

Occupational coding: principles and examples

Morris, Robert J.

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Morris, R. J. (1990). Occupational coding: principles and examples. *Historical Social Research*, 15(1), 3-29. <https://doi.org/10.12759/hsr.15.1990.1.3-29>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>

Occupational Coding: Principles and Examples

*Robert J. Morris**

Abstract: A clear statement of the principles upon which good practice for the coding and categorisation of occupational titles should be based is necessary for this increasingly used source. Coding at heart trades comprehensibility for loss of information. The nature of the inferences involved must remain clear. Each code must be based on one set of occupational titles and not incorporate other sources on an occasional basis. Each list of titles needs its own code. Comparison of the titles from a poll book (voting list) and a trade directory for the early 1830s in Leeds, in the north of England, indicated that the titles varied according to the function of the document. Codes should follow the rules of good social science practice as well as being responsive to the quality and context of the document and the purpose of the study. Guidance was derived from a variety of 19th century studies which ranged from an Owenite journal to Mayhew and Booth. Although comparability was rejected as a primary aim of category making, close attention was paid to the major coding systems in use over the past fifty years. The outcome was the multi dimensional code quoted here, suited to the study of a 19th century industrial town, but the principals behind this code had a wider applicability.

Hence all society would appear to arrange itself into four different classes:

- I. Those that will work
- II. Those that cannot work
- III. Those that will not work
- IV. Those that need not work

Under one or other section of this quadruple division, every' member, not only of our community, but of every other civilized State, must necessarily be included; the rich, the poor, the industrious, the idle, the

* Address all communications to Robert J. Morris, Dept. of Economic and Social History, William Robertson Building, George Square, Edinburgh EH8 9JY, Great Britain.

honest, the dishonest, the virtuous, and the vicious - each and all must be comprised therein.....

To arrange the several varieties of work into »orders«, and to group the manifold species of arts under a few comprehensive genera - so that the mind may grasp the whole at one effort - is a task of a most perplexing character. Henry Mayhew, *London Labour and the London Poor*, London 1861, vol.4, pp.3-4

Occupational titles are amongst the richest and most rewarding types of information used by historians. They have become increasingly important with the growing use of nominal list processing and prosopography. A variety of occupational codes have lain behind important advances in political history and in the understanding of the change over time in demographic processes and family and household structures.

Many of the problems related to the use of occupational titles have a long history in demography, sociology and other forms of social analysis. For historians the past ten or fifteen years has produced a wealth of experience. With the increased use of machine readable data that experience has intensified the need for a clear statement of the principles upon which good practice in occupational coding should be based. This need has intensified because of the increase in the quantity of data and titles which emerge from even the most modest of studies.

This paper sets out the principles upon which occupational codes for historical sources should be devised if historians are to maintain clarity and control as they deal with the large amounts of information which are now part of many projects. The principles outlined here are based upon the experience of a study of the middle classes of Leeds in the north of England during the 1830's. This was a major woollen textile, engineering and market centre for the West Riding of Yorkshire. (1) This experience is supplemented by an examination of the use of occupational coding by historians and other social scientists, as well as by an examination of the variety of principles upon which 19th century social analysts grouped occupations.

The Leeds Trade Directory of 1834 produced 2338 titles from 9131 entries. The 2% sample taken by from the 1851 census of Britain has produced 13,191 occupational titles from the coding of around 90,000 entries from the total 400,000 people in the sample. The final total for the sample is likely to be between 25,000 and 30,000. (2) Although these totals included quite small variations, such as different farm sizes attributed to those in agriculture, and variations in spelling and nomenclature describing the same job, the amount of information is still massive. Dr P J Corfield identified 2,000 separate occupations amongst the 30,000 individual entries taken from trade directories in British towns in the 1770s and 1780s. (3)

At its heart coding is a process which trades comprehensibility in return for the destruction of information. (4) The human mind for some reasons

prefers to work in threes (high, low, middle), can scan a table of up to 20 or so lines with effort but certainly does not regard a list of 2 or 3 thousand titles as anything more than an index. In the initial stage of any investigation a large number of categories should be retained. These should be coded so that subsequent merging will be easy, for example

farmers = 10
market gardeners = 11

From the start each title should have a unique code so that revision possible at all stages. Ideally, there should be no pre coding of data. In the examples discussed, the occupational titles were entered exactly as they appeared in the document and this text became the object of semi automatic machine coding from a directory of codes and titles. These principles of no pre coding and the potential for the reversibility of analytical mediation which it makes possible gives the historian a control of the relationship between analysis and evidence equivalent to the ability to go back to the details of a political speech or personal letter.

