DANIEL COURGEAU

GENERAL INTRODUCTION

To give the subject of this volume—multilevel analysis—its proper place within the general context of the social sciences, it is useful to start with an overview of the issues and questions raised by these disciplines. Compared with the natural sciences, the social sciences are far from fully constituted, and it is important to realise how their relevance can be improved by taking multiple aggregation levels into account.

First, the social sciences begin with the observation of a human behaviour or phenomenon, and then seek to structure it into different fields, which will constitute the specific object of each science. As a rule, the objects are defined independently of the vantage-point and scale that we can adopt to observe them. For example, the object of economics is the production, distribution, and consumption of wealth, but there is no indication of whether the level chosen is the individual, a market, a firm, or a nation. Likewise, the object of demography is the quantitative study of human populations, their variations, and their condition, but demographers do not specify whether the study is at the level of a family, a local population, or a national population. In other words, the distinction between levels precedes the object of each of these sciences, and we shall see that they are all subjected to it.

Second, the social sciences need to discover the categories that will provide suitable starting points for their development. It may be tempting for any individual, who routinely experiences these various social facts, to remain content with their apparent meaning and with a naive explanation of the lived experience in its immediacy—either because the individual already realises its meaning, or because he or she feels its absence and is preparing to search for it among similar lived experiences (Granger, 1994). This is true of the many phenomena studied in the social sciences, such as the birth of a child in demography, a price rise in economics, the fact that a person develops AIDS in epidemiology, or the fact that a person solves a problem in psychology. Far from being convinced of the complexity and opaqueness of these phenomena, naive observers see them as being fraught with explanation, because of their possible experience of similar circumstances and because of their personal knowledge of the conditions and consequences of their everyday actions. But this explanation, specific to each individual, will differ from that of other individuals with other experiences. As a result, it will not provide any schematisation that can be adopted by all and be publicly intelligible. The social sciences must therefore set aside such explanations and identify conceptual categories that will allow them to objectivate human experience, even provisionally. Although these nascent sciences, by comparison with the “non-social” sciences, have not yet identified the categories with sufficient clarity, we can assume that the objectivation process is under way. We shall return to this point in the volume’s conclusion.

Third, we need to realise that all scientific knowledge assumes a splitting of reality into a concrete aspect and a virtual aspect—the latter being a fairly abstract image of the events explored by the scientist. For the social sciences, the observation of individuals, groups or societies provides the "reality" that they will study. But unlike in the physical sciences, where there is only one type of virtuality—an abstract construct based on the formal properties of physical phenomena—virtuality in the social sciences undergoes a split as well. The first virtuality obtained is "an abstract construct, external by nature to the conscience of human actors albeit very *real*; it provides a scaffolding for the knowledge of facts" (Granger, 2001, p. 191): by placing this virtuality in the foreground, the social sciences attempt to *explain* human behaviours, developing the structure that will enable social scientists to describe the phenomena observed. The second virtuality resulting from the split is experienced by the actors: it is a complex of thoughts, affects, and intentions that make it possible to *understand* the behaviour of a given individual, without being entirely accessible to the social scientist. We are dealing here with a "clinical knowledge" that aims to grasp human facts in their singularity, in their individuality and not in their generality. However, it is this experienced virtuality that will confer meaning upon the abstract construct that the social scientist is trying to achieve. As we shall see, it is from life stories told by individuals themselves that biographical analysis will be able to take shape: the stories will provide the raw material from which we can construct a process underlying all these biographies.

We will not elaborate here on this "clinical knowledge" of singular human facts, which psychoanalysis and some currents of psychology and sociology are seeking to attain. The process of capturing the individual, in this case, "consists in constructing and superimposing ever more detailed networks of conceptual representation, each of which represents only a generic virtuality. It is the assumed convergence of the superposition of these grids that would bring us closer to an understanding of the individual. But it would, at best, impart only a limit meaning to this scientific reality of the individual. When measured against its concrete aspect, the limit meaning remains indefinitely incomplete" (Granger, 2001, p. 206). We shall therefore give priority to a virtual theoretical representation—divorced from the procedures of an individualised and unique capture of phenomena—and a scientific description of human reality, using concepts chained together in causal relationships, which will lead to models. Among the sciences examined in closer detail in this volume are demography, economics, epidemiology, education, human geography and social statistics. These are in fact the main social sciences for which multilevel modelling proves to be the most relevant.

It is only when social scientists attempt to observe behaviours and objectivate the environment where they occur and the manner in which they operate that they will face the problem of aggregation—in a space that is both physical and social—and the problem of choosing a suitable time scale. There will be many questions as to the choice of the right aggregation level: Should the observation focus on aggregate or individual behaviours? Will the methods used to identify the relationships between the values measured be the same or totally different depending on the level of observation? Can several aggregation levels be used simultaneously? and so on. The social scientist will also need to address the issue of time: Will it be historical time, in which the events studied unfold, or, on the contrary, the time lived by the individual who experiences the events? Will the observation point be a precise moment of that "lived" time, in order to explain the behaviours occurring then by conditions prevailing immediately beforehand? Or, on the contrary, will the observation span an individual's entire life, involving constantly changing conditions? These issues have been present from the very beginning of social-science research. We shall address them throughout this volume, and try to find satisfactory solutions.

The multilevel approach—which has recently gained ground—tackles the issues from a fresh angle. Within the framework of a single model, it seeks to achieve a synthesis connecting individuals to the society in which they live. For this purpose, it uses intermediate levels, which can vary from one science to another: for example, class and school, in education; the village, the town, and the region, in human geography; the family, the household, and the contact circle, in demography. This approach recognises that the grouping of individuals according to these various levels introduces an influence of the group on its members and, conversely, an influence of members on the group's future. Ignoring this relationship may lead to an incorrect analysis of individual behaviours and an equally incorrect analysis of the behaviours of the entire group. Only by recognising these reciprocal influences can we arrive at a more correct analysis of behaviours. The aim of this volume is therefore to explore the contributions of the new approach to various social sciences, to dissect the methodological assumptions on which it is based, and to see if it helps to improve the state of knowledge in those sciences.

This multi-author volume is not simply a collection of independently-written papers. It is the product of close communication between the specialists involved in clarifying the advantages of multilevel analysis: their exchanges shed new light on the approach. We also asked a philosopher of science for a more epistemological contribution—which proved highly relevant—to our methodological work. Our joint efforts lasted more than two years and culminated in a three-day meeting at INED in March 2001. The forum gave us the opportunity to compare our different approaches—often in an impassioned spirit—and to agree on a fuller and more written-up version of our proceedings. We were also able to present the social sciences in a more varied manner so as to avoid excessive repetitions. Of course each author retains sole responsibility for his or her presentation; their opinions, which diverged on certain points, have been included here so as to highlight the constructive side of our disagreements.

Before opening the discussion to our contributors, we will try to outline a very broad opposition between holism and individualism in the social sciences, and to indicate the richness and relevance of the explanations they allow, as well as the apparent incompatibility of their premises. Next, we show how more numerous and more complex levels of aggregation can emerge. We conclude with a prelude to the synthesis provided by multilevel analysis, on which the subsequent chapters will elaborate, and give a short presentation of the scope of this volume.

1. OPPOSITION BETWEEN HOLISM AND INDIVIDUALISM

The distinction between holism and individualism stems from the fact that a social system can be viewed from two opposite perspectives: either as a totality endowed with specific properties, irreducible to those of its members, or as a set of individuals, such that all social phenomena resolve into individual decisions and actions, without involving any supra-individual factors.

