Environment and Planning B: Planning and Design, 1992, volume 19, pages 613-637

Visualizing people in time and space

D Dorling
North East Regional Research Laboratory, Department of Geography, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, England
Received 8 April 1991; in revised form 6 July 1991

Abstract. This paper is an introduction to novel ways of rendering spatial-temporal systems to gain insight into and a holistic view of the more complex aspects of their distributions. Colour images can depict multivariate distributions for more than 100000 spatial units from a census. The changes between two or three censuses can also be displayed at this resolution. Other techniques can be employed to explore flows of people, for example, to work or to new homes or from other countries. Three-dimensional and animated displays can be used to portray the integrated geography of space and time. Examples including the distributions of employment, voting, and movement are given and can be combined. All illustrations are of national distributions using the finest resolutions of spatially and temporally defined data that are currently available.

In this paper I report on the earliest results of the visualization work being carried out in the North East Regional Research Laboratory. I therefore concentrate on the reasoning behind this research, where the initiative has come from, and where it is hoped it will proceed in the future. The paper is divided into two broad sections; in the first half I describe the theory that links this visualization methodology to conventional spatial concerns in the social sciences; in the second half I describe what has already been completed and what is hoped will be completed in the near future. The illustrations presented in this paper are, of necessity, poor-quality black and white copies. They represent the first attempts of our initiative to create new methods of visualizing large amounts of information about people over space in time for the release of the 1991 census small-area statistics.

Theory

Visualization is learning through the creation and observation of abstract images. It is most useful for analysing large amounts of data for which meaningful spatial projections can be determined. Alternative definitions proposed in the literature include:

"to form a mental image of something incapable of being viewed" (*Collins English Dictionary*);

"the act or process of interpreting in visual terms or putting into visible form" (*Webster's New Collegiate Dictionary*);

"Visualization offers a method for seeing the unseen" (McCormick et al, 1987, page 3).

Visualization is an attempt to take the graphic far beyond the realm of illustration. Only a picture can successfully describe the complexity of geographical and historical information. However, the written word suggests the processes which maintain that picture. Most graphics in social science are illustrations that distort what they attempt to portray to such an extent that a better explanation of the information is required in the text to supplement them. Visualization in social science aims to produce graphic descriptions that accurately portray the information they purport to contain.

True examples of visualization in social science do not illustrate the text—the text illustrates them. That is what makes these pictures different (Wood, 1985).

History

Most papers in the current explosion of material on visualization state that the field began in 1987 with the report by McCormick et al (for example, Anderson, 1989; Manuel, 1988). Others trace its origins back further to the early work in computer graphics of the 1960s (Neal, 1988); the 19th-century explosion of statistical graphics; their origins in the preceding centuries, and even to the first uses of maps and charts in Mesopotamia or China (for example, see Beniger, 1976; Bertin, 1981; Pickett and White, 1966; Royston, 1970; Tufte, 1983).

In my opinion, the current era of visualization began when researchers first realized that it was impossible to describe and study complex spatial arrangements without the use of graphics that could only be produced by computer. The first well-known work of this kind was produced by Mandelbrot (1980) and christened 'fractal geometry'.

Mechanical measuring devices created enough information to warrant the use of the graphical form. Later increases in the collection of population statistics led to their graphical presentation (Beniger and Robyn, 1978, page 2). As a result, the growth of the field has been strongly led by the availability of both data and advanced computing equipment. Just as the early pioneers required printed logarithmic graph paper on which to scribe their work, the advent of the graphics window has been a precursor to today's developments.

The report on "Visualization in scientific computing" (McCormick et al, 1987) marked political acceptance of the field in academia and the influx of substantial investment. The motivation for this action appears to have more to do with the US government's attempt to repel the success of Japanese industry than with the promotion of understanding (Frenkel, 1988, page 113). Nevertheless, the result has been a huge research effort attempting to exploit the visual image (Rosenblum, 1989, page 14).

Origins

The origins of visualization in social science lie in at least three separate fields—none of which has a dominant influence. These are thematic cartography, statistical (exploratory) graphics, and computer graphics. Their combination produces a very different field by taking from each what is most imaginative and jettisoning what is dogmatic, to achieve a single aim—the exploration of an idea through visual media (Feinberg and Franklin, 1975).

Cartography is a case in point. The section of the field concerned with mapping information in the social sciences refers to its work as thematic mapping (Petchenik, 1979, page 5), although the distinction is not clear. Less clear is the purpose of this work (Hsu, 1979, pages 118-119) and murkier still are the means by which that purpose is achieved. An enormous literature exists on the details of the task, but the substance of the activity is generally lacking. It is lacking because there is no agreement on precisely what is being mapped—a failure of the original definition. Interesting comments in this area are made by Wright (1938), Meuhrcke (1969; 1972), Bertin (1978), Kretschmer (1978), and Monmonier (1980).

The reasons for this failing appear to revolve around the historical subservience of thematic mapping in the field as a whole (Barbour, 1983; Robinson, 1982). This led to the adoption of an inappropriate methodology—principally that techniques suitable to the mapping of physical features on the surface of the earth could be borrowed (and slightly altered) to 'map' phenomena of interest to social science.

These phenomena, by definition, are concerned primarily with the lives of people who play out their existence in a very different metric (space) from that populated by rivers, seas, and mountains. At a simple level the lives of people are not static entities which can be viewed as if they were timeless. Similarly, their lives are not dominated by physical distances, but by distances measured in time and association.

This is accepted in the majority of social science in which little relevance is placed on the physical geography of society in its studies (Muehrcke, 1981). It may well be the origins of geography (in exploration, territory, and the control of land—upon which labour came free) which are really to blame. Nevertheless, the anachronism that is thematic mapping today remains—becoming embedded in the supposedly modern world of desktop mapping and geographical information systems (GIS). More recent recognition of some of these problems can be found in Tikunov (1986; 1987), Salischev (1987), Lyutyy (1989), Phillips (1989), and Visvalingam (1989).

Whereas the rest of human geography has moved forward from the simple maps of physical features and trivial spatial associations, one part—risen from the ashes of the 'quantitative revolution'—is trying to resurrect these mistakes of the past by linking people to Cartesian coordinates, calculating straight-line distances between them, and wrapping Theissen polygons around them. This is done with little concern for what (or even who) it is they are really studying. The primary interest appears to be in applying the latest 'tool' of GIS just to see what it can do. It is hard enough justifying the use of computers in serious social science research when it is done well, without being shackled by the quirks and limitations of some of the worst pieces of software in operation today (Taylor, 1990).

Detail

Computer graphics would not have enabled the Mandelbrot set to be found, had a coarse and meaningless set of areal units been imposed upon the complex plane. Statistical graphics would have failed to show relationships among variables that are distributed logarithmically, say, if the graph scales had not been changed to a more appropriate metric—even if it is more difficult to comprehend immediately. So human cartography fails in using units that are too coarse and a Euclidean metric that distorts and not by employing misleading choropleth map class intervals, garish colours, or placing legends in the wrong place. The real problems are not the superficial ones which receive so much attention and to which so many master's theses are directed. It is as if the field of cartography is being perfected rather than undermined (Jenks, 1975; Petchenik, 1985; Robinson and Petchenik, 1975).

These claims are not new, although they have traditionally been presented more as a doubt than as a condemnation. It is in part the very fact that such fundamental problems have been known for so long, while so little has altered, that makes criticism appropriate (Barabba, 1975, pages 24-25; Guptill and Starr, 1984; Szegö, 1987; Taylor, 1983; 1985; 1987; 1988). These claims are important to cartography, for if they are true, then the emerging field of visualization in social science will pass it by, leaving it to the study of its own history. Goodchild (1988, page 311) subtly presented a similar argument to this from within the discipline, stating that "its true potential lies in less conventional methods of analysis and display and in the degree to which it can escape its traditional constraints".

Commercial interest lies in persuading us to buy the most expensive machines and packages. This is not concordant with academic interest, despite academics' apparent desperation to obtain the 'latest' machines (Lathrop, 1988; 1989). These may well do something very fast, in perfect colour, with millions of pieces of information. The substantive interest is in what is being done so quickly (often something which is not worth doing at all), why use all those colours, and what

value does the information really hold? Visualization as a way of doing research does not require 16 777 216 hues, ten million pixels, fifty million individual records, and always or necessarily involve animation. It has to inform us, to teach us something we did not already know rather than blind us with ever-cascading information (Buttenfield and Ganter, 1990).

The concern in this paper is to channel the excitement and resources of visualization (Carlbom and Potmesil, 1988; Winkler et al, 1987, page 29) to an environment that is rich in the data it seeks to portray—the social sciences. If this is not undertaken within these sciences, it will be done from without and the results may not be to their liking (Prueitt, 1987, page 6). In fact the call for this research has already been made:

> "The general consensus in the scientific visualization field is that a broad commonality exists among the visual needs of all numerically intensive sciences ... we are keenly awaiting its application to fields with a shorter history in numerical computing, such as econometrics and the social sciences. Will users from these fields find this environment appropriate for their needs?" (Upson et al, 1989, page 41).

What is required is a closer look at what visualization promises and why this promise is not being met by conventional graphical methodology in the social sciences—in particular by cartography.