Because the power of machine readable nominal listings enables the historian to handle huge quantities of information, it is vital that the nature of the inferences involved remain clear and distinct. In a modern machine readable environment, quite modest studies can easily accumulate 20 or 30 thousand items of information. Thus the cumulative effect of unrecorded inferences may be substantial. If each interference is not separate then its effect cannot be checked and if necessary reversed. Thus if a code is a code of occupational titles, then only information from the occupational title should be used. For example the 'baker' may be an employer, small master, self employed or wage labour. The 'baker' may be engaged in production, distribution or both. Other information about an individual may enable an estimate to be made. Thus the presence of domestic servants or a high rateable value may imply that the individual is an employer. A rate book description of property as 'shop and bakehouse' implies engagement in both production and distribution. Such judgements are quite separate from occupational coding. Although it is clear that the 'baker' is engaged in food production, such a title can only be coded in a very general way in terms of status.

At this stage it would be wise to note that a major assumption lies behind most studies which use occupational titles, namely that the title gives reasonably accurate information about occupation. This is a classic concept indicator problem. (5) Indeed the careful and pedantic use of the phrase occupational title provides a warning that this assumption and the risks it carries lies behind much of the work done with occupational titles. One extreme and generalized form of distortion evident in 19th century sources concerned women's occupations. The occupational information in

the 19th century census of Great Britain has long been recognized to reflect a specific adult male concept of work as a full time activity devoted to a specified and limited range of activities and providing the vast majority of that individual's income. This was especially misleading when women's work was considered.

The clerks employed in classifying occupations in 1881 were given clear instructions

Those females only to be abstracted who are returned as following some specific occupation.

All males ... (except under fives and 'scholars') ... are to be abstracted, whether they are returned as with or without specific occupation.

The results of these attitudes showed clearly when it came to the agricultural section

A farmer's son, or close male relative, aged 15 or upwards, and living in the farmer's house, and not described as of any other specific Occupation, must be regarded as engaged in agriculture, and ticked to the Heading provided in order 7, sub order 1. But the corresponding female relatives are not to be ticked at all. (6)

The attribution of an occupational title was a value loaded exercise. The great range of part time, casual and home based work characteristic of many women was poorly recorded. Still less was attention given to work, again mainly female, which failed to enter the cash economy, as the reference to farming suggests. In general, the values concealed in the attribution of occupational titles were the dominant values of the society or institution which produced the document. Thus few prostitutes appear in the census and no pimp or brothel keeper amongst the lists of businesses in the directory.

Multiple occupations were common in the 18th and 19th centuries. These are poorly recorded in most nominal listings. It was a characteristic of many early entrepreneurs that they had multiple business and property interests. (7) Casual labourers frequently moved from one sector of the economy to another in search of work. Thus the people Mayhew wrote about in London might move from the docks in spring to the gas works in winter with a break for harvesting in late summer. (8) Some sources, like the trades directories did admit complexity and give more general clues as to the nature of multiple occupations. Others like the poll book and even the census were simplifications.

If the nature of the source itself is considered, then the issue becomes more one concerning the purpose for which the document was constructed than one of distortion. The nature and purpose of each source had a major influence on the way in which occupation was recorded. Although the Trade and Post Office Directories may have had some 'social register' function, they were in the main utilitarian documents. Individuals gave infor-

mation to the directory makers to ensure that they could be located in an increasingly complex world of business and commerce. (9) In the poll book the purpose of the occupational title was to ensure correct identity. It was a seeking for status and recognition. The comparison of occupational titles given for the same individual in different documents is an instructive and exercise. It is an exercise which gives a great deal of information about the nature of those documents, the nature of the society and economy which produced them and the historical methodologies now being applied to nominal list analysis.

As a result of the different purpose of the two documents, there was a divergence of occupational attribution. The nature and significance of this can best be explored through a particular example. If attention is restricted to the 2366 Leeds cases in which entries from the 1834 Directory and and 1834 Parliamentary Poll Book were successfully linked and which had occupational titles in both sources, then some idea of the consistency with which titles were attributed can be gained. In terms of occupational status, consistency was over 70% in most cases, with some important exceptions.

Table One; Comparison of the Occupational Status Codes in the Poll Book and Directory for Leeds 1834.