In the social sciences, the initial opposition is between society and the individual—although, as we demonstrate later on, the levels are far more diverse. The important point here is to examine in greater detail, on the basis of this initial distinction, how behaviours are taken into account and what consequences result from it.

Society

It seems preferable to begin with the social structure or form, which is already viewed as essential in some of Aristotle's writings. For the philosopher, the State as community (πόλιζ), under whatever government, "is by nature clearly prior to the family and to the individual, since the whole is of necessity prior to the part; for example, if the whole body be destroyed, there will be no foot or hand, except in an equivocal sense, as we might speak of a stone hand; for when destroyed the hand will be no better than that" (*Politics*, book I, part 2, trans. B. Jowett) Considered as a whole, the community is not an artificial or conventional form, but originates in the demands of human nature: a man who cannot belong to a community "must be either a beast or a god".

In fact, for Aristotle, the individual cannot be the object of any science. He clearly states: "But none of the arts theorise about individual cases. Medicine, for instance, does not theorise about what will help to cure Socrates or Callias, but only about what will help to cure any or all of a given class of patients: this alone is business: individual cases are so infinitely various that no systematic knowledge of them is possible" (*Rhetoric*, book I, part 2, trans. W. Rhys Roberts). It should be noted here that Aristotle often uses the term "art" (τέχνη) as a substitute for the term "science" (επιστημη), although he occasionally distinguishes between the two: art is more oriented toward "necessity or pleasure"; science is disinterested and aims not to indulge in the pleasures or necessities of life, but rather to discover the structure of things. Incidentally, the modern concept of social science is not present in Aristotle's thought (Granger, 1976).

Closer to us, the macrolevel *par excellence* is society or the State, rather than the community. To take a society as the macrolevel is to regard it as a perfectly defined and organised whole, clearly distinct from the sum of individuals who compose it, and displaying a powerful internal integration. We can thus deal with this society independently of other contemporaneous societies, and we can treat the social phenomena to be studied as external to individuals. Moreover, these phenomena are of a different nature than individual states of conscience. By contrast, we can compare different societies and highlight their distinguishing features.

We have seen earlier that the purpose of all social sciences is to explain a certain number of behaviours and to analyse the structures in which these phenomena appear. The behaviours and structures are specific to each science, for example: mortality, fertility, nuptiality, and migration, for demography; production, and consumption of wealth, for economics; the dissemination in space and time of diseases, for epidemiology. When we view phenomena at the level of a society, the concrete aspect is represented by the statistical reality of the facts observed in that society. We can classify the facts into two categories, which provide an explanatory framework: (1) the facts that will represent the origin of social facts and the initial conditions observed; (2) the facts that will represent the results obtained under these conditions. The aim here is to use a *model*—which will constitute an abstract virtuality—to describe not only the summary results, but also the processes that lead to these results from the initial conditions.

The origin of social facts must be sought in the formation of the social environment in which they occur. The initial conditions will therefore be supplied by the main characteristics of this environment, which can lead to the phenomena studied and are observed prior to them. The conditions can be measured by statistics describing the state of the society under

examination at a given moment. The events studied, meanwhile, can be measured by the proportions of individuals who have experienced them in the following period, which is often very short. For example, the percentages of individuals having displayed a given behaviour (proportion of suicides; proportions of migrants, of persons who have contracted a particular disease, of farmers who have given up farming, etc.), will be linked to certain characteristics that may or may not lead to these behaviours (shares of Catholics and Protestants to explain suicide; percentage of managers or, on the contrary, of farmers to characterise migration; percentages of individuals living in insalubrious conditions or on the contrary in uncontaminated locations to characterise the propensity to contract a given disease; percentages of farm labourers or, on the contrary, of farmers on large holdings to characterise exits from farming).

Thus, when we start from society as an organised whole in order to produce a set of effects under social constraints, our aim will be to show the way in which the society produces a given economic, demographic, social, or other type of fact. More specifically, it is by linking the observed facts to the society of which they are a diverse expression that we will be able to explain and find a basis for their reciprocal effects (Franck, 1994).

Durkheim (1897/1930) sought to relate social facts to the society in which they occur, in order to explain and find a basis for the effect of the religious system, household system, and political system on suicide. To his end, he established a network of links between different factors representing these systems—factors for which we can perform suitable aggregate measurements (such as the percentages of Protestants or divorcees in each age group). His method for comparing suicides in different categories of individuals is that of concomitant variations already advocated by Mill (1843), which closely foreshadows what we now call a linear regression. Durkheim shows how, for different sub-populations, suicide varies as a function of religion, and of the domestic and political characteristics of the society in which the individual lives. However, it is not these characteristics themselves that explain the greater or lesser frequency of suicide, but the social structure itself in which the individuals live. He concludes that "suicide varies in inverse proportion to the degree of integration of the social groups to which the individual belongs" (Durkheim, 1897/1930, p. 223), i.e., the more structured the society, the fewer suicides will occur in it.

Likewise, demographers have long given priority to the analysis of aggregate data. This is possible, for example, by using civil-registration records to study a phenomenon in the year following a census. The phenomenon can then be related to the set of characteristics measured thanks to the census. First, the census provides data on various populations exposed to the risk of the events; one can thus calculate the corresponding rates for different regions, districts, or population categories covered by the civil records (by occupational category, for example). Again, we can also use linear-regression methods for more detailed analyses: Puig (1981) examines the immigration and emigration rates of French regions, measured by a question on the place of residence in the previous census; he relates the rates to several aggregate characteristics of these regions (percentage of farmers, unemployed, etc.). Puig effectively identifies a link between ratios, while assuming that it proxies the influence of these characteristics on an individual's decision to migrate "based on a trade-off between his or her resources and location preferences" (p. 49). We will examine the validity of such a hypothesis later.

Likewise, in epidemiology, the theory of miasmas developed in the first half of the nineteenth century can be regarded as a holistic approach to public-health issues associated

with urban conditions, poverty, and hazardous occupations (McMichael, 1999). For this purpose, epidemiologists relied on statistical averages characteristic of a natural and physical environment. The result was a series of measures to improve sewage systems, water supply, regulations on housing standards, and—more generally—sanitation developments affecting public health.

In all these instances, society is regarded as a system composed of different categories such as religious, occupational, or political. The system explains why an observed social fact is the cause of a given social effect, or why it produces another given social fact. The problem is to define and bound this system properly by identifying the appropriate aggregate characteristics, which correspond to the collective states existing in the society (suicide rates, percentages of Protestants and Catholics, proportion of bachelors and widows, etc.). We will then be able to consolidate the relationship between these characteristics, such as the respective influence of Catholicism and Protestantism on suicide rates.

Another factor underlying this approach is historical time: as noted earlier, we will observe the situation at a given instant to explain the phenomena that occur at that time on the basis of conditions prevailing immediately before. The approach gives precedence to the analysis of concomitant phenomena and relationships observed at that moment: period analysis in demography, static analysis in sociology, structuralism in anthropology, etc. Of course a change from one period to the next is possible, as the structures have changed and the macro effects can also evolve. Again, however, these changes take place only at the aggregate level, without involving individual behaviours that occur in a "lived time".

The paradigms or rather the research programs that sustain such an approach in each social science must all, therefore, regard the individual as a non-relevant unit, and consider that only the individual's membership in different groups or categories will influence the occurrence rates of the phenomena studied. Of course, these paradigms will contain other elements specific to each social science. This defines a *methodological holism* in which some of the facts studied are a function of the social science examined, while others may be common to several of the sciences.

