paris-La Haye: Mouton.

Guralnik, J. M., Branch, L.G., Cummings, S. R. *et al.* (1989). Physical Performance Measures in Aging research. *Journal of Gerontology*, **44**, 141–146.

Hill, T. P. (1977). On Goods and Services. *The Review of Income and Wealth*, December.

Hill, T. P. (1997). *Tangible, Intangibles and Services: A New Taxonomy for the Clas-sification of Output*. CSLS Conference on Service Centre Productivity and the Productivity Paradox, Ottawa, Canada, April 1997.

Hirschman, A. O. (1977). *The Passions and the Interests: Political Arguments for Capitalism before its Triumph*. Princeton: Princeton University Press.

Huntington, S. P. (1993). The Clash of Civilizations? *Foreign Affairs*, **72**, 51–83.

Lazarsfeld, P. F. (1965). Le vocabulaire des sciences sociales. In *Méthodes de la sociologie*. Eds. R. Boudon & P.F. Lazarsfeld. Vol **I**. Paris-La Haye: Mouton.

Levorato, A., Rozzini, R. & Trabucchi, M. (1994). *I costi della vecchiaia*. Bologna: Il Mulino.

OECD (1992). *Employment Outlook*, Paris.

OECD (1993). *Social Policies Ministerial*, Paris.

Opit, L. J. (1993). The Measurement of Health Service Outcomes. *Oxford textbook of health care*, **10**, OLJ Londra.

Sabbadini, L. L. (1996). Trasformazioni familiari e rilevazioni statistiche. *Rassegna Italiana di Sociologia*, **XXXVII**, n.2.

Saraceno, C. (1994). *Un familismo ambivalente: le politiche della famiglia in Italia*. International Conference "Mutamenti della famiglia nei paesi occidentali", Bologna, Italia, October 1994.

Vittadini, G. (1997). *Una metodologia statistica per la performance evalutation dei servizi alla persona di pubblica utilità*, Convegno SIS '97, La Statistica per le Imprese, Torino, Italia, April 1997.

Williams, T. F. (1989). Current Status of Biomedical and Behavioral research in Aging. In *Health Care for an Aging Society*. Eds. S. Andreopulos & J.R. Hogness. New York: Churchill Livingstone.

Résumé

Le changement sociale, qui prends directions et formes diverses et qui ne peut en aucune façon être assimilé ou rédui à une seule dimension, est souvent accompagné par la perception d'une compréhension insuffisante et d'une manque de contrôle. Il y a un besoin frustré de maîtriser la complexité et l'instabilité, tout en séparant le volontaire de l'involontaire, l'intentionnel du non voulu, les opportunités des risques, pour arriver aux causes réelles et dominer les implications uncertaines du changement sociale. Le changement sociale nous prends au dépourvu et confus.

Dans ce contexte, la statistique est généralement considérée un instrument fondamental de la connaissance, mais aussi une partie même du problème! Dans les débats publics, ainsi que dans la litérature spécialisée, la capacité de mésurer les phénomènes au moyen de la statistique courante et les indicateurs vient de plus en plus mise en doute. Les données—on dit—manquent, en particulier les données longitudinelles; leur qualité (précision, pertinence, opportunité, comparabilité, etc.) doit être améliorée; les indicateurs ne fournissent pas des signaux d'alerte précoce, ni l'évaluation de l'accomplissement des politiques, ni une indication précise des résultats. Donc, on ne peut pas se servir de la statistique comme base croyable et opportune pour le processus décisionnel des individus, des organisations et de gouvernements, ni pour comprende ces décisions. En quelques ca, la statistique a été accusée de donner une image trompeuse et fausse de la réalité: est-ce que nous measurons l'extension réelle de l'exclusion sociale et du chômage? Est-ce que nous capturons entièrement la qualité de la vie

et la dégradation de l'environnement? La mauvaise mesure a été jugéè par quelques commentateurs comme la responsable de la fausse mise au point de l'inflation et des politiques de stabilisation, de la science et de la technologie, du chômage et de la pauvreté. Le paradoxe de la productivité, l'économie informelle, le défaut de mesurer le bien-être et la qualité de la vie urbaine, se sont des exemples où la statistique ne semble pas apporter de réponses complètes et satisfaisantes à la demande d'information et de connaissance. Notre thèse démontre que, d'une façon tout à fait independante des mesures techniques et des processus de production des données, l'état incomplet du cadre cenceptuel peut expliquer la mauvaise mesure relativement à phenomènes sociaux complexes (multidimensionnels) et dynamiques. C'est alors aux théories sociales, aux explications et aux interprétations, que les statisticiens doivent se diriger pour en venir aux prises avec les nouveaux défis de la mesure sociale.

Nous développerons cette thèse à l'aide de l'analyse de quelques cas où la question de la mesure peut se rapporter à des difficultés aussi bien théoretiques que empiriques. Le trou statistique, qui se fait danc connaître à travers la mauvaise mesure ou la mesure difficile des phenomènes sociaux, est étroitement lié au trou de la science sociale. Seulement la collaboration entre les statisticiens et les sanvants sociaux peut amener un progrès continu de la science sociale et une amélioration de la qualité de la statistique sociale.

[*Received September 1997, accepted July 1998*]