Information

It is digital information and the move away from its summary by numerical analysis in textual form which sustain visualization. It can be argued that the accelerating availability of such information was the primary influence in the emergence of the field and its invasion into new areas (Kennie and McLaren, 1988, page 737). Numerical information is one thing with which social science abounds (thanks to the 'inquisitiveness' of government and others). We do not need to invent equations or simulations to create the detailed information that gives visualization its beauty. Unlike other sciences, we do not rely on a scattering of weather stations, a sample of boreholes, or the emissions of particles which are scanned to give estimations of the material in a brain or the vegetation of a nation. We know the state of our object of analysis in great detail—when and where the people which make it up were, and what they were doing there.

More time and money is spent collecting information about people in our society than on any other subject. Yet to date that information receives the least attention from the form of analysis most suitable to analyze it—visualization. There have been many studies of what the future holds in this area of social science. Today that future is here, the machines now blink at formerly impossible problems and we have more data than we know what to do with. What is lacking is the imagination to exploit the limits of what is possible and available. Visualization provides a channel to express new ideas and different interpretations.

Equipment

It is common to suppose that such studies will require expensive equipment and highly trained people as was true for the state of the field in computer science as it emerged. The use of 'supercomputers' is a good example. Fortunately, it is beginning to be realized that for this work they are unnecessary and may even hinder development, although the field appears to be being used to justify their continued existence (Salzman and von Neumann, 1987, page 106). Most researchers use graphics workstations, but increasingly machines verging on the home computer market are being adopted. I prefer the cheaper option, despite the more expensive options that are readily available. Again, it is imagination—not 'pixel resolution'—that has to be exploited to create new images for the future rather than crisper ones

of the past. As for the highly trained people, most home computer programmers are children and enthusiasm is as great an asset as experience in these uncharted waters.

This research is concerned with examining what it is possible to visualize with machines that could not be achieved with pen and paper. This includes working to levels of detail as fine as the eye can see (if necessary), incorporating time in the image to raise the detail of our subject from two to three dimensions, and transforming to metrics that are more appropriate for the spatial study of society than those currently provided by physical geography (Finamore, 1982; Hsu, 1981; Monmonier, 1977; Tikunov, 1988; Tikunov and Yudin, 1987). The combination of these three aims provides our first clear glimpses of the information we have been collecting for years, but have been unable to see. Finally, this work opens up the possibility of examining relationships between previously isolated data sets without resorting to crude summary measures or suspiciously dissociated statistical techniques.

Future

There is, though, a more important question to answer than why past techniques are inappropriate. Namely, why should we wish to use visualization in social science? This requires a definition of this work that should suffice to conclude the rejection of past practice.

Visualization used to be called computer graphics, but the practical mechanics of the process soon became too simple to warrant technical expertise on how the pictures were produced. Instead, work began on the question of why images were produced and what they could show (Staudhammer, 1987, page 24). The computer had freed manipulation of the image just as the pen frees manipulation of text. Ideas could now be conveyed easily, and more importantly, discoveries made in a medium far more expressive and manipulatable than the written or spoken word. The image is not a substitute for writings, for it blatantly could not be described in a thousand words. Rather, with visualization, it should work in tandem with the text, each supporting the other.

Visualization is above all a way of doing research—a methodology. In these early days we are still exploring the method as much as the subject, constantly being surprised by how much more we can show, and how much more there is to see. It allows us to: (1) produce results and studies which do not oversimplify very complex structures while still presenting coherent information, (2) make comparisons which show the general trend and simultaneously allow for variations, and (3) use the best channel there is to our mind and to the minds of others.

If the development of visualization in social science is to grow most successfully, its roots must be recognized equally in thematic cartography, statistical graphics, and computer graphics. A working knowledge of all three disciplines is required to fashion a combination in which (just like the images the subject creates) the whole is far more than the sum of its parts. Visualization arises from, and creates, a picture of gestalt.

> "In conclusion, visualization should not be viewed as the end result of a process of scientific analysis, but rather as the process itself. More than simply the application of techniques for displaying data, visualization can be used as a paradigm for exploring regions of untapped reservoirs of knowledge. ... Visualization is not new, but its awareness by the general scientific community is" (Wolff, 1988, page 35).

Practice

In this section of the paper I describe work towards the general goal of visualization in social science that I have already completed at the North East Regional Research Laboratory. The work involves deciding the object of study, creating meaningful spatial metrics, amassing the large quantities of data required, choosing the correct hardware, and writing or buying satisfactory software to render or animate the results. Most importantly it involves deciding how the results should be displayed.

Cartograms

Area cartograms are images of geographical space transformed from that of the physical world so that the size of areal units is in proportion to some variable—usually their population. An algorithm has been designed to create area cartograms of over 100 000 areal units. The creation of cartograms is a necessary prerequisite to almost all visualization in human geography because the use of conventional maps makes images (derived from distributions based upon peoples) meaningless. At the most trivial level it is not possible to see physically the majority of the information being displayed upon a decent number of units (and that which is visible is heavily biased by the land areas of those units).

More fundamentally, the world of physical geography is of little (and decreasing) importance to what determines and perpetuates the distribution of attributes among people across space. Instead, the results of those distributions are spatially distinct and obvious once the social spatial distances between people are made clear. Spacing people evenly across the page or screen gives each one equal representation in the graphics produced; our conventional images are not only misleading, but biased toward representing the particular groups in society who have most access to land.

The creation, implementation, and justification of the algorithm are described more completely in another paper (Dorling, 1990). Briefly, areal units are represented by circles, the sizes of which are in proportion to their populations. These circles are arranged so that none overlap and almost all touch those with which their areal units share a common (physical) boundary. Thus an approximation to a continuous transformation is achieved which can produce cartograms of many more units than other algorithms written to date.

Resolution

A high spatial resolution—the use of many small areal units—is preferable as large units average the attributes of people over enormous populations. The average unemployment rate for London is almost as vague a measure as that for the whole country because it requires the combination of areas of greatly varying fortunes. The need for high resolution is not an attempt to approximate individual-level data. Rather, it is an attempt to gain insight about the mix of people in different areas—which has a value of its own right—showing us who is living near to whom, rather than what so-and-so is doing. To be of use, though, we have to try and get to the levels at which people do live 'together' as communities, neighbourhoods, or estates. Many people have never been to the other side of the city in which they live, let alone know it well.

In Britain, wards are at roughly the level of (if not perfectly demarcating) areas seen as socially meaningful, in which many children will go to the same secondary school and people have similar (spatial) opportunities, housing, and jobs. They also have the advantage of being the most basic units for which political choices are made by people (that is, electing local councillors) and for which information between censuses is disseminated. For the decennial censuses, wards are divided into enumeration districts (EDs—approximately 'street' or 'block'), for the finest

reporting purposes (graphic treatment of this is discussed in Aangeenburg, 1976; Census Research Unit, 1980; MacEachren, 1982; 1987; Rhind, 1983).

The cartogram algorithm has been successfully applied to all the conventional sets of areas in Great Britain (that is, counties, districts, constituencies, office areas, and so on) and also to the 10444 'wards' of the 1981 census and the 129211 populated EDs. With the ward cartogram it is possible to discern individual spatial units, whereas the ED version produces a continuous mat of shades or colour,