Indeed, it is a holism of this kind that enables us to envisage an interdependence of social facts in a given social structure. Accordingly, a particular social fact that we want to study may appear in different proportions in different regions studied, as can the preceding social facts to which we are trying to connect it. By contrast, to the extent that this structure characterises the society as a whole, a linear relationship between these proportions should emerge. In demography, for example, if we establish a relationship between regional emigration rates and percentages of farmers, the relationship should always involve the same parameters for all regions (Courgeau, 2001). We will thus be able to estimate the migration probabilities of the farmers and of other categories, that are independent of the region in which they live, by regressing the regional emigration rates on the regional percentages of farmers.

If, however, these hypotheses are not confirmed, we cannot take the conclusion for granted. All we can say from such an analysis is that a high unemployment rate leads to a high emigration rate, which may involve persons in employment, unemployed persons, or persons outside the labour force. This type of fallacious inference leads to what is customarily referred to as *ecological fallacy*, which consists in trying to detect individual behaviours by looking at aggregate measures (Robinson, 1950). Robinson showed, for example, that the

correlations between two characteristics measured on a binary basis among individuals (being black and illiterate in the United States), or by proportions in regions (proportion of black and illiterate population), were generally not identical and could even carry opposite signs.

The individual

The other approach centres, instead, on the individual. However, given the diversity of meanings that the social sciences have assigned to individualism (Birnbaum and Leca, 1986), it is important to state at the outset that we shall set aside sociological, economic, legal, ethical, and philosophical individualism—described and discussed in greater detail in Valade (2001)—and focus exclusively on *methodological individualism*. This consists in explaining an observed phenomenon not as if it were determined by the society examined, but on the contrary as the outcome of individual actions or attitudes. It is essential, for example, to "reconstruct the motives of individuals concerned by the phenomenon in question, and to understand the phenomenon as the result of the aggregation of individual behaviours dictated by these motives" (Boudon, 1988, p. 31), all the more so as the individualism is rational here. Such an approach can be used for all phenomena, whether they belong to the realm of sociology, demography, economics, or any other social science.

It is important to realise that methodological individualism appeared in our western societies far later than holism, as it was largely derived from the ideas developed in the early Classical age, when "the autonomous individual constituted the ultimate unit of the social sciences, and all social phenomena were resolved into individual decisions and actions whose analysis in terms of supra-individual factors would be useless or impossible" (Valade, 2001, p. 370). However, its introduction raised a host of problems, which we now need to examine in detail.

Earlier, we noted the force of Aristotle's argument that "individual cases are so infinitely various that no systematic knowledge of them is possible". The individual is indeed intimately linked to the actors' virtual experience, comprising thoughts, affects, intentions, etc. not directly accessible to social scientists. Incidentally, this is why we do not elaborate here on the "clinical knowledge" of human facts. How can we, in such conditions, envisage the formalisation of a virtual individual as a theoretical object open to comprehensive modelling?

Our starting point is the observation of individual lives, by means of a biographical compendium that supplies all the events of use to the social science concerned, and provides an accurately dated record of the individual's existence. This observation in no way enables us to estimate individual random processes, whose probabilistic structure would be specific to each individual tracked. It does seem hard to assume that two individuals, even if similar in many ways, automatically follow the same path. Moreover, as we can only observe one realisation of this process for each individual—his or her own actual life course—we have no way of identifying its probabilistic structure. This is entirely consistent with Aristotle's earlier-quoted observation that an individual process is not identifiable.

We must therefore modify the interpretation of this process. For this, we shall distinguish between two stages in the development of a truly individual approach. As the full, complex process cannot be the object of scientific inquiry, the social scientist first needs to specify a *theoretical model*, characterised by only a small number of events. This theoretical model can

be regarded as "a filter that retains from complex phenomena only that which must figure in the object of research" (Franck, 2001, p. 288).

Conventional economics, for example, relies on the postulate that agents display a strong rationality, which can be defined by explicit axioms. This makes it possible to elaborate a pure theory, derived from the consequences of the set axioms (Walliser, 2001). However, the strong-rationality postulate can be rejected in favour of a limited-rationality postulate, which leads to different theoretical models.

Similarly, the demographer can set up a theoretical model that constitutes a complex causal mechanism. To study the mortality of a population, for example, we can assume that "the age at death of an individual would be dependent on the states he or she has undergone in his or her life, on the time spent in each, and on their sequence" (Duchêne and Wunsch, 1991, p. 112). These states—characterised by the individual's education, family status, occupation, etc.—are assumed to act in synergy on the effect. Here, therefore, we examine not the effects of isolated characteristics but the incidence of the transitions between states in an individual life.

In a second step, we will seek to test such a theoretical model with the aid of a special *empirical model*, situated at an intermediate level between the complex process and theory. Let us again begin with observed reality, which comprises a number of individual paths. From this observation, can we estimate a probabilistic process that takes into account all the information contained in these paths? To the extent that any random process can be seen as a distribution of probabilities across a set of paths, we can say—in this case—that we observe the same random process repeatedly. Now, the probabilistic structure of the underlying process becomes identifiable from the observation of these different paths. We thus identify a collective process, which can be as complex as we like.

In the search for individual random processes, two *individuals observed* by the survey, possessing identical characteristics, have no reason to follow the same process. By contrast, in the search for a process underlying the population, two *statistical individuals*—seen as units of a repeated random draw, subject to the same selection conditions and exhibiting the same characteristics—automatically obey the same process. We can thus see more clearly how the use of observed biographies, which constitute the statistical reality of the human facts studied, can now be transformed into an abstract description of human reality by means of concepts deliberately stripped of at least some of the concrete circumstances of a virtual life-experience. These concepts fall into a sequence governed by the logical relationships of the process identified, forming a biographical model.

From the observation of a set of individual cases, we can use such an analysis to describe a mechanism that will link the phenomena studied to the individual characteristics, whether or not they are time-dependent. We now need to show what abstract relationships exist between the elements of a process that organises the life of the population studied. But for this enterprise we must replace the aggregate approach by an individual approach to human societies. This calls for new data-collection procedures and analytical methods. Let us now examine briefly how this is taking place in selected social sciences.

The economist Léon Walras (1874/1926) held that individuals respond independently to market prices, which form the only link between them. A collective entity—the auctioneer—matches consumers' orders (which define their effective demands) against the producers'

supply, and can thus determine pure-competition prices: to reach equilibrium prices, the price of the goods whose effective demand exceeds effective supply must rise, and the price of the goods whose effective supply exceeds effective demand must fall. This approach is, however, based on very heroic assumptions that define a market operating under perfect competition, and hence an underlying economic structure: perfect market fluidity; immediate, full information for consumers and producers; free entry into the market, etc. Moreover, there is no statistical or temporal dimension involved here. Walras himself realised, however, the arbitrary nature of some his assumptions, when he stated that the calculation must stop at a certain point. This point, in his view, was marked by the appearance of free will, i.e., the emergence of a virtual life-experience no longer directly accessible to the investigator, as we noted earlier. Naturally the conditions of this model have been enlarged to the case where all individuals consciously interact with each other (Kirman, 1997). This enlargement, however, is still thwarted by the virtual life-experience or free will, even if its boundaries have receded.