Occupational group	% of Poll Book titles found in equivalent category in the Directory	% of Directory titles found in equivalent category in the Poll Book
Agriculture	69	86
Retail and Processing	86	73
Dealers	41	73
Commerce	71	78
Banking and Finance	91	91
Agents and Travellers	58	52
Clerks and Bookkeepers	82	73
Manufacturing	65	69
Craft	75	80
Professional	84	70
Medical	96	94
Legal	98	93
Religion	100	100
Services	74	76
Construction	14	50
Independent Income	80	81

The comparison of production showed the same sort of levels of consistency if not slightly higher. An examination of the 'deviants' in terms of

the occupational status revealed a great deal about the nature of the documents and methodology being used. In a small number of cases the unique names logic which had been used for record linkage led to errors. It was technically possible that two identical and unique names could be present in each list referring to different people. A second technical cause of error arose when the occupational code of the poll book agreed with the second element of a double occupational code in the directory, but not with the first code with which it had been compared. These second code equivalents were not deviants in a real sense.

That left three major causes of deviation which arose from the nature of the documents and their economic and social context. In cases of status overlap, the source compilers disagreed over the description of cases which lay on status boundaries, as in the case of 'wholesale dealers' and 'merchants'. Secondly, there were genuine double and different occupations which arose from the multiple sources of income with which many individuals sought to counter the insecurities and fluctuations of the economy of Leeds. Finally, the imperfect division of labour evident in the economy created many deviations when an individual named one aspect of his business in one source and another in the other. Many shopkeepers and publicans had a stake in the manufacturing or craft section of the economy. It is worth examining some of the deviants in more detail. The professions were the most consistent. The small number of deviants which could not be attributed to technical factors suggested that the variation between the Directory and the Poll Book reflected a move from accurate description of job content to status seeking.

directory title	poll book title
glass and china dealer	glass merchant
foreign wool agent	woolstapler
solicitor	gentlemen
attorney's clerk	solicitor
accounts secretary	accountant
to the Savings Bank	

Builders had a very poor rate of equivalence because of the structure of the building trade. There were very few capitalist building employers. Most houses were built by building trade craftsmen who set up as an entrepreneur sub contracting work to other craftsmen. In the prestige conscious environment of the poll book there were 20 builders. In the Trade Directory they felt it was wiser to offer more accuracy concerning economic activity and the specific trade was more likely to appear. Of the 20 in the poll book, there were in the Directory

joiner and builder	10
joiner	3

bricklayer and builder	4
joiner and sawyer	1
joiner and timber dealer	1
shopkeeper	1

There were a large number of cases of increased status between the directory and poll book titles. 22 moved from distribution to commerce and 10 agents and travellers became merchants (commerce). All were essentially descriptions of the same job with different status claims. Imperfect division of labour also played a part. Eleven poll book 'merchants' admitted to getting their hands dirty in a variety of textile trades when it came to the Directory. Wine merchants became distillers and the glass merchant became a cut glass manufacturer.

Shopkeepers were readier to claim they were manufacturers in the Poll Book than in the Directory. Presumably the shopkeepers and ale dealers did not want to deter customers by admitting in the Directory that they were flax dressers, carpet manufactures, ironfounders or clothiers. Another large group concealed their shopkeeping activities from the poll book. There were 55 cases coded as craft in the poll book and distribution in the Directory. 24 of these were tailors who became drapers in the directory. Some were multiple occupations but here we find a value system very different from that of the high status middle class which concealed manufacturing in a status seeking environment. The craftsmen drew status from his skill and concealed the shopkeeping outside the utilitarian needs of the directory.

Two lessons can be drawn from this examination of equivalence and divergence between two sources of occupational titles. The purpose of the document is the likeliest guide to the risks being taken in making an inference about occupation from occupational title. Secondly, the question, which occupational title is right and which wrong is often not a useful question. Where the two occupational titles differ, they give different aspects of an individual's claims on the economy. They were different presentations of self to the world in which the document was created.

Aims and Principles

Several aims should be borne in mind when selecting and designing a code.