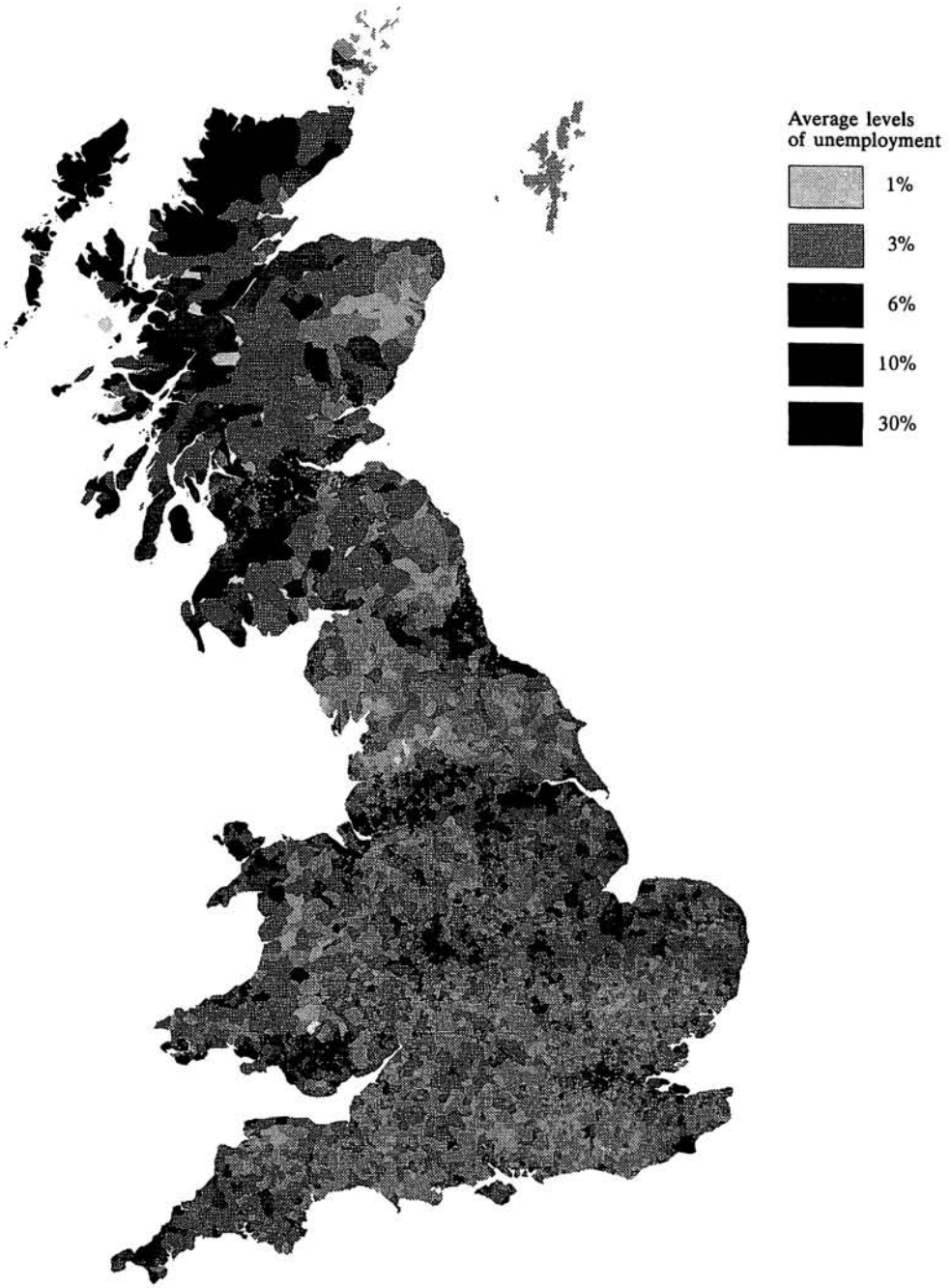


Figure 1. The concentration of unemployment by ward, 1981.

each tiny pixel representing approximately 100 people (Sibert, 1980). There are over one million pixels in the images (just under half are blank 'sea'). An average ED would be made up of four of these, the areal units tend to be distributed hexagonally in the cartogram as the median number of neighbouring areal units (for almost any set) is six.

Figure 1 shows a conventional map of unemployment by ward. Figure 2 shows the same distribution (using the same units and shading) upon a population cartogram.

Figure 2. The distribution of unemployment by ward, 1981.

Figure 3. The distribution of Irish born in Britain 1981, by enumeration district, on a population cartogram.

The differences illustrate both the utility of the cartogram and the inherent bias of the map which draws the eye towards large unpopulated regions, away from the cities where people live. Figure 3 shows the distribution of Irish-born people upon the ED cartogram. The half of the country in which over-average numbers of Irish-born people lived is shaded black.

This work has been extended to include colour (unfortunately not reproduced here, see Dorling, 1992). The printed images are made up of six million eight-colour pixels, dithered to produce a continuous effect, the most complex and enlightening variation of which is trivariate shading (Bertin, 1983; Dutton, 1982; Levkowitz, 1988; Olson, 1975; 1981; 1987). Here the three subjective primaries of red, yellow, and blue represent three variables of the areal units—say proportions voting for three different political parties. The colours then mix to the subjective secondaries of orange, green, and purple and all possible shades in between, including white, which implies minimal values for all three and black meaning saturation.

The images can undergo spatial smoothing to simplify their messages (though this may detract from the sincerity of the message). This work extends the early investigations into numerical generalization, made more meaningful, for as the smoothing is 'averaging' equal numbers of people with their nearest neighbours rather than with whomsoever should live within so-many metres. It should be noted that this research has its origins almost exclusively in the seminal work of Tobler on cartograms, continuous chloropleth mapping, numerical generalization, and many other topics (of particular relevance are Tobler, 1959; 1961; 1963; 1968; 1973; 1979; 1985; 1987; 1989).

Change

The research into depicting national 1981 census distributions has been extended to investigate patterns in the 1971 census and between the censuses. The majority of 1981 EDs vary little (or not at all) from their 1971 counterparts. The standard software which accesses the 1981 census offers change information at the level of amalgamation EDs, designed to maintain constant boundaries, called tracts. There is general dissatisfaction with the level of error in the conventional combined census tract approach. It is also not extendable to the census information for 1991 (on which work is beginning as this paper was revised) and so an alternative approach had to be found (McKee, 1989; Norris and Mounsey, 1983).

Indiscriminate aggregation is one of the worst forms of error. This statement is patently obvious to me but not to many who use the census, and would require another paper to explain adequately. This is not to advocate working at the smallest resolution possible, although that has been done here and work at the level of the individual is tempting. In fact much of our most recent work has been at the ward level. What is to be avoided is the unimaginative attitude that, as long as the physical boundaries coincide, the results are valid. Two ways of avoiding aggregation as soon as a spatial boundary wobbles between time periods have been employed in this work.

The most comprehensive method is to keep the two sets of boundaries separate, to treat the situation just as it happened. This has been done to create a set of changing (but linked) parliamentary constituencies over the major redistributions between 1955 and 1987. Such a set requires detailed work (mostly manual still) to contend with changes of name as well as the creation, joining, and disappearance of places in population space.

An automated method is required if the two sets of EDs are to be linked between the censuses. This is simply achieved by finding all the nearest 1981 neighbours of each 1971 ED and the nearest 1971 neighbours of each 1981 ED.

Figure 4. The changing distribution of total population in Britain 1971–81: rates by enumeration district, on a population cartogram.

More than just the first neighbour of each set could be selected, but the simplest approach was found to be more than adequate. In general there is largely one-to-one correspondence, but the linkage is flexible enough to cope when, for instance, a series of 1971 districts must link to a single 1981 survivor (after, say, 'slum' clearance), or when a 'new estate' of 1981 EDs must link to a single 1971 'green field'. Even (the rare) complete redistributions can easily be accommodated with none of the aggregation of dozens of EDs that use of tracts can involve. Figure 4 shows the simple distribution of EDs which (roughly) gained or lost people over the 1970s.

It is vitally important to remember that any misallocation of people from an ED caused by this procedure will only misplace them to a neighbouring district in the next census—a visual error of only one pixel in a million (if we assume the district would be coloured differently as a result). More importantly still, at the level of EDs (and higher) we are not seeing changes in people, but changes in who is living in particular houses and flats. Few of the inhabitants of most 1971 EDs will still reside in those areas by 1981. People move, and we all too easily forget that in our excursions into the pedantry of census definitions. The spatial smoothing mentioned above not only instantly (at the first pass) removes any memory of these minor misclassifications, but, if done with a wide enough bandwidth, will encompass many of the changes due to migration (which is mostly local).

The downside of being able to investigate all this information so easily is that it has to be collected first. Although the census has been taken and stored and numerous years of work have been spent making it accessible to academics, actually obtaining this data is harder than analyzing and visualizing it. The 1981 census can only be accessed through archaic mainframe packages and from magnetic tape. Reading the values of a single variable for all EDs takes a long time and many changes of tape.

For the 1971 census I wrote my own simple access routine. The entire 100% population census for all cells and every ED in Britain can easily be compressed to a single 29 MB file; the 10% population data can be compressed to a 10 MB file. These figures represent rates of 95% compression over the space taken up on tape—with no information being lost. Reading any set of figures is done in a matter of seconds from a single disc file. The implications of this technical detour are important for future work in this area of social science. Put simply, there are no more 'large data sets'.

Method

Once the cartogram has been created, the data gathered and linked, there is still the question of how to show the changes. Deviation from the expected (national) change for an ED was chosen. So, if the national proportion of elderly went up by 20% over the ten years, only those districts showing a rise greater than this would appear darkly coloured; those with a lesser rise, lightly coloured. For trivariate shading only the quartiles of the three distributions were distinguished, still generating sixty-four possible shades. The images are thus of relative differences and changes, and the general (detailed) spatial relationships between these distributions. There is, of course, no problem of districts containing too few people, as on the population cartogram, such districts would be too small to see!

All the major census variables relating to people were painted on these high-resolution colour cartograms. Their ages, gender, children, birthplaces (England, Scotland, Wales, Northern and Southern Ireland, Asia, or Africa including the Caribbean), employment (active, inactive, student, or retired), occupation ('professional', 'intermediate', or 'supervised') and so on. This was done for 1971, 1981, and also

for the distributions of change (Dorling, 1991). The whole process was then repeated for Greater London alone and enlarged for the area showing most distinct structure. The next stage is to extend the study as soon as the results of the 1991 census are available. The technical problems associated with this will be trivial (as they have been solved for two censuses already and the same basis will be used again). The theoretical difficulties of dealing with such fine change between so many areas is more difficult, but can doubtless be overcome.

Means

It is important to stress that the eradication of the 'large data set' was, appropriately, not made possible by using a large computer. An Acorn Archimedes home microcomputer was used for all this work on visualization (though a Sun Sparc station was useful for data collection from the aging mainframe). This hardware is within the budget of a postgraduate student—you can take your work home. The software used is a collection of mostly free or extremely cheap applications, simple enough for a child to use, powerful enough in its generality to outclass options that cost fifty times as much today. It is the bypassing of problems which do not deserve to exist (in this decade) which makes use of these machines such an advantage. For instance, the operating system ensures that every application can print to every type of printer I have access to in the University of Newcastle (though some minor tinkering still helps).