Similarly, it is only in the early 1980s that demographers were able to adopt an approach effectively based on individual data supplied by the World Fertility Survey (WFS), INED's "Family, occupational, and migration biography" survey (also known as "Triple biography" or "3B"), and similar programs. The aim was no longer to take a snapshot of a population at a given moment, as in a census, but on the contrary to try to document over time either the history of the fertility of each woman in a sample, as in the WFS, or the more complex history of several types of phenomena tracked simultaneously in several areas of an individual's life, as in the 3B survey. But at the same time the methods for analysing aggregate data—essentially the linear regression methods—no longer allowed such biographies to be analysed. This obstacle was overcome by the introduction of new methods, initiated a few years earlier by probabilistic analysts (Cox, 1972); their development went hand in hand with their adoption in many fields, particularly the social sciences. Thanks to these methods of event history analysis, we can now examine the unfolding over time of several demographic phenomena while taking into account their possible reciprocal influences and, simultaneously, the role of various individual characteristics—time-dependent or not—on these behaviours (Courgeau and Lelièvre, 1986).

Meanwhile, by the late 1940s, medicine had succeeded in controlling the most important infectious diseases such as tuberculosis, smallpox, plague, and typhoid. Epidemiology was then confronted by new chronic diseases whose rapid spread and totally unknown origins were sufficiently frightening to replace earlier epidemics (M. and E. Susser, 1996). The chief weapon to combat these diseases was found in the vast cohort studies on smoker's cancer, coronary heart diseases, and so on. The studies accumulated a large set of characteristics of the persons untouched by the disease, of those already affected, and of those who contracted it in the course of the longitudinal study. The methods for observing and analysing such biographical data flourished, made it possible to identify the main risk factors, and helped reduce them. All these studies, performed in the last fifty years, seek to link the occurrence of a disease to various risks, using multivariate methods and survival analysis. This has made it possible to highlight the multicausal nature of public-health problems more effectively.

In all the examples we have given, the initial observation no longer involves society as a whole, but focuses on a number of individual cases from which we can estimate the structure of a given process. It is this process that now characterises the occurrence of one or more given phenomena. We will thus be able to describe the causal structure of the processes studied, such as the effect of a particular characteristic on the risks of disease in epidemiology, of migration in demography or human geography, of passing an exam in

education, and so on. More generally, we should be able to describe with precision the interactions between many phenomena and the characteristics of the society studied.

This time, however, we will not find a historical time underlying the approach, as in the case of holism, but a time experienced by the individuals studied, which may even be frozen, as in the Walrasian model mentioned earlier or in a stationary Markov process. The individual time scale may run from the date of birth or a founding event such as marriage (for the analysis of legitimate fertility) or contamination (for the study of the outbreak of a disease). The event history approach allows the events that will shape the phenomena studied to be introduced as time-dependent characteristics. We will therefore be able to track evolutionary processes: for example, "the evolutionary processes of the determinants of agents under the influence of time, their past actions, and the actions of others" (Walliser and Prou, 1988, p. 201), in economics, or a process describing the sequence of fertility-period events in demography.

The paradigms or rather the research programs that underpin such an approach in the social sciences will now consider individual biographies as the empirical material on which they will work; however, their research objects will be the processes that give meaning to the biographies. It is important to remember here that, under the terms of this methodological individualism, we cannot grasp the virtual life experience of these actors themselves, but only an abstract process. Each of these paradigms, of course, will contain elements specific to each of the sciences in question.

In fact, it is a methodological individualism of this kind that enables us to contemplate an interdependence between different characteristics for a given statistical individual, situated in a given social structure and a given population. Consequently, if we break down the total population into different parts, the statistical individuals in each of these sub-populations will no longer have any reason to follow the previous overall process. They will now follow as many different processes as there are sub-populations. For example, if we establish a relationship between the probability of migration of farmers compared with other occupations in the total population, the relationships estimated for each region will no longer yield the same parameters as under the conditions of methodological holism (Courgeau, 2001). Under the latter, statistical individuals taken from the overall population will consistently follow the same process, whatever their region of residence. By contrast, we can say that statistical individuals picked from different regional sub-populations will no longer have any reason to follow the same process. Yet there will not be the slightest contradiction between these two samplings, as we will not be dealing with the same statistical individuals. In this case, the concomitant-variations method is not valid; we need methods using individual responses to estimate the parameters of each regional model.

It is fair to say that this approach enables us to factor in the influence of some individual characteristics on behaviours: the characteristics form a combination of variables whose identified effects on a given cohort and in a specific environment provide an explanation of these behaviours. Here, however, as we no longer consider the context in which these behaviours occur, a risk of *atomistic fallacy* appears. Indeed, there is no reason why this context—comprising both family and environment—should have no influence on observed behaviours, and it seems fallacious to examine individuals divorced from the constraints of the society and environment in which they live.

2. HOW ARE THE TWO APPROACHES LINKED?

We are now in the presence of two approaches: the aggregate or "macro" approach versus the individual or "micro" approach. On the face of it, they appear to be antagonistic in every way, both epistemologically and technically—i.e., in the measurement, analysis, and interpretation of results. Let us begin by examining each of these points in greater detail.

Insofar as the aggregate approach starts with the assumption that the essential factor is the social structure, the latter cannot be defined except through the fullest possible observation of society. Moreover, the structure basically evolves in historical time; it can remain static for long periods yet change quickly in times of crisis. We will therefore give precedence to a synchronous measurement method, with repetitions in the course of the historical time of the observation. This makes it clearer why the principal method for measuring aggregate data is the population census, which gathers a comprehensive set of information on a country's population. The information can, for example, be linked to civil-registration data for demographic studies. Other sources—comprehensive or not—are available for this population, allowing a measurement of phenomena not documented by the censuses. Often, the sources consist of period surveys of large numbers of individuals, which capture a number of phenomena and furnish evidence of the links between them. For example, the Labour-Force Surveys conducted in European Union countries provide an annual overview of a population's activity measured through such variables as employment and unemployment, wages, occupation, educational attainment, and job search.

As the individual approach emphasises individual actions and decisions, these cannot be identified except by the most detailed possible observation of behaviours provided by a biographical database. Moreover, the time scale to be considered will be the individual's personal time frame, not historical time. We can thus see more clearly why the measurement of individual data, because of the scope of the information that it involves, leads to retrospective or prospective biographical surveys that can cover only a small number of respondents, given the level of detail of the questions required. The surveys supply all the events in an individual's life that are relevant to the planned study; the events are situated in diachronic time. As a rule, they are highly detailed biographical surveys that will capture the dates of occurrence of events in such different areas as family life and working life. Sometimes the events will be located in a physical and social space, along with a certain number of the individuals' characteristics, which can also change over time.

As we can see, the two types of measurement are very different, but they will nevertheless capture two aspects of an identical reality, specific to each social science. They are most often used separately, as the measures rest on different concepts such as holism or methodological individualism. However, if we could transcend these concepts, we could envisage a simultaneous treatment of both measures. We shall discuss this later. For now, let us turn to the analytical methods.