- The first must be the purpose of the project. In the example used here the intention was to study patterns of association and political and public action amongst the middle classes of the 1830s, thus the code made fine distinctions between merchants, manufacturers, craftsmen (makers), retailers and professional people. Studies interested in the

behaviour of the 'labour aristocracy' in Britain paid more attention to skill differentials and to white collar groups. (10)

- Second was the influence of the local economy. The woollen textile economy of the West Riding of Yorkshire provided a wide range of economic positions. Attention to this was a fruitful base for analysis. Small though important areas of the economy like printing and metal working received little sub division. In an economy like Edinburgh with little textiles and a large a varied printing and publishing sector the extent of sub division would be reversed.
- Third was the nature of the document. In the example used here, the trade directory and the parliamentary poll books were the major source of information. By their nature they excluded the bulk of the wage earning and manual labour population. Thus little attempt was made the differentiate different skill levels or the relationships of manual labour to capital.
- A code should ensure that it exploits the richness of the document. Thus the Directory often gave multiple occupations and a second code was included to represent this. Certain manuscripts of the 1851 census indicated whether an individual was master, journeyman or apprentice. In the 1891 census individuals were entered as employer, employee or self employed thus giving a valuable additional dimension to the code. Equally, no attempt must be made to invent information. Hence the need to take care with titles like 'baker'.
- This introduces another principle namely the need to respond to the specific economic and social context in terms of period as well as locality. In the 1830s, the imperfect division of labour created many problems. Thus many shopkeepers were still engaged in processing and production. The category retail in fact was fully expressed as 'those engaged in retail and associated processing and production activities'. Again merchanting and manufacturing activities were imperfectly divided and the mixed category was introduced. Anachronism is always a major danger. Some cases are easy to deal with. The category 'electrical engineering' provided by the Industrial Classification of 1971 can readily be dropped in 1832, but should 'surgeon' be allocated its current high status or be 'translated' as general practitioner. Was a 'teacher' wage labour, professional or white collar in the period before the Kay Shuttleworth reforms. (11) In the end a category of 'service and lesser professional occupations' was created to deal with this group.
- The context of the document is also important. Thus a 'baker' in a trade directory of the 1830s can be classified as engaged in 'retail and production' because we know that most entries in the trade directory refer to those engaged in businesses and that in the 1830s distribution and production were rarely separate. Note that this decision relates to the

context in which the title was found and not to the characteristics of the individual to whom the title was attributed.

- At a technical level many decisions were in part a response to the small numbers problem. It is useless to choose categories which in analysis produce cell totals so small that there would be a high probability that the scores and their variations were a matter of chance. In general, categories which regularly produced cell sizes below 30 were suspect. Those which never produced cell sizes over 10 were useless. Thus in the initial codes used for the Trade Directory an attempt was made to provide categories for white collar workers (clerks), manual labourers (there were 12 out of 9101 entries) and government and local authority employees (0.1% of the total). Such an attempt could easily be justified in terms of the aims of the project but retaining them in analysis was usually pointless.
- Lastly, there is an issue which too often dominates coding decisions, namely that of comparability. In Britain, the choice of Booth-Armstrong (12) or of a modified OPCS Occupational Code has often been made because it is there and seems to offer a ready means of comparison. Such a decision often introduces rigidities, an inability to serve the detailed needs of the project. The apparent consistency of using the same code may conceal differences of judgement that derive from the nature of the document or economic context. Comparability is a desirable and welcome aspect of any analysis. It should not be purchased at any price. It should be entered into with a full awareness of the risks involved.

As the categories themselves were produced, they were designed to observe the basic principles of any social science coding.

- Only one dimension or type of information should be dealt with at any one time. The failure to do this causes more confusion in reading occupational tables than any other difficulty, hence the need for multi dimensional coding which is the central assertion of this paper. (13)
- They were exclusive. In other words no title should appear in more than one category. In many codes the failure to observe this derives from a failure to observe the principle of handling one sort of information at a time. Thus codes which included 'retailers' and 'drink' or 'food' in the same dimension are a frequent source of confusion.
- Codes should be inclusive. Thus every title should have a place even if only in that catch all category 'others'. In the Leeds directory this left a number of awkward categories on the edge of the analysis, such as quarry owners, and transport. The nature of the poll book entailed the creation of a category 'general manufacturing' because of the large number of individuals called 'manufacturers' or 'millowners'. Alt-

though the context of the Leeds economy implied that the bulk of this group were in textiles it seemed wise to keep them separate for analytical purposes. The need for inclusivity also created a major category of 'rentier' income earners which although not an occupation was a distinctive form of relationship to the economy frequently referred to by the directory.