If you can write short programs (in a child's programming language) you can create any two-dimensional image you can imagine—and many three-dimensional and animated ones as well. You have to do a bit better than many of today's so-called geographic information systems in how you organize the information, but then many of them were designed over twenty years ago. I wasted several weeks using one particular package, before probing into its internal data structure to discover the mess that the badly structured code had to contend with. Remarkably in 1991 it is still quicker, easier, more efficient, and much more productive to write programs to handle geographic information from scratch the way you want it. This assumes, of course, that you do not wish to do exactly what the standard packages provide. Too few people who use them question the value of what they are doing and ask whether they are doing it because this is how they wish to examine the information or simply because it is how they are allowed to do it by the package.

Unemployment

The first subject to be studied outside the census in this visualization research was unemployment. Unemployment varied greatly between the censuses. Apart from being a serious social problem it also offers a fine space-time series to investigate. Figures are available (after laborious extraction) for over 800 amalgamations of office areas and for every month of at least the last twelve years—producing a good-sized spreadsheet. Deviations from the 'expected' rate can be plotted on a cartogram for each year as has been done in figure 5, the crude results for counties are shown in figure 6. It should be noted that these illustrations were based on the latest available information of two years ago. The picture may have changed markedly in the meantime.

A problem of showing deviations from the expected level for a particular area and year is the tendency to highlight the situation at the beginning and end of the series. This is most clear in the example, as the 'celtic fringe' of the late 1970s is replaced by the 'inner cities' of the late 1980s. The confused pictures in between result from that fact that every area must at some time do worse than its average rate.

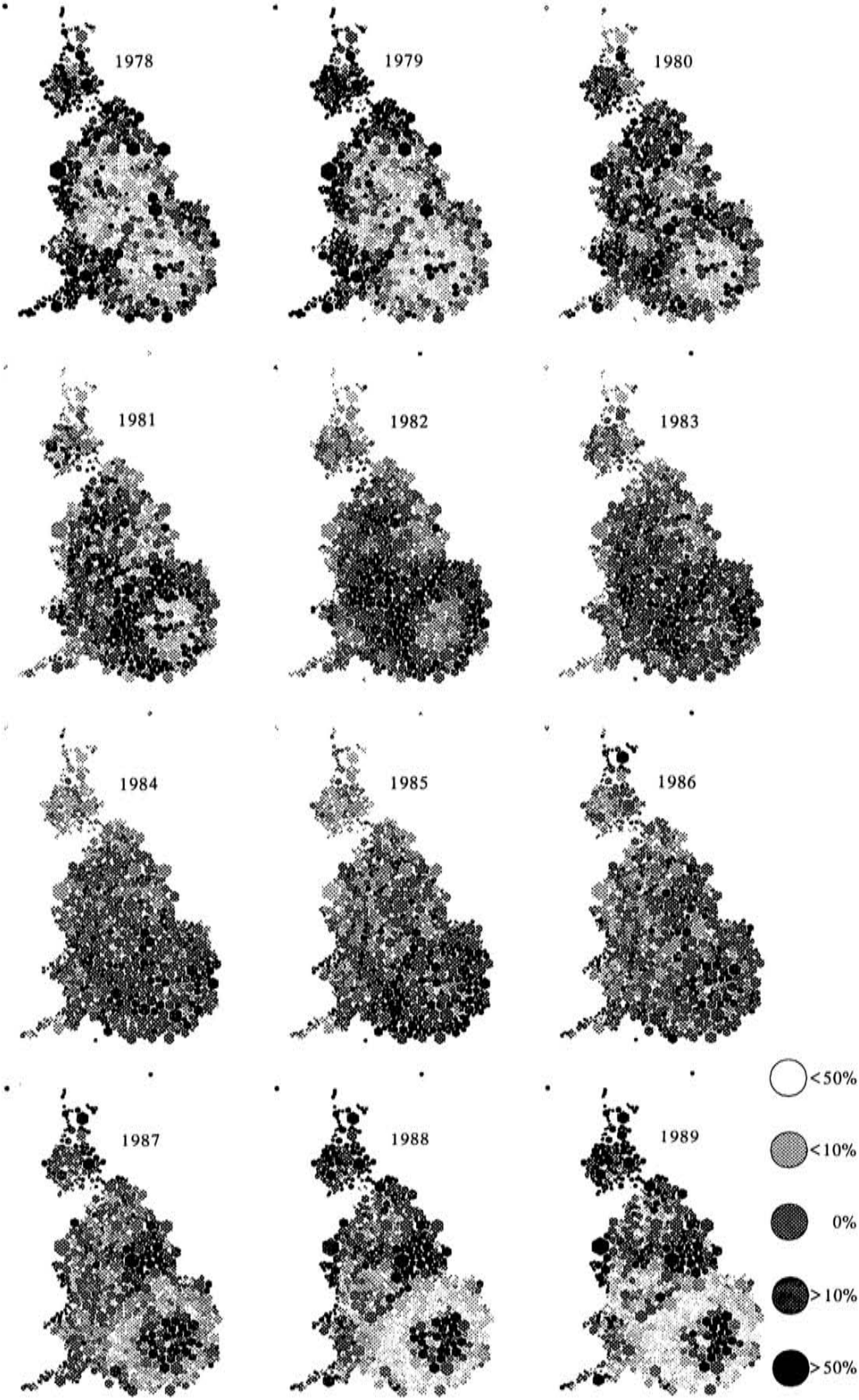


Figure 5. The space–time trend of unemployment in Britain, 1978–89 (on 1981 Amalgamated Office Area population cartograms, the scale indicating deviation from space–time independence).

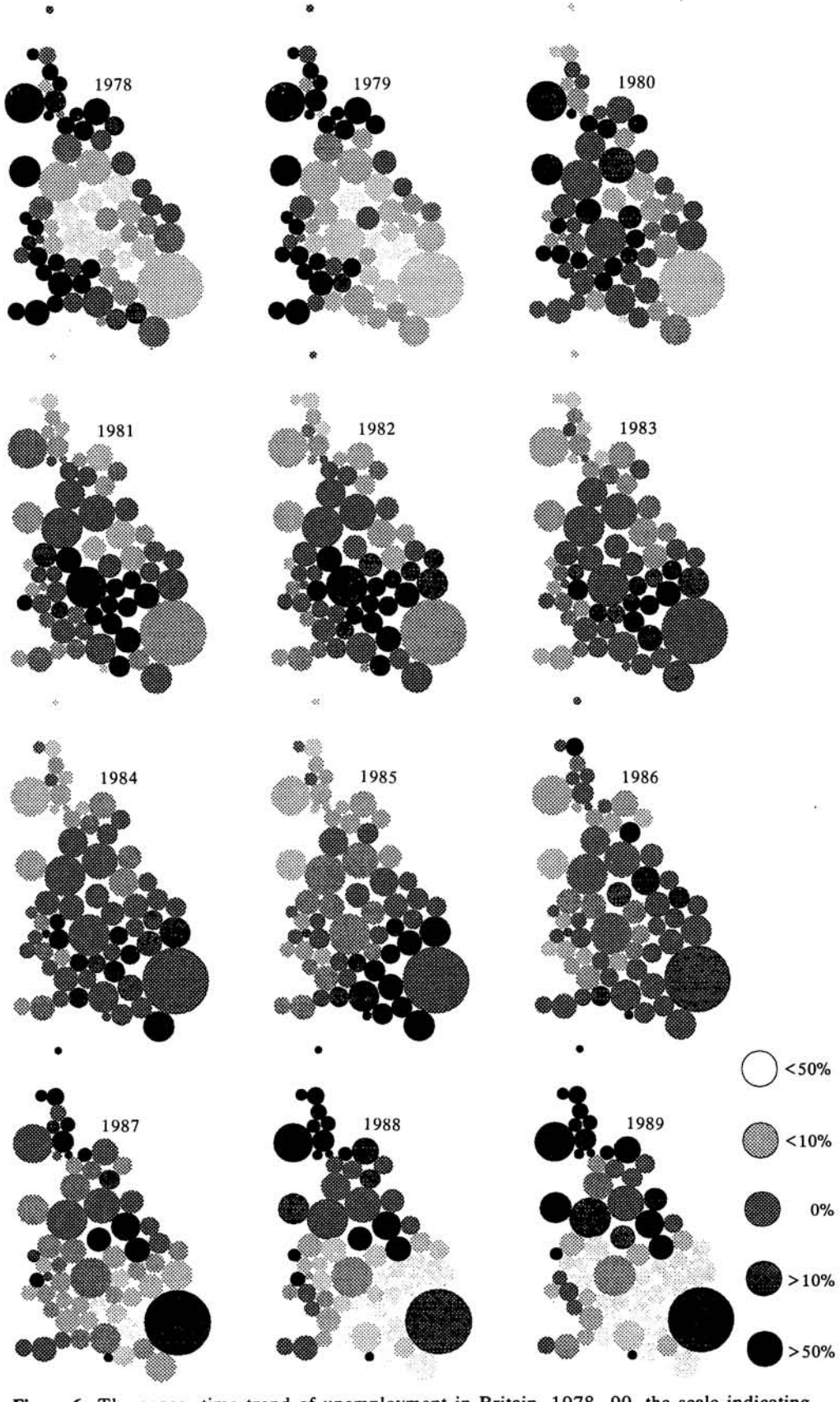


Figure 6. The space–time trend of unemployment in Britain, 1978–90, the scale indicating deviation from space–time independence, on 1981 county and Scottish region population cartograms.