For aggregate data, the analysis of social structures leads to the identification of concomitant variations between relevant indicators: "the mere parallelism of the values taken by the two phenomena, provided that it has been established in an adequate number of sufficiently varied cases, is the proof of a relationship between them" (Durkheim, 1895/1937, p. 129). In this case, to show the possible links between an aggregate characteristic that we want to analyse and other aggregate characteristics that may help explain the phenomenon studied, the best method is to use regressions—in particular, linear regressions. The units of

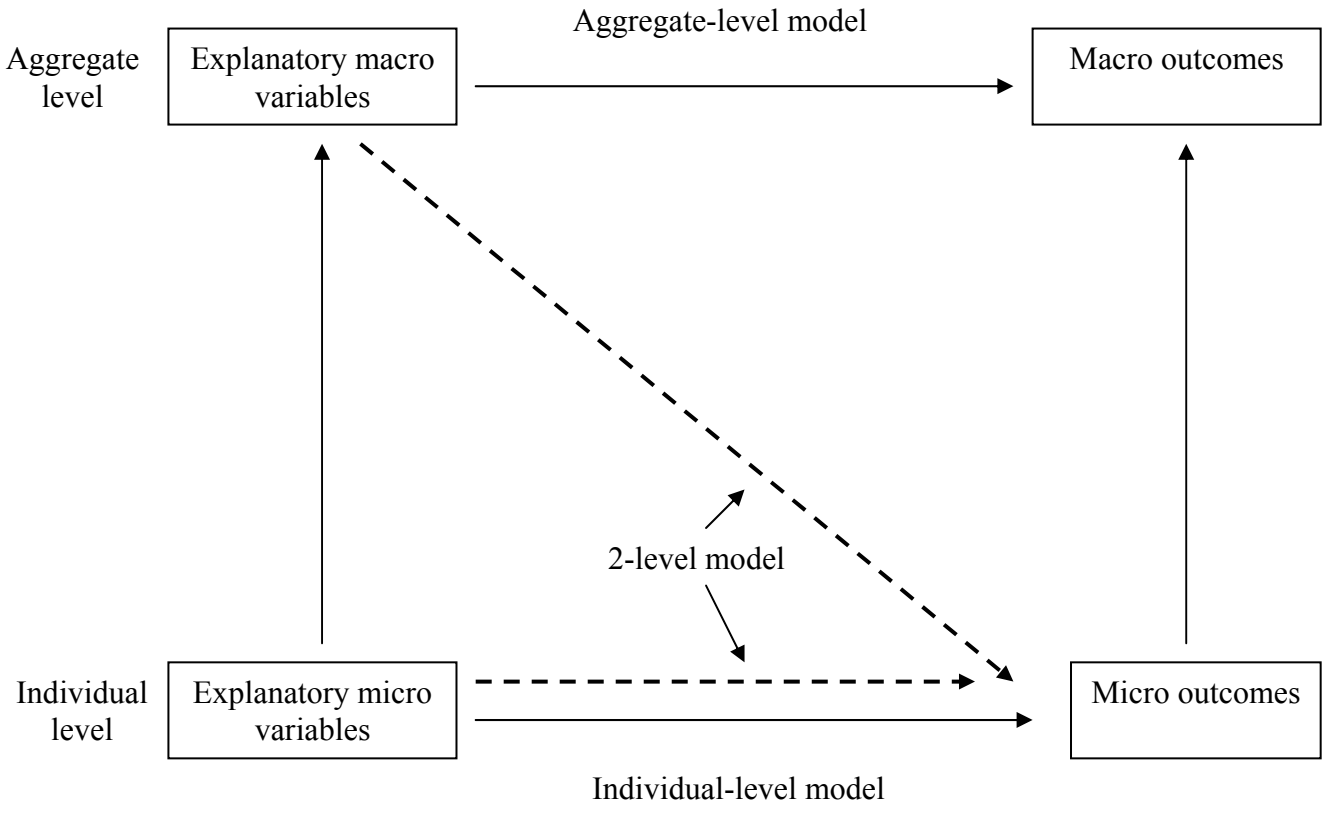
analysis generally consist of groups of individuals living in the same district, the same region, etc., especially when we are working on comprehensive census data. Within the group, we do not know the joint distribution of the individual variables. All we know is the marginal distribution of each of these characteristics. We will be able to identify the links between them by means of regression analysis on the groups—which, as explained earlier, we regard as arbitrary for our purposes. Such an analysis will also be conducted in a historical time frame, typically at a particular point, in order to interlink the social facts.

For individual data, on the contrary, the aim is to connect a life-long behaviour to personal events or characteristics that can also vary over time. Other methods must therefore be used. Here are some of the most common: logistic regression methods for binary data or polytomous data observed at a specific moment (binary: the individual has experienced a given event or not; polytomous: the variable examined may exist in several forms); event history analysis methods for binary characteristics observed throughout individuals' life courses (the date of an event they have experienced); or longitudinal analysis methods for continuous data such as an individual's size throughout his or her life. However, linear-regression methods can still be used for continuous individual data (such as an individual's score on a test). In all these cases we can find links between the individual characteristics by working on a "lived time".

Again, we can see how different these analytical methods are, and how hard it is to use them simultaneously. The first relate aggregate characteristics to each other; the second interlink individual characteristics. The first will place them in a historical-time context, the second in a time lived by the individual. We must therefore try to transcend these incompatibilities in order to establish models that will reconcile the analytical methods.

Let us set aside, for the moment, the time dimension of the phenomena studied, which we will examine in greater detail later. Figure 1 illustrates the two types of model and gives a clearer picture of the difficulties we face in trying to move from one to the other.

Figure 1: Connection between different models



The aggregate-level model, represented by the top horizontal solid-line arrow, shows the result obtained at the macro level, for example the proportions of individuals having experienced the event studied: the proportions are explained by the rule and institutions governing these behaviours in the society considered, which are measured, for example, by aggregate characteristics. In contrast, the individual model, represented by the lower horizontal solid-line arrow, links the obtained result to the micro level—for example, the probability that a statistical individual will experience the event, a probability explained by the individual characteristics governing the behaviour.

Using simple, concrete examples, we shall now attempt to describe the difficulties encountered when trying to move between these two levels in other social sciences. Here, we deal with the shift from the individual level to the aggregate level, as indicated by the two vertical solid-line arrows of figure 1. We shall not examine all the possibilities in detail, but we will show, in a few selected cases, the reasons for divergence between individual and aggregate models.

First, let us assume that all the individual characteristics are continuous data, for example, in education, the score in a test taken by an eleven-year-old pupil as a function of the pupil's score at age eight (Goldstein, 1995). We may regard the children's school as the aggregate level at which we can compute the average scores of pupils aged eight and eleven. The aggregate variables are, again, continuous variables. We should also recall that these schools are used, at the aggregate level, solely to estimate overall parameters, which are assumed to be independent of their choice.

Can we assume that the relationship observed between individual scores is reproduced identically between aggregate scores? Supposing, for example, that a linear regression is correctly demonstrated between individual scores, can we estimate the parameters of a similar regression on the aggregate data? We can show mathematically that the parameters of a regression on aggregate data cannot, in most cases, be deduced from those of a regression on individual data, as they depend on other quantities linked to the grouping used (Duncan *et al*., 1961). For these quantities not to affect the shift from individual to aggregate regressions, very special conditions must prevail—which is rarely the case. For example, if the scores at age eight are identical in each class and if each class has the same number of pupils, then the parameters of a regression on individual data and on aggregate data are identical. These parameters, therefore, may broadly be described as independent: in some cases, the parameter may be positive at the individual level and negative at the aggregate level. Moreover, if the relationship between the scores is more complex than a linear relationship, then the relationship between the aggregate data will not even be of the same kind as the relationship between individual data. For example, when the score at age eleven is linked to the score at age eight by an exponential relationship, it is easy to see that the aggregate scores will not be linked by an exponential relationship. We must therefore recognise in all these situations, where a linear or non-linear regression applies at the individual level, that the relationship at the aggregate level will measure something totally different from the aggregation of individual behaviours.