Guidance

A great deal of guidance can be provided by contemporaries. The earliest schemes of occupational coding were used by the 18th century national income statisticians, King, Massie and Colquhoun. They were interested in economic power and its distribution and accumulation. Thus their schemes give especially valuable in terms of the middle and landowning classes. In the example given here Colquhoun does substantial justice to the wide variety of relationships to the economy held by those whose income derived from rent, profit and taxes, the division between merchant and manufacturing capital, the separation of 'inland traders' from others. The recognition that 'Merchants and bankers' derived their income from land and the funds as well as trade are all distinctions which have shown recent analytical value. (14) The scheme was poor regarding wage labour. The very large class 14 lumped 5.2 million people together. Colquhoun did recognize the distinction between labourers and paupers, as well as the distinctive relationships of criminals, debtors and lunatics in asylums to the economy and the state, but once the 'labouring poor' were safely encapsulated within the market economy (class 14) he lost interest in the enormous variety of experience within that group. His original calculations were based upon 47 groups which he reduced to 21 for his summary tables. The 21 categories were

1. Sovereign and Family.
2. Peers, Country Gentlemen, Freeholders of lands and houses, mines, minerals, funds and public incomes.
3. Persons with Colonial and East India property, funds etc., including foreign incomes.
4. Merchants and Bankers deriving income from trade, funds and land.
5. Ship Owners deriving income from freights and other property.
6. Manufacturers of all kinds including ships, houses and works.
7. Inland Traders, Shopkeepers including publicans of all kinds, trading on capitals.
8. Agriculture, Farmers, Graziers, Dealers in cattle.
9. Established and Dissenting Clergy.
10. Liberal Professions - law, physic, literary and fine arts.

11. Persons employed in education, including the universities.
12. Persons employed in theatre and music.
13. Civil and Military labourers for the state and its defence.
14. Labourers in agriculture, manufacturers, commerce, navigation and fisheries.
15. Hawkers and Pedlars.
16. Persons employing capital in lunatic asylums.
17. Lunatics supported in asylums.
18. Persons in prison for Debt.
19. Vagrants, Gipsies, common prostitutes and criminals in and out of prison.
20. Paupers, including those who get supplements in aid of wages.
21. Persons included in the above who have money in the funds for selves, widows, orphans or minors or charities.

Gregory King's account of English society has been criticized in recent years for underestimating income levels especially amongst the wealthy, and for underestimating the contribution of trade and industry to the national income. It remains a value laden traditionalist perception of English society at the end of the 17th century. (15)

By the 1830s, the growing number of enquiries into pauperism and public health, the move of the statistical movement and medical topography into quantification and the critique of contemporary society by a range of radical movements had produced a variety of occupational classifications. The most characteristic was that used by Chadwick in 1842 report.

1. Gentlemen and Persons engaged in Professions.
2. Tradesmen, Farmers, Graziers and Shopkeepers.
3. Operatives, Mechanics, Servants and Labourers.

Chadwick provided no formal discussion of the nature of his three social classes. It was clear from the text that he defined the three groups by their huge differentials in property, income and education which enabled them to command the better housing, sanitation, cleanliness, medical, clothing and food which ensured the equally substantial differences in biological survival. The scheme was used by Dr W A Guy in an 'Abstract of the Professions and Occupations of 4312 males of the age of 15 years and upwards who died of cholera in London during the epidemic of 1848-49'. (16) Some occupational titles had clear locations, architects (class 1), grocers (2) and dustmen and scavengers (3), but the bakers and butchers had considerable ambiguity. Dr Guy used the London Post Office Directory to help him solve this problem and to help him calculate the probability of individuals from each group appearing in the cholera death lists.

This ratio (of deaths to the living for each occupational group) is obtained, in the case of tradesmen, by dividing the number following each

trade (as given in the Post Office Directory for 1840) by the number of deaths. In the case of such of the working class as follow occupations of the same name with that borne by tradesmen, the number of the living is obtained by subtracting the number of tradesmen from the total given in the census, such total comprising both the employers and the employed.'

Two lessons come from these schemes; the contemporary perception of a fundamental division within the middle classes that was recognized in Gibbon Wakefield's phrase the uneasy classes, (17) and a contemporary confidence in a good trade or post office directory as an indicator of the division between masters, entrepreneurs or profit takers and wage takers where titles themselves had ambiguity.