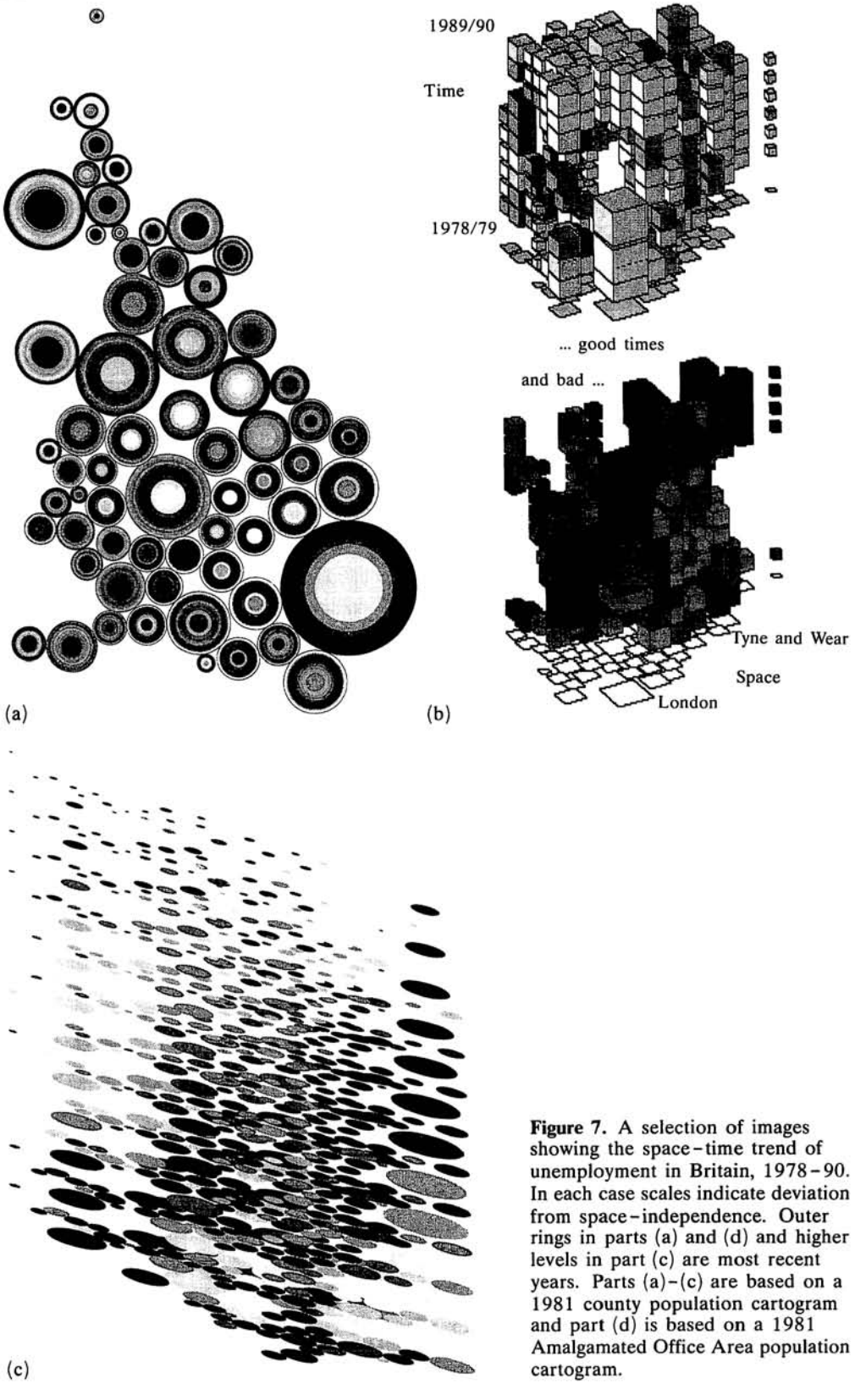
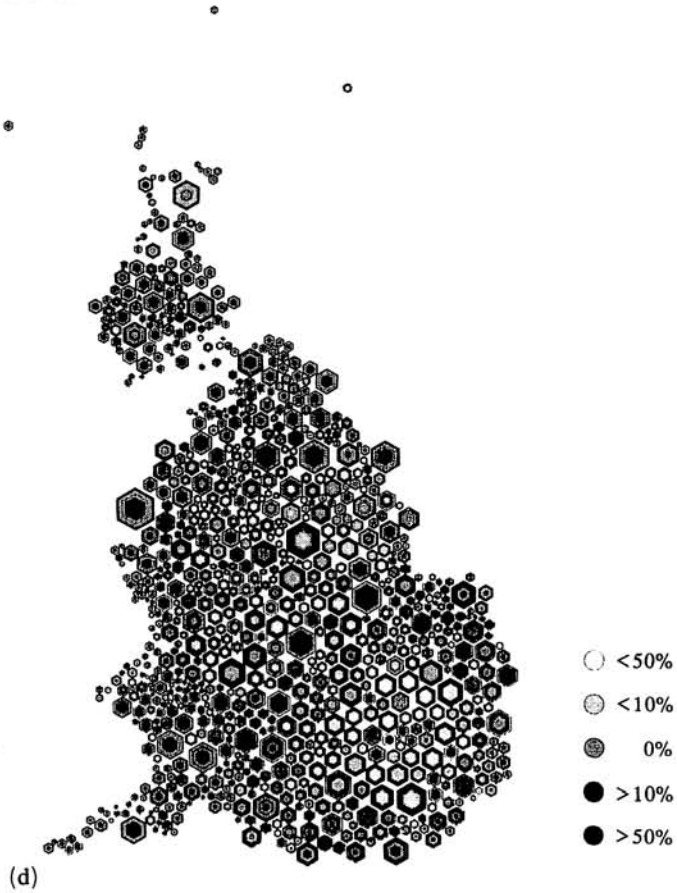


Figure 7. A selection of images showing the space–time trend of unemployment in Britain, 1978–90. In each case scales indicate deviation from space–independence. Outer rings in parts (a) and (d) and higher levels in part (c) are most recent years. Parts (a)–(c) are based on a 1981 county population cartogram and part (d) is based on a 1981 Amalgamated Office Area population cartogram.



(d)

Figure 7 (continued)

For example, when the rest of the country was also suffering, Liverpool was faring 'relatively' well in the early 1980s, though it was still probably worse off than it was before or after. This opens up the whole question of how we should compare times and places; against the national picture and decade, or against recent years and nearby places?

More sophisticated images were attempted, but failed to show the complex three-dimensional regional-temporal structure (a selection is shown in figure 7). We are currently in the process of reworking this research to address the problems raised above to find better solutions for exploring three-dimensional structures than multiple two-dimensional images. We also hope to extend the work from single measures to explore more complex relationships in the 1980s between the sexes, part-time and full-time employment, types of industry, and even to include such things as house prices and local voting patterns (shown in Dorling, 1991).

Voting

General elections have already been studied in this research programme at the national levels for the ten elections between 1955 and 1987 (and also the effect of the 1979-83 redistribution). This work was novel in that it represents the first such study at the constituency level—using the constituencies that existed at the time of each of the elections—and involved extensive animation work. Cartograms based on electorates (which changed dramatically over time) were employed. And trivariate shading was used to show the mix of the voting between the three main parties.

Who won, who came second, the vote, the direction of change in the vote, the seats which changed hands, the level of abstentions were all drawn for each election. They were also all used as key frames in animations, some series more successfully than others. The most enlightening animations were ones in which the spatial relations were dispensed with and the distributions of constituency votes (by party) were shown as bouncing points inside a triangle. Figure 8 shows the most recent four elections on these triangles; without animation the movements cannot be represented clearly (Dorling, 1992). Figure 9 shows the key frames used for the animation of the changing spatial distribution of levels of turnout.