Let us now assume that we are working on binary individual data, i.e., the individual possesses or does not possess a particular characteristic, as regards both the variable to be explained and the explanatory variables. For example, in demography, let us suppose that we are seeking to relate an inter-regional migration to the fact that the individual is a farmer or not (Courgeau, 2001). We can treat the individual's region of residence as the aggregate level. This enables us to compute percentages of migrants and farmers in each region. We will then see a change in the nature of the variables and hence in the models applicable at the individual and aggregate levels. There is a shift from a dichotomous model (logit or probit) to a regression model involving aggregate characteristics. How can we tie the parameters of these models together? For example, we have been able to show that if we estimate a logit model at the individual level, then the model at the aggregate level is a linear regression between the percentages, whose parameters are expressed in a fairly simple but non-linear manner as a function of the individual parameters (Courgeau, 2001). In the case of the Norwegian data used in that article, the parameters estimated at the individual and aggregate levels will be very different and even contradictory, for at the aggregate level the migration probability of farmers exceeds that of other occupations, whereas at the individual level it is lower. Again, these differences can be explained only by simultaneously including in the same individual model the fact of being a farmer and the percentage of farmers living in the region. We are no longer dealing with an aggregate model.

Lastly, let us assume we are working on an individual biographical model, which will be seeking to estimate the effect of various individual characteristics, most of them binary and potentially time-dependent, on the instantaneous probability of a given event. The instantaneous probability measures the limit, when the time interval considered tends toward zero, of the ratio between the probability of the event's occurrence and the length of that time interval t. In that case the instantaneous probability will no longer even have an equivalent at the aggregate level, unless we define an instantaneous probability estimated on each of the sub-populations of the areas. If we follow this approach, the characteristics to be examined must be individual characteristics, not the areas' aggregate characteristics. This is equivalent to working at the individual level, using the division into areas as well. As we can see, the result is that we take into account the aggregate and individual characteristics to explain an individual behaviour. Again, the model is not of the aggregate type.

As the three examples above have shown, we cannot treat an aggregate-level model as directly deducible from a model estimated at the individual level. The criteria for grouping individuals are too far from being strictly random, and the resulting models generally do not resemble individual models. In addition, the aggregate-level model may incorporate more general characteristics that cannot be linked to individual characteristics. For instance, the population density of a locality or region, or the number of hospital beds there, have no equivalent at the individual level. They can, however, be aggregated at higher levels: the number of hospital beds in a region is the sum of the number of beds in each locality of the region. Other collective characteristics are well defined at a given aggregation level, but cannot be aggregated at higher levels: the political colour of a municipality—defined, for example, by the party affiliation of its mayor—cannot be aggregated with that of neighbouring municipalities, which may cover a broad spectrum. It therefore does not exist at the individual or regional level.

Conversely, it is yet more difficult to determine the parameters of the individual model from those of the aggregate model, even assuming a very simple aggregate model. In particular, the fact that information is poorer at the aggregate level than at the individual level

will prevent the switch: the percentages of farmers and migrants in a given region tell us nothing about a migrant's occupation.

In consequence, we will be able to regard these two types of models—one for the social structure, the other for individual behaviours—as practically independent. Let us now introduce the time dimension, and revisit the problems encountered when trying to interlink these different approaches.

We can say that, at the macro level, we will obtain snapshots of the social structure, taken—for example—at the successive censuses examined. We must then interlink these snapshots taken in historical time to show the broad changes in the society studied during the period. For instance, the theory of demographic transition will seek to find links between the long-term changes in a country's fertility and mortality by relating them to other equally aggregate characteristics such as international migrations and the role of institutions. We may therefore describe this theory as a process that occurs at the societal level, and we can compare it with similar processes occurring in other societies. These societies may be either of the same kind (for example, developed countries) or very different in their histories (such as the developing countries as distinct from the developed countries).

At the micro level, by contrast, we are dealing with an individual process to be described within the "lived time" of the statistical individual. The period will begin with an event such as the individual's date of birth or date of entry into the labour force. It will describe the sequence of events as a function of their inter-relationships as well as of individual characteristics, which may or may not change over time: number of siblings, successive occupations of each, etc. We may also assess the effect of economic or political events on that generation's behaviour: for example, how did the Second World War influence the fertility behaviour of women born in 1920, or how did the Great Depression of the 1930s influence the occupations of workers born in 1910? Lastly, as we will possess the timelines of successive generations, we will be able to interlink their histories in order to show the changes in behaviours over time: for example, how have consensual union partnerships spread in the European countries from one generation to the next since the end of the Second World War, and which population categories have provided the momentum for their dissemination? This approach introduces another view of the evolution of a society, via a different process than the one used in the aggregate approach.

Could we then connect the two processes by introducing a dual historical temporality—or rather a social temporality and an individual temporality—that would allow a joint analysis of these two processes? This is no easy task, given the separate and even conflicting development of the two approaches, which diverge in all aspects, but their conjunction would represent a breakthrough in the social sciences. However, before outlining a scenario for such a recomposition, we must continue our examination of the levels that need to be considered.

Thus far, we have envisaged only two main levels from which to study human phenomena: society and the individual. A closer look at a society readily shows us that other intermediate levels exist between them, and that we need to position ourselves at those levels as well in order to better understand the society in which they operate and the individual behaviours they can generate. Should we not also study the time scales that these new levels will generate?

3. A PLURALITY OF AGGREGATION LEVELS AND A PLURALITY OF TIME SCALES

Progress in the social sciences has led to the use of multiple levels, creating the need to define their purposes and to examine temporalities that are multiple as well. This diversity of levels and time scales further complicates the shifts between them but, as we shall see later, it will free us from the contradictions—and the need to choose—between holism and methodological individualism.

Let us take an example from education. Between a social structure, corresponding to the national education system, and an individual level, corresponding to the personal education process, we will find intermediate levels, whose importance is far from negligible. The first such level involved is the class, to the extent that the teacher(s) and pupils in the class can play a major role in the success or failure of an individual pupil. Some teachers may focus on the class as a whole by trying, for example, to raise all pupils to the same level. Other teachers, instead, may neglect the weaker pupils to allow the best ones to over-achieve, thereby widening the gaps between them. Other levels as well may influence the pupil: the family (the child's place among his or her siblings, and parental choices concerning the child's further course of study), the school attended (in particular, whether the school is public or private), the neighbourhood, etc. We can thus identify several levels—ranked in a hierarchic order or not—that will exert diverse effects on children's behaviour.

Likewise, when demographers or population geographers study inter-regional migrations, instead of adopting an individual or national perspective on these migrations, they can use the region as intermediate level. Depending on the region, several regional characteristics can play a crucial and different role in these population flows. At the national level, we regarded average regional unemployment or wage rates as means of measuring a broader effect of unemployment or wages on inter-regional migration rates. Here, we will assess the specific effects of those same characteristics on the individual probabilities of emigrating from each region. The same is true if we examine migrations between municipalities, districts, etc., which will involve levels representing municipal divisions, district divisions, and so on. Other, non-geographic levels may also exert an influence: the individual's household may restrain migration if it includes many members with different activities in the region; by contrast, the household may trigger migration for individuals living alone and with few ties to the region.

It is easy to generalise this emergence of multiple levels in all the social sciences, and it is increasingly necessary to examine them simultaneously. Indeed, we must realise that these realities are not ontologically separate, and that it is essential to find out how the microstructures fit into the macrostructures and vice versa.

At the same time, the temporalities to be taken into account will multiply and diversify. It may not be enough to analyse the time linked to the individual's life, as intermediate time scales may play a major role. Likewise, historical time can be broken down into periods of different significance.