The scheme was the ancestor of the very successful scheme devised by the registrar general in the early 20th century to relate status and life style positions to demographic fortunes. The Registrar General Five Classes (plus 2) were very successful for the purpose for which they were devised, although they were less clear when applied to other forms of social behaviour. They have been subject to much criticism lately because of the lack of homogeneity of behaviour within each class. Such criticism misses some of the nature of a code as a trade between comprehensibility and loss of information, and should perhaps concentrate in assessing how far the powerful increase in the ability to comprehend demographic and other forms of behaviour was worth the undoubted loss of information on the variety of behaviour within each major group. (18) The long and uneasy life of this classification may well be extended simply by its ability to provided long run comparisons of the relationship between status and other demographic and social variables. (19) Although it echoes Chadwick's first two classes in Classes I and II, its inability to make distinctions between the variety of middle class positions makes it unattractive for many historical studies. (20) In any case, the work involved in adapting this code to pre 1900 historical situations would be greater and involve more risks, than devising the purpose directed, context sensitive codes advocated here. The early codes relating to production or product were of very little value. They incorporated some odd theories. (21) They incorporated hidden assumptions about the nature of an individuals occupation as being directed to a well defined area of the economy. This causes especial confusion over titles like 'servant'. (22) Almost always 'production' is best handled by modifications of recent SIC codes with the addition of some 'general' categories

The nineteenth century codes then give substantial guidance on status and on relationships to resources brought to the market (including professional status and aristocratic prestige). Because status is central to these codes, their value laden nature is especially important. This was shown most clearly by the nature of the classification used by the Owenite radicals

'FIRST CLASS the labouring population, the producers of all wealth
SECOND CLASS distributors, superintendents and manufacturers; necessary but too numerous" - included farmers, capitalists, merchants, bankers, shopkeepers, clerks, hawkers and physicians

THIRD CLASS government. Much too numerous and expensive' - included the royal family, judges, the military services, paupers, lunatics and those imprisoned for debt

'FOURTH CLASS instruction and amusement, indispensable and eminently useful' - included clergymen, education as well as theatres and concerts

'FIFTH CLASS the most wealthy and least useful' - included nobility, bishops, landowners and fundholders. (23)

This is not a classification that many historians would want to use, gathering together royalty, the military and lunatics in one class and clergymen and actors in another. It is a valuable reminder of the manner in which values and perceptions create classifications. This status hierarchy was based upon the labour theory of value. Fundamental was the merging of wage earners and many middle class occupations as the 'producers of all the wealth'. As will be clear from other schemes contemporary perceptions of occupational positions did not and sometimes could not make a clear distinction between middle and working class, between profit taking and wage taking positions in the economy.

One of the most perceptive schemes produced during the 19th century was devised by Henry Mayhew as a result of his study of London. It is little used because it is tucked in at the start of volume four along with the prostitutes and criminals. (24) In its published form it was impossibly chaotic. In a study which stands on the edge of social science, Mayhew showed that he clearly understood the needs of occupational classification. The categories devised for the Great Exhibition of 1851 were rejected as 'neither distinct nor do they include the whole', and for 'the confounding of processes with products'. Mayhew began with the four classes quoted at the start of this article. There was a clear acknowledgement of the labour theory of value and some reference to the political economists like J S Mill. As he elaborated his code it demonstrated three aspects of value to historians. He attempted to come to terms with service occupations and made a division between those which created human capital and those which served immediate needs. He examined the variety of relationships of labour to capital and went far beyond the skill hierarchies that form the basis of subsequent codes. Lastly he outlined some of the variety of forms of business organization and relationships to capital ranging from the Joint Stock Company to the Penny Capitalist. Those who will work were divided into

- I The **Enrichers. The collectors, extractors or producers of exchangeable** commodities.
- II **Auxiliaries.** The Promoters and Distributors of the produce.
- III **Benefactors.** The producers of human **capital.**
- IV Servitors. Render temporary service, '... amusers, protectors and servants.

The classification of the enrichers was multi dimensional. The dimensions were

- Nature of the product.
- The mode by which the operatives are paid.
- The places at which they work.
- Those who employ them.
- Those they themselves employ.
- Skill levels reflected in pay.

The modes by which the operatives were paid were

- Day Workers.
- Piece Workers.
- 'Lump' or contract workers.
- Perquisit workers, as waiters.
- 'Kind' or truck workers, as north of England farm servants.
- Tenant Workers, who lodge with or reside in houses belonging to their employers.
- Improvement workers, like apprentices and learners who are rewarded by the instruction they get whilst working.
- Tribute Workers, where a share of the proceeds goes to the workmen as Cornish Miners, whalers.

The operatives themselves could act as employers in three major ways.