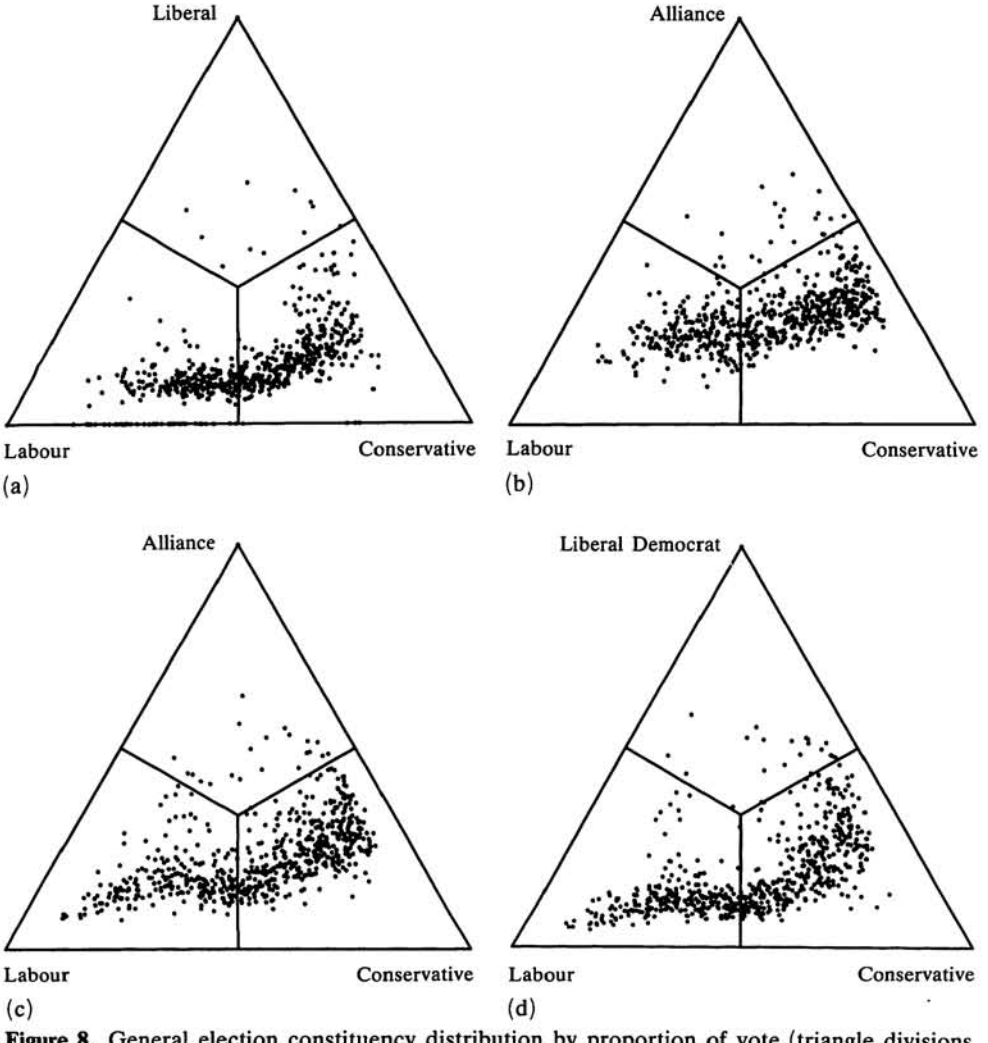


Figure 8. General election constituency distribution by proportion of vote (triangle divisions represent winning portion for each main party): (a) 1979, (b) 1983, (c) 1987, (d) 1992.

Flows

The last subject of this paper concerns the possibilities of visualizing flows of people, the last major type of problem to be addressed in this programme of research. Flows in the social sciences represent the most challenging situation because they usually contain much more complex structures than those in other fields to which visualization has been applied. So here again we have to imagine new ways of depicting the information before we can see it.

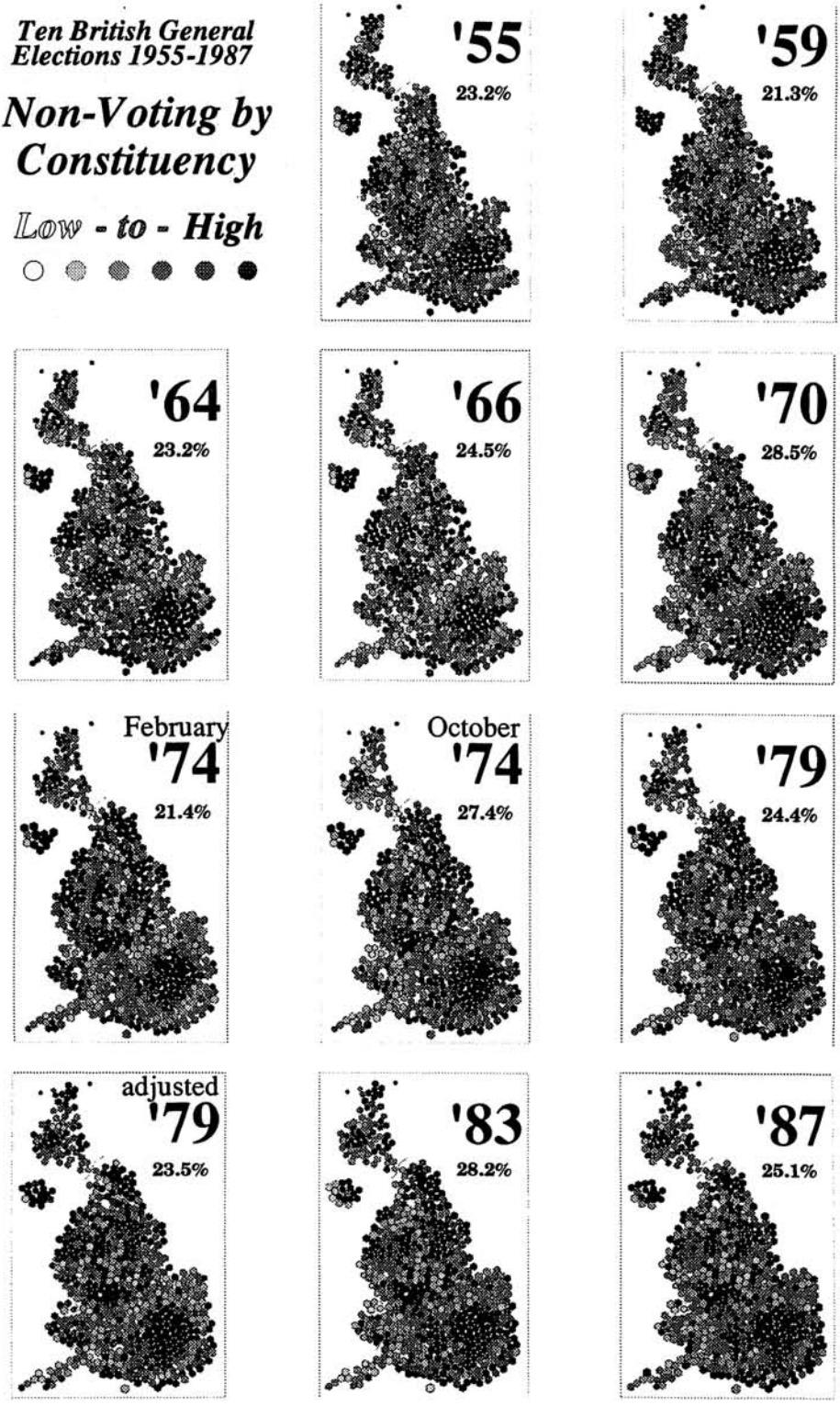


Figure 9. The key frames used for the animation of the changing spatial distribution of levels of turnout. Percentages under the year give the national average abstention rates.

The two best known categories of flow in geograpy are divided by their temporal resolution: travel to work and migration. The first is obviously more localized than the second which forms a very complex matrix across the country. By the visualization of flows we do not mean depicting the average distance or direction of flow for a given set of areas. Rather we are attempting to show both the actual 'streams' of movement and their 'ensemble' simultaneously. This is not easy.

The technical difficulties of collecting and storing the data are also not trivial. The 1981 census presents the complete (excluding nuances with Scotland) ward-to-ward matrix of flow of both migration and travel to work. Collecting these through the standard procedures is practically impossible. Luckily previous work by the Centre for Urban and Regional Studies has made these data relatively accessible. Compaction can achieve very high levels of performance as the matrices are extremely sparse and the flows vary little in magnitude. Complete matrices describing migration disaggregated only by sex have been converted into files of about 1.5 MB in size with no loss of information.

Figure 10 shows the results of simply plotting a line of varying thickness between every pair of wards where more than ten people travel to work (out of a 10% sample). It shows the collection of links which create our towns and cities but is very confused and extremely crowded—overemphasizing rare long-distance commuting, completely hiding the vast majority of journeys within towns. Many schemes are in the pipeline to explore this information in further detail.

The population cartogram base is an obvious way to open up the cities (Phillips, 1977). It destroys the comprehension of Euclidean distance, but then it is the time

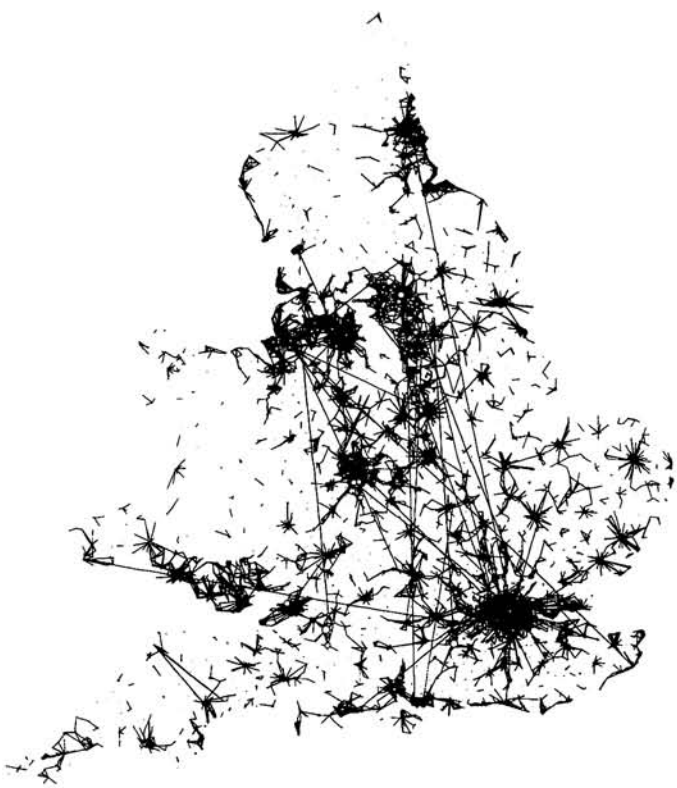


Figure 10. 1981 journey-to-work flows of over ten people between wards from the 10% sample.

to get to work rather than the miles which is important. Even more important are the patterns made by where the people at work live, rather than by how far they had to go to get there. The population cartogram is also a very crude approximation to travel time equalization. A myriad of tiny arrows can show direction of flow, their length indicating distance, their size indicating quantity, their colour ...?