For instance, when demographers want to study a woman's fertility, they can analyse the woman's age and compute age-specific fertility rates. But the demographer may take the view that fertility is linked to the formation of the couple, in which case it is preferable to calculate the rates specific to the duration of the live-in partnership—a temporality that, no doubt, better reflects the woman's fertility. Again, however, for births above order one, it may be preferable to focus on the interval between births rather than the duration of the partnership,

and so on. The choice between these time scales is not easy, and it would be better to examine them all simultaneously.

Moving in the other direction, it seems harder to discern a plurality of time scales under a single historical temporality. However, it is useful to introduce different periods into the latter, for example to characterise the changes in a population's fertility. It may be worth considering durations computed from key dates in the population's history that may influence its fertility. Examples include wars or economic crises; major discoveries in the field studied; the advent of the pill for the history of fertility, or major changes in the social rules concerning fertility, such as the legal restrictions on abortion in Romania in 1966.

To better understand human phenomena, therefore, we need to examine a wide diversity of levels and a great variety of time scales. The terms "micro" and "macro" become totally relative, and a level regarded as micro in an analysis can become macro in another. For instance, while the class represents a more aggregate level than the pupil, it will serve as a micro level relative to the school. This relativity of levels with respect to each other is now clearly visible. More important, however, it is essential to realise how closely interlinked these levels are, and how they can no longer be treated separately. We can no longer qualify any one of the levels as more fundamental that the others; even less can we state that it is independent of the others. We must therefore turn to the study of the inter-relationship between levels. This inter-relationship is the focus of the recomposition described in our closing section.

4. TOWARDS A RECOMPOSITION AND A MULTILEVEL SYNTHESIS

After breaking down the object of the social sciences into its various levels and temporalities—which, as we have seen, can seem mutually contradictory—we must try to reconstruct an overall object that will allow a synthesis of the approaches identified here. This summary approach has been called for by many social scientists for years (Alexander *et al*., 1987; Huber, 1991; Janssen, 1998). In this section, we shall give a very broad outline of a multilevel synthesis for use in different social sciences. To this end, we shall apply some of the concepts that were introduced in our discussion and will enable us to carry out the recomposition.

To begin with, we believe the concept of statistical individual is crucial for describing a more general process affecting the entire population. The concept allows us to link together the analyses undertaken at different aggregation levels, by accurately separating the virtual experience of these players from the virtual construct of the social sciences. There are no obstacles left to asserting that statistical individuals can be exposed to the effects of their own characteristics as well as to the structural constraints of the social system in which they live (Giddens, 1984). Such constraints are not exercised independently of individual motives and reasons—as in Durkheim's holism—but in a manner that is both enabling and restrictive. We can thus avoid the two types of fallacy described earlier: the ecological fallacy and the atomistic fallacy. The risk of ecological fallacy is removed, as the aggregate constraints will measure a construct that differs from their equivalent at the individual level. They occur not as a substitute, but as a structural constraint that can affect the behaviour of an individual subjected to it. Simultaneously, the atomistic fallacy disappears when we make proper use of the context in which the individual lives.

More profoundly, by identifying a plurality of levels, we discard the dualist approach, which pits society against the individual. Under these conditions, "it no longer makes sense to

choose between *holism* and *atomism*, and, as regards the social sciences, between *holism* and *individualism*" (Franck, 1995, p. 79), for our aim now is to study how these different levels interconnect. At the same time, we want to find out how to fit a historical time scale and different individual time scales together into a single model, as indicated earlier. Multilevel analysis, therefore, effectively paves the way for a new approach in the social sciences.

The new methods to be introduced should make it possible to treat levels in a hierarchic order (pupils in classes, classes in schools, etc.) as well as more complex nestings (individuals classified by type of residential neighbourhood and type of workplace, which are in turn ranked into districts and regions). It should also be possible to generalise these contextual models—in which individual results obtained in the groups composing a given level are treated as independent—into true multilevel models where the result of an individual situated in a group can depend on the results obtained by the other members of the group. Lastly, these methods should make it possible to accommodate the multiple time scales, both historical and individual, in which the events occur.

It should be noted that the multilevel model itself, as we have outlined it, operates at the level of the statistical individual, but leaves room for the effect on the individual's behaviour of characteristics measured at different aggregation levels, as well as the interactions between individual and structural constraints. Such a model is shown in figure 1 as two dotted lines connecting the individual and aggregate characteristics to the expected result. Of course, we leave open the possibility of interactions between these characteristics—not only at a given aggregation level but especially between these levels—and the possibility of drawing on far more aggregation levels than the two depicted in the diagram.

We may legitimately ask, however, if this approach is the only solution to the problem of the recomposition of the object of the social sciences. Might it not be necessary to use different models to explain the changes in structural constraints? For example, in economics, how should we analyse the birth, functioning, and death of institutions such as the market or central planning (Lesourne, 1991)? What, if any, interactions exist between these institutions and individual behaviours: how individual actions may break down an institution? Another valid question is whether—in view of the multiplicity of time scales to be incorporated into the analysis—several types of models should not be used to grasp the full complexity of the phenomena studied. These are some of the issues raised by the introduction of multilevel models.

5. OUTLINE OF THIS VOLUME

After this general introduction, that develops the distinction between holism and individualism in the social sciences, and also discusses the link between the two approaches and the need for a multilevel analysis, a first chapter by Harvey Goldstein presents the main ingredients of such an analysis, taking examples from educational sciences. He introduces the basic notations used throughout the book and covers aspects such as the units and levels to be considered. He presents the basic linear multilevel model, and elaborates this with examples of cross classified and multiple membership models. While this chapter is concerned mainly with educational data, almost all the issues covered apply to other disciplinary areas. The chapter provides the main tools for undertaking a multilevel analysis. In particular it proposes new ways of simultaneously modelling responses at different aggregation levels, without setting a logical priority of one over another.

The second chapter by Daniel Courgeau presents the history of demographic thinking, as an example, from the methodologist's viewpoint. He shows how, among other things, the increasing availability of detailed data has changed one's focus from the aggregate period approach, to longitudinal and individual life histories, and finally to multilevel analysis. He stresses the more general role of time and space in the social sciences, giving a wider scope than a demographic one to this history. He shows how the multilevel approach allows one to solve some of the problems encountered when using other approaches, and argues for a renewed use of the Bayesian probabilistic viewpoint.

The third chapter by Ana Diez-Roux provides a stimulating discussion of the potentialities and limitations of multilevel analysis in public health and epidemiology. She discusses the challenges raised by the use of such an approach in epidemiology, in particular the danger of adopting the methods for their own sake. She also points out limitations of the present multilevel models which often ignore complex chains of causation and non-linear structural relationships, although the contemplation of these issues within a multilevel framework is currently of great interest. This cautious approach emphasises the need for using a model not as a complete picture of the real world but as an imperfect way of understanding some aspects of the world.

The fourth chapter by Mark Tranmer, David Steel, and Ed Fieldhouse focuses on the use of multilevel models in exploring small area population structures. They deal with the practical issues of mixing individual and spatial data from various sources, and present detailed examples of their approach to British data, including nested and cross-classified models. They point out the pitfalls and solutions when using geographic area data, a common approach in such fields as economy, demography, public health and human geography.

In the next chapter, Bernard Walliser presents a discussion of organisational levels and time scales in economics. This theoretical exposition gives a complete overview of the main paradigms. He shows that despite the modest role of holism in economics, this research field cannot be regarded as wholly informed by methodological individualism. He introduces useful distinctions between other methodological forms, such as weak methodological holism or individualism, autonomous or crossed methodological interactionism. He finally argues for empirical economic models to be confronted by theory.