- Family Workers could be hired.
- Sweaters or piece master workers employed labour at under the standard rates.
- Garret Masters employed mainly apprentice labour

Mayhew provided considerable guidance for those who seek to organize occupational information about business and other activities which depended for their income in authority over land and capital. Most made their appearance under Auxiliaries and under 'those who need not work'. The categories were often confused and not worth quoting in full, but extracts from these lists provide major guidance for those concerned with the middle classes and the margins of the middle class.

- Administrative Employers, they supply wholesale and retail dealers.
- Executive Employers, work directly for the public as builders.

Distributive Employers, retail what they produce as tailors,
shoemakers and eating house keepers
Middlemen Employers, sub contractors

Distributors of Production

Dealers engaged in buying and selling on their own account (these
were again subdivided into merchants (related to overseas trade),
wholesalers, retails and contract purveyors

Agents

Lenders and letters out of commodities

property including housing of money
(from bankers to pawnbrokers)

Carriers

Amongst Those who need not work were

Landlords

Fundholders

Shareholders

Annuitants

Mayhew's perceptions were influenced by the London economy and by his concern for the 'moral economy' of standard wage rates which was under considerable pressure from 'sweaters' and 'cutting employers'. (25) There were major gaps in his scheme but his awareness of the different types of relationships and authority patterns provides substance for many distinctions which historians need to make amongst the middle classes and provides a well documented warning of the blurring of the dividing line between capital and labour.

The last scheme which needs consideration was used by Charles Booth in his study of London in 1888. His eight class scale recognized the regularity as well as the level of earnings and also related this to consumption and cultural patterns such as servant keeping and drinking.

- H Upper Middle Class, 'the servant keeping class'.
- G Lower Middle Class, 'shopkeepers, small employers, clerks and subordinate professional men a hard working, sober and energetic class'.
- F Higher Class Labour, 'the best paid of the artizans together with those of equal means'. At between 30/- and 50/- a week. The group included foremen, first hand lightermen, trusted employees who sometimes shared in profits. Their sons were clerks. Their daughters served in the better shops. If the wives worked they kept a shop or a laundry.
- E Regular Standard Earnings. Most artizans. The better street

- sellers and dealers, a large part of the small shopkeepers. Some small employers.
- D Small regular Earnings. Poor. Never over 21/a week. Dock and Gas works labourers. Children worked if possible. Little actual want unless the wife drinks.
 - C Intermittent Earnings. 'the victims of competition' Stevedores, waterside porters; known for 'improvidence'.
 - B Casual Earnings. Very Poor. ' ... is not one in which men are born and live and die, so much as a deposit of those who from mental, moral and physician reasons are incapable of better work.'
 - A The Lowest Class of Occasional Labourers, Loafers and Semi Criminals with 'little regular family life amongst them.'

The value judgements built into this scheme were clear as was the merging of several dimensions, but it was an excellent example of a purpose designed scheme which fulfilled its aim with great success in Booth's study of poverty. (26) This survey of contemporaries attempts to understand the occupational structures they lived in confirms the richness and variety of information carried by an occupational title. Sociologists and historians have attempted to deal with this in a variety of ways. Government Statistical Services have provided British historians and social scientists with two important dimensions for industrial coding. The occupational code related to the nature of the work task, whilst the Standard Industrial Classification was determined by the nature of the output, (see below) The most popular coding system amongst British historians, known as the Booth-Armstrong code, is at its best as an industrial code. The status dimension was organized into an equivalent of the Registrar General Five 'social classes' with the rigidities and limitations that imposed. (27) Other historians have favoured extended multi dimensional codes. Anderson working with the 1851 census devised a multi dimensional code, based upon an explicitly Weberian scheme of stratification, sensitive to the context and content of the document and containing some elements which made comparative work easier. Thus the major dimensions of the code were i) economic activity(retired, employed), ii) personal resources brought to the market (skill, professional training), iii) place in a hierarchy of control iv) information on partners, employees, acreage where appropriate, and v) a unique code which enabled the title to be linked to the Booth Armstrong scheme, the Standard Industrial Classification of 1970 and the census classification of 1881. (28) Sociologists with access to questionnaire responses have been able to develop more elaborate status based codes. The bulk of these are based upon a 'reputational approach', by asking respondents to rank occupations according to perceived status. One of the most recent has taken a 'relational approach', using 'patterns of asso-

ciation in non work situations', thus relating status more clearly to life style and the social interactions related to that style. As historical data bases become more complex such studies based upon the behavioural evidence available to historians will become possible but this is some way into the future. (29)