Single arrows for each area will not show the links between areas. The fewer spatial units there are, the more elaborate the technique can be. The National Health Service central register gives migration flows between ninety-eight family practitioner areas for every year since 1976. Recent experiments suggest it may be possible to show the relative strengths of all migration streams between as many areas as these (approximately 10 000 flows). The technique must be able to show the 'cascading' of people from Scotland and Wales to London, and the diffusion from London to the rest of the South, simultaneously with all the relations between other pairs and groups of places, in perspective. What is even more ambitious is to try then to show how this situation has changed over the course of the last decade.

Animation

Another paper (Dorling and Openshaw, 1992) is concerned exclusively with the subject of animation. In particular, work on depicting the space-time incidence of crime and cancer are discussed. These examples generally use time in the animation to substitute for real time (speeding up days or years into seconds). However, animation can also be used in visualization to explore essentially static situations.

The first situation in which it can be useful is to zoom and pan in and around a high-resolution two-dimensional image: highlighting areas of particular interest, showing the entire situation, and moving across the country. This may be more useful in an illustrating capacity rather for visualization, but may still be very effective. The animation work on general elections is being rerecorded with these techniques in mind.

The problem with conventional animation in social science is that it is not possible to compare times as you compare areas; one period follows the next and is quickly forgotten. The second suggested use for static animation is to explore three-dimensional space-time structures. The space-time image is created—not as a two-dimensional surface—but as a truly three-dimensional structure of objects (or three-dimensional surfaces) in space. Then a camera is flown around and through the scene, circling the most striking features, contrasting other constructions, providing a general overview and detailed examination of a complicatedly structured situation.

Integration

Too much time has been spent in this programme of research (and this paper) on the technical problems of collecting and storing the information to be viewed. This is partly because of the complex and ever-changing set of areal units to be dealt with—one set for every new set of variables it appears. But it is also because of the tedium of accessing and learning the many antiquated packages and machines that provide and store the data. Once all the sets are compressed and held in a few small files on the home micro the interesting work can begin.

Integration is the final promise of visualization to social science. After all, all of this information is about the same set of people who lived in Britain through the 1980s; should its sum not tell us more than its parts? Conventional spatial research dissects and subdivides the field, finally showing the coarse distribution of a single variable over a few areal units, for a single time period, distorted to show land area, and to cap it all shaded with cross-hatched lines to give the reader a headache (Blakemore, 1987).

I have shown through this work that simple trivariate shading allows at least three variables to be compared for almost any number of areal units. More complex possibilities lie just around the corner. If we can draw arrows to depict migration streams, why not colour them by some aspect of where the migrants are coming from, or where they are going to, or both? Why not place these arrows upon the plane showing the near-continuous distribution of other census variables? Why not animate the images to rotate the shading slowly to represent other variables, allow the arrows to move, to restructure over time, or whatever.

This research programme does not call for a complete break with the past. There are many good ideas in the past which have been overlooked more recently. One of particular interest is Chernoff faces, but extended now to 'Chernoff crowds' (Rahu, 1989). A face for every parliamentary constituency, say, using the 633 circles of the mainland cartogram makes it simple (Dorling, 1993). The size of the face is the number of voters, the colour is the vote, the smile (or frown) unemployment, the eyes show age structure (big for young) and so on. Animation would allow you to zoom and pan around the country. The crowd could even be brought to life to show how its 'mood' changed over time!

There is no longer a technical problem to doing such work, anything that can be imagined can be implemented. The problem is in deciding what to do first, and envisioning these new possibilities (Dorling, 1991).

References

Aangeenburg R T, 1976, "Small area data as a constraint to computer graphic analysis", in *Proceedings of the Workshop on Automated Cartography and Epidemiology* National Center for Health Statistics, Office of Health Research, Statistics, and Technology, US Department of Health, Education, and Welfare, Hyattsville, MD, pp 59-66

Anderson G C, 1989, "Images worth thousands of bits of data" *The Scientist* **3**(1) 16-17

Barabba V P, 1975, "A challenge to cartographers" *Auto-carto 2* Proceedings of the International Symposium on Computer-Assisted Cartography, American Society for Photogrammetry and Remote Sensing and the American Congress on Surveying and Mapping, Bethesda, MD, pp 15-26

Barbour M, 1983, "The use of maps: the topographic and thematic traditions contrasted" *The Cartographic Journal* **20** 76-86

Beniger J R, 1976, "Science's 'unwritten' history: the development of quantitative and statistical graphics" Abstracts of the Annual Meeting of the American Sociological Association, New York; copy available from the British Library, Boston Spa, Wetherby, W. Yorks

Beniger J R, Robyn D L, 1978, "Quantitative graphics in statistics: a brief history" *The American Statistician* **32** 1-11

Bertin J, 1978, "Theory of communication and theory of 'the graphic'" *The International Yearbook of Cartography* **18** 118-126

Bertin J, 1981 *Graphics and Graphic Information-processing* (Walter de Gruyter, Berlin)

Bertin J, 1983 *Semiology of Graphics* (translation of *Semiologie Graphique* by W J Berg) (University of Wisconsin Press, Madison, WI)

Blakemore M, 1987, review of *Maps by Computer: The 1981 Census Atlas of Brighton and Hove* in *Geography* **72** 276-277

Buttenfield B P, Ganter J H, 1990, "Vizualization and GIS: what should we see? what might we miss?", in *Proceedings of the 4th International Symposium on Spatial Data Handling* volume 1, pp 307-316

Carlbom I, Potmesil M, 1988, "Modelling and display of empirical data" *IEEE Computer Graphics and Applications* **8** 10-11

Census Research Unit, 1980 *People in Britain: A Census Atlas* (HMSO, London)

Dorling D, 1990, "A cartogram for visualization" NERRL RR 90/4, Centre for Urban and Regional Development Studies, University of Newcastle upon Tyne, Newcastle upon Tyne

Dorling D, 1991 *The Visualization of Spatial Social Structure* PhD thesis, Department of Geography, University of Newcastle upon Tyne, Newcastle upon Tyne

Dorling D, 1992, "Stretching space and splicing time: from cartographic animation to interactive visualization" *Cartography and Geographical Information Systems* **19** forthcoming

Dorling D, 1993, "Cartograms for visualizing human geography", in *Visualization and GIS* Eds D Unwin, H Hearnshaw (Belhaven Press, London) forthcoming

Dorling D, Openshaw S, 1992, "Using computer animation to visualize space-time patterns" *Environment and Planning B: Planning and Design* **19** 639-650

Dutton G, 1982, "Bivariate construction of equal-class thematic maps" *Auto-carto 5* Proceedings of the 5th International Symposium on Computer-Assisted Cartography, American Society for Photogrammetry and Remote Sensing and the American Congress on Surveying and Mapping, Bethesda, MD, pp 285-296

Feinberg B M, Franklin C A, 1975, "Social graphics bibliography" Bureau of Social Science Research Inc., Washington, DC

Finamore P M, 1982, "The value of cartograms as communication devices" MSc thesis number 60933, Indiana University, Bloomington, IN

Frenkel K A, 1988, "The art and science of visualizing data" *Communications of the ACM* **31** 110-121

Goodchild M F, 1988, "Stepping over the line: technological constraints and the new cartography" *The American Cartographer* **15** 311-319

Guptill S C, Starr L E, 1984, "The future of cartography in the information age", in *Computer-assisted Cartography Research and Development Report* Ed. L E Starr, International Cartographic Association, Commission C, pp 1-15; copy obtainable from the British Library, Boston Spa, Wetherby, W. Yorks

Hsu M L, 1979, "The cartographer's conceptual process and thematic symbolization" *The American Cartographer* **6** 117-127

Hsu M L, 1981, "The role of projections in modern map design", in *Cartographica Monograph 27. Maps in Modern Geography—Geographical Perspectives on the New Cartography* Ed. L Guelke **18** 151-186

Jenks G F, 1975, "Contemporary statistical maps evidence of spatial and graphic ignorance" *Auto-carto 2* American Society for Photogrammetry and Remote Sensing and the American Congress on Surveying and Mapping, Bethesda, MD, Proceedings of the International Symposium on Computer-Assisted Cartography, pp 51-60

Kennie T J M, McLaren R A, 1988, "Modelling for digital terrain and landscape visualisation" *Photogrammetric Record* **12** 711-745

Kretschmer I, 1978, "The pressing problems of theoretical cartography" *International Yearbook of Cartography* **18** 33-39

Lathrop O, 1988, "Data visualization: the emergence of a new industry", in *Electronic Imaging '88* International Electronic Imaging Exposition and Conference, 2-6 October, World Trade Center, Boston, MA, Institute for Graphic Communication, 77 Rumford Avenue, Walthum, MA, pp 1091-1093