The sixth chapter by Robert Franck deals with philosophy and epistemology, while it considers very practical issues raised by the previous chapters. First, he carries out an examination of several types of causal determination and of their impact on multilevel analysis. Then, he clarifies the nature of levels, by the aid of the philosophical concept of hylemorphism. This concept allows us to understand the emergence of systems, at a higher level, from factors acting reciprocally at the lower level. Under these conditions multilevel analysis may be enlarged to deal with the influences of factors over the emerging system, and to the influence of the emerging system over the same factors, leading to new procedures for the analysis of system functions.

The volume ends with a series of conclusions by Daniel Courgeau. He raises some more general issues on the non-experimental versus the experimental approach (Wunsch, 1994; Goldstein, 1998), on the nature of the concept of probability used (Greenland, 1998, 2000), and on the difficulty of defining relevant levels and their interconnection in the wider sense. In other words, how should we implement a fuller theory of human behaviour?

It is now up to the authors to show—for the social science area represented by each—the means of introducing such a multilevel analysis, of taking into account the various

aggregation levels and complex time scales as best as possible, and of addressing the issues left unsolved (or, on the contrary, the issues raised) by this approach.

REFERENCES

Alexander, M., Giesen, B., Münch, R., & Smelser, N. J. (Eds.). (1987). *The micro-macro link.* Berkeley, Los Angeles, London: University of California Press.

Aristotle (1885). *Politics* (B. Jowett, Trans.), from http://classics.mit.edu/Aristotle/politics.html

Aristotle (1954). *Rhetoric* (W. Rhys Roberts Trans.), from http://classics.mit.edu/Aristotle/rhetoric.html.

Berthelot, J.-M. (2001). Les sciences du social. In J.-M. Berthelot (Ed.), *Epistémologie des sciences sociales* (pp. 203-265). Paris, Presses Universitaires de France.

Berthelot, J.-M. (Ed.). (2001). *Epistémologie des sciences sociales.* Paris: Presses Universitaires de France.

Birnbaum, P.,& Leca, D. (Eds.). (1986). *Sur l'individualisme. Thèmes et méthodes*. Paris: Presses de la Fondation Nationale des Sciences Politiques.

Boudon, R. (1988). Individualisme ou holisme : un débat méthodologique fondamental. In H. Mendras & M. Verret (Eds.), *Les champs de la sociologie française* (pp. 31-45). Paris: Armand Colin.

Courgeau, D., & Lelièvre, E. (1986). Nuptialité et agriculture. *Population, 41*, 303-326.

Courgeau, D. (2001). Vers une analyse biographique multiniveau. In Christine M. (Ed.), *Actes des journées de méthodologie statistique* (pp. 375-394). Paris: INSEE méthodes 101.

Cox, D.R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, *B34*, 187-220.

Duncan, O.D., Cuzzort, R.P., & Duncan, B. (1961). *Statistical geography.* Illinois: The Free Press of Glencoe.

Duchêne, J., & Wunsch, G. (1991). Life histories, a new approach to the study of mortality. In Jozwiac, J., Kotowska, I.E. (Eds.), *Usefulness of demographic models* (pp. 111-120). Warsaw: Institute of Statistics and Demography.

Durkheim, E. (1930). *Le suicide.* Paris: PUF (original work published 1897).

Durkheim, E. (1937). *Les règles de la méthode sociologique.* Paris: PUF, (original work published 1895).

Franck, R. (1994). Les explications causale, fonctionnelle, systémique ou structurale, et dialectique sont-elles complémentaires ?. In R. Franck (Ed.), *Faut-il chercher aux causes une raison ? L'explication causale dans les sciences humaines* (pp. 275-302). Paris - Lyon: Librairie Philosophique J. Vrin.

Franck, R. (1995). Mosaïques, machines, organismes et sociétés. Examen métadisciplinaire du réductionnisme, *Revue Philosophique de Louvain*, *93*, 67-81.

Franck, R. (2002).Conclusion of Part IV. In R. Franck (Ed.), *The explanatory power of models. Bridging the gap between empirical and theoretical models in the social sciences* (pp.285-294). Boston/Dordrecht/London: Kluwer Academic Publishers.

Giddens, A. (1984). *The constitution of society.* Cambridge: Polity Press.

Goldstein, H. (1995). *Multilevel statistical models.* London: Edward Arnold.

Goldstein, H. (1998). *Model for reality: new approaches to the understanding of educational processes.* (Professorial Lecture). London: Institute of Education,.

Granger, G.-G. (1967). Epistémologie économique, in Piaget, J. (Ed.), *Logique et connaissance scientifique* (pp. 1019-1055). Paris: Editions Gallimard.

Granger, G.-G. (1976). *La théorie aristotélicienne de la science.* Paris: Aubier.

Granger, G.-G. (1988). *Essai d'une philosophie du style*, Paris: Editions Odile Jacob.

Granger, G.-G. (1994). *Formes, opérations, objet.* Paris: Librairie philosophique J. Vrin.

Granger, G.-G. (2001). *Sciences et réalités.* Paris: Éditions Odile Jacob,.

Greenland, S. (1998). Probability logic and probability induction. *Epidemiology, 9*, 322-332.

Greenland, S. (2000). Principles of multilevel modelling. *International Journal of Epidem*iology, *29*, 158-167.

Huber, J. (Ed.) (1991). *Macro-micro linkages in sociology.* London: Sage Publications.

Kirman, A. (1997). *Interaction and markets.* (DT-GREQAM, n° 97A02). Marseille, Groupement de Recherche en Economie.

Janssen, M. (1998). Microfoundations In Davis, J.B., Hands, D.W., and Mäki, U. (Eds.), *Handbook of Economic Methodology* (pp. 307-310). Cheltenham: Edward Elgar.

Lesourne, J. (1991). *Economie de l'ordre et du désordre.* Paris: Economica.

McMichael, A. J. (1999). Prisoners of the proximate : loosening the constraints of epidemiology in an age of change. *American Journal of Epidemiology*, *149*, 887-897.

Mill, J. S. (1843). *A system of logic, ratiocinative and inductive, being a connected view of the principles of evidence and the methods of scientific investigation.* London: Parker.

Puig, J.-P. (1981). La migration régionale de la population active. *Annales de L'Insee*, *44*, 41-79.

Robinson, W.-S. (1950). Ecological correlations and the behaviour of individuals. *American Sociological Review*, *15*, 351-357.

Susser, M., & Susser, E. (1996). Choosing a future for epidemiology: I Eras and paradigms. *American Journal of Public Health, 86* (5), 668-673.

Valade, B. (2001). De l'explication dans les sciences sociales: holisme et individualisme. In J.-M. Berthelot (Ed.), *Epistémologie des sciences sociales* (pp. 357-405). Paris: Presses Universitaires de France.

Walliser, B. (2001). La science économique. In J.-M. Berthelot (Ed.), *Epistémologie des sciences sociales* (pp. 117-147). Paris: Presses Universitaires de France.

Walliser, B., & Prou, C. (1988). *La science économique.* Paris: Editions du Seuil.

Walras, L. (1926). *Eléments d'économie politique pure ; ou Théorie de la richesse sociale.* Paris: R. Pichon et R. Durand-Auzias (original work published in 1874).

Wunsch, G. (1994). L'analyse causale en démographie. In R. Franck (Ed.), *Faut-il chercher aux causes une raison ? L'explication causale dans les sciences humaines* (pp. 24-40). Paris - Lyon: Librairie Philosophique J. Vrin.