Lathrop O, 1989, "The current state of data visualization as a buiness", in *Conference Proceedings, 10th Annual Conference and Exposition Dedicated to Computer Graphics* NCGA Philadelphia, April, pp 282-286; copy available from the author, Cognivision Inc., Westford, MA 01886

Levkowitz H, 1988 *Color in Computer Graphic Representation of Two-dimensional Parameter Distributions* PhD thesis number DA8824759, University of Pennsylvania, Philadelphia, PA

Lyutyy A A, 1989, "Current problems of geographical cartography" *Mapping Sciences and Remote Sensing* **26** 1-13

McCormick B H, DeFanti T A, Brown M D (Eds), 1987, "Visualization in scientific computing—a synopsis" *IEEE Computer Graphics and Applications* **7** 61-70

MacEachren A M, 1982, "Thematic map accuracy: the influence of enumeration units size and compactness", paper presented at ACSM-ASP 42nd Annual Convention, March, Denver, CO, American Society for Photogrammetry and Remote Sensing and the American Congress on Surveying and Mapping, Bethesda, MD, pp 512-521

MacEachren A M, 1987, "The evolution of computer mapping and its implications for geography" *Journal of Geography* **86** 100-108

McKee C, 1989, "The 1971/1981 census change files: a user's view and projection to 1991", SERRL Working Report 13, Department of Geography, Birkbeck College, University of London, London

Mandelbrot B B, 1980, "Fractal aspects of the iteration of $z \to \lambda z(\lambda - z)$" *Annuals of the New York Academy of Sciencs* **357** 249-259

1988, "Graphics visualization is a real eye opener" *Electronics* **61** 91-92

Monmonier M S, 1977, "Maps, distortion and meaning", Resource Paper 75-4, The Association of American Geographers, Washington, DC

Monmonier M S, 1980, "The hopeless pursuit of purification in cartographic communication: a comparison of graphic-art and perceptual distortions of graytone symbols" *Cartographica* **17** 24-39
Muehrcke P, 1969 *Visual Pattern Analysis: A Look at Maps* PhD thesis, Department of Geography, University of Michigan, Ann Arbor, MI
Muehrcke P, 1972, "Thematic cartography", Resource paper 19, Association of American Geographers, Washington, DC
Muehrcke P, 1981, "Maps in geography", in *Cartographica Monograph 27. Maps in Modern Geography—Geographical Perspectives on the New Cartography* Ed. L Guelke **18**(2) 1-41
Neal M, 1988, "When did scientific visualization really begin?" *IEEE Computer Graphics and Applications* **8** 8-9
Norris P, Mounsey H M, 1983, "Analysing change through time", in *A Census User's Handbook* Ed. D Rhind (Methuen, Andover, Hants) pp 265-286
Olson J M, 1975, "The organization of colour on two-variable maps", in *Auto-carto 2* Proceedings of the International Symposium on Computer-Assisted Cartography, American Society for Photogrammetry and Remote Sensing and the American Congress on Surveying and Mapping, Bethesda, MD, pp 289-294
Olson J M, 1981, "Spectrally encoded two-variable maps" *Annals of the Association of American Geographers* **71** 259-276
Olson J M, 1987, "Color and the computer in cartography", in *Color and the Computer* Ed. H J Durrett (Academic Press, Orlando, FL) pp 205-220
Petchenik B B, 1979, "From place to space: the psychological achievement of thematic mapping" *The American Cartographer* **6** 5-12
Petchenik B B, 1985, "Value and values in cartography" *Cartographica* **22** 1-59
Phillips D J, 1977, "The visual effectiveness of flow maps and cartograms", MA thesis, Graduate Faculty, University of Georgia, Athens, GA
Phillips R J, 1989, "Are maps different from other kinds of graphic information?" *The Cartographic Journal* **26** 24-25
Pickett R M, White B W, 1966, "Constructing data pictures", paper presented at Seventh National Symposium on Information Display, October, Society for Information Display, Boston, MA, pp 76-81
Prueitt M L, 1987, "A window on science" *IEEE Computer Graphics and Applications* **7** 4-8
Rahu M, 1989, "Graphical representation of cancer incidence data: chernoff faces" *International Journal of Epidemiology* **18** 763-767
Rhind D, 1983, "Mapping census data", in *A Census User's Handbook* Ed. D Rhind (Methuen, Andover, Hants) pp 171-198
Robinson A H, 1982 *Early Thematic Mapping in the History of Cartography* (University of Chicago Press, Chicago, IL)
Robinson A H, Petchenik B B, 1975, "The map as a communication system" *The Cartographic Journal* **12** 7-15
Rosenblum L J, 1989, "Visualizing oceanographic data" *IEEE Computer Graphics and Applications* **9** 14-19
Royston E, 1970, "A note on the history of the graphical presentation of data", in *Studies in the History of Statistics and Probability* Eds E S Pearson, M G Kendall (Charles Griffin, London) pp 173-181
Salishchev K A, 1987, "On restructuring thematic cartography: trends and problems" *Mapping Sciences and Remote Sensing* **24** 285-291
Salzman D, von Neumann J, 1987, "Visualization in scientific computing: summary of an NSF-sponsored panel report on graphics, image processing, and workstations" *International Journal of Supercomputer Applications* **1** 106-108
Sibert J L, 1980, "Continuous-colour choropleth maps" *Geo-Processing* **1** 207-217
Staudhammer J, 1987, "Supercomputers and graphics" *IEEE Computer Graphics and Applications* **7** 24-25
Szegö J, 1987 *Human Cartography: Mapping the World of Man* (Swedish Council for Building Research, Stockholm)
Taylor D R F, 1983, "Some conclusions and new challenges", in *Progress in Contemporary Cartography, Volume II: Graphic Communication and Design in Contemporary Cartography* Ed. D R F Taylor (John Wiley, Chichester, Sussex) pp 285-307
Taylor D R F, 1985, "The educational challenges of a new cartography" *Cartographica* **22** 19-37

Taylor D R F, 1987, "The art and science of cartography: the development of cartography and cartography for development" *The Canadian Surveyor* **14** 359-372
Taylor D R F, 1988, "Cartography past, present and future" *New Zealand Cartographic Journal* **18** 9-20
Taylor P J, 1990, "Editorial comment GKS" *Political Geography Quarterly* **9** 211-212
Tikunov V S, 1986, "Thematic contents of maps: some theoretical approaches in the development of modelling methods" *The Cartographic Journal* **23** 56-65
Tikunov V S, 1987, "Some debatable issues in cartography" *Mapping Sciences and Remote Sensing* **24** 313-319
Tikunov V S, 1988, "Anamorphated cartographic images: historical outline and construction techniques" *Cartography* **17** 1-8
Tikunov V S, Yudin S A, 1987, "Use of transformed images in urban construction studies" *Mapping Sciences and Remote Sensing* **24** 202-208
Tobler W R, 1959, "Automation and cartography" *The Geographical Review* **49** 526-536
Tobler W R, 1961 *Map Transformation of Geographic Space* PhD thesis, Department of Geography, University of Washington, Seattle, WA
Tobler W R, 1963, "Geographic area and map projections" *The Geographical Review* **53** 59-78
Tobler W R, 1968, "Transformations", in *The Philosophy of Maps* Ed. J D Nystuen, DP 12, Michigan Inter-University Community of Mathematical Geographers, Wayne State University, Detroit, MI, pp 2-4
Tobler W R, 1973, "Choropleth maps without class intervals?" *Geographical Analysis* **3** 262-265
Tobler W R, 1979, "A transformational view of cartography" *The American Cartographer* **6** 101-106
Tobler W R, 1985, "Interactive construction of continuous cartograms", in *Auto-carto 7* Proceedings of the 7th International Symposium on Computer-Assisted Cartography, American Society for Photogrammetry and Remote Sensing and the American Congress on Surveying and Mapping, Bethesda, MD, Washington, DC, page 525
Tobler W R, 1987, "Experiments in migration mapping by computer" *The American Cartographer* **14** 155-163
Tobler W R, 1989, "Frame independent spatial analysis", in *The Accuracy of Spatial Databases* Eds M Goodchild, S Gopal (Taylor and Francis, Philadelphia, PA) pp 115-122
Tufte E R, 1983 *The Visual Display of Quantitative Information* (Graphics Press, Cheshire, CT)
Upson C, Faulhaber T, Kamins D, Laidlaw D, Schlegel D, Vroom J, Gurwitz R, van Dam A, 1989, "The application visualization system: a computational environment for scientific visualization" *IEEE Computer Graphics and Applications* **9** 30-42
Visvalingam, 1989, "Cartography, GIS and maps in perspective" *The Cartographic Journal* **26** 26-87
Winkler K H A, Chalmers J W, Hodson S W, Woodward P R, Zabusky N J, 1987, "A numerical laboratory" *Physics Today* (October) 28-37
Wolff R S, 1988, "Visualization in the eye of the scientist" *Computers in Physics* **2** 28-35
Wood D, 1985, review of *Graphics and Graphic Information-Processing* in *Cartographica* **22** 118-121
Wright J K (Ed.), 1938 *Notes on Statistical Mapping* (American Geographical Society and Population Association of America, New York)