Outcome

The richness and complexity of the information which can be derived from an occupational is so great that any thorough, extensive and consistent analysis using such titles must involve a multi-dimensional code which keeps distinct the several separate types of information which might be derived and does not attribute or anticipate any hierarchy within these types. The two major dimensions involve the nature of the work done and the nature of the product. Thus the woolstapler, the cotton merchant and the grain merchant will share many experiences and interests in common in matters of accounting, relationships to labour, concerns over finance and enforcing contracts. In other matters the grain farmer, his labourers and the merchant will be the relevant group for analysis and explanation. The two major dimensions are best described in the recent words of Office of Population Censuses and Surveys.

The occupation of a person is the kind of work which he or she performs, due regard being paid to the condition under which it is performed ... the nature of the factory, business, or service in which the person is employed has no bearing upon the classification of his occupation

The Industry in which an individual is engaged is determined (whatever may be his occupation) by reference to the business or economic activity in which his occupation is followed ... the (industrial classification) has regard only to the nature of the service or product to which his labour contributes. (30)

status or resources brought to the market (skill, professional
economic activity status
compatibility (anachronism; info demands

The coding itself was done by semi mechanical means. All occupational titles were extracted from the machine readable version of the two documents concerned and the frequency of each noted. Each title was given a unique code, although trivial differences were amalgamated thus

Music Teacher
Teacher of Music
Music Professor
Music Preceptor

were all given the same code. The file of occupational titles was then annotated line by line and the resulting file used for the coding of the individual entries in the document files. Thus the file for the Parliamentary Poll Book began as follows

139	50	503	215	Clothier
309	30	112	208	Innkeeper
228	10	100	132	Farmer
378	45	810	132	Merchant
74	30	131	129	Butcher
487	30	991	122	Shopkeeper

the codes being, a unique identifier, an occupational code, an industrial code and the frequency of the title in the source. The Trade Directory gave a fuller description of occupations. For this source a second set of codes was provided for dual occupations. Thus Joiner and Builder became

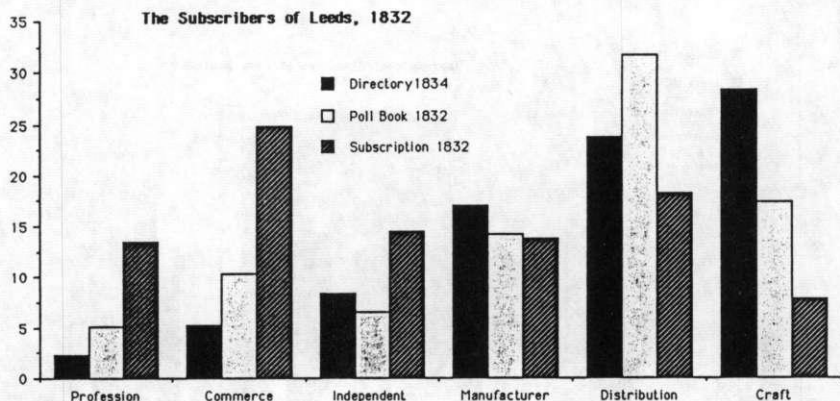
55	651	70	723
----	-----	----	-----

In other words, craft, wood; building entrepreneur, building general. Note that the "building entrepreneur" has been retained as a separate type of relationship to capital because of the distinctive structure of capital in an industry dominated by many small and often temporary units of entrepreneurship. (31)

The purpose of this coding was to enable the 'occupational titles' to be grouped into categories. The codes as such attributed a variety of qualities to the individual titles. The categories which were derived from these codes through their ability to group and summarize qualities in a variety of ways provide the analytical leverage. The flexibility of the computer and its ability to perform repetitive operations with considerable ease gave a flexibility and variety to categorization lacking in most previous hand driven methodologies. At the end of this process any code is only as good as its ability to organize the raw material in such a way that useful historical conclusions

- ability to discriminate

The subscription was gathered after a public meeting called during the winter of 1831-32 to consider the distress amongst the unemployed poor threatened by the cholera epidemic of that year. The result of the enquiry made possible by this code that although the manufacturers dominated the directory population and the shopkeeper (dealer-producers) dominated the poll book population, it was the commercial and professional men who dominated the public action of the subscriptions. This was part of a larger demonstration of the importance of this group in the creation of middle class public culture during the 1830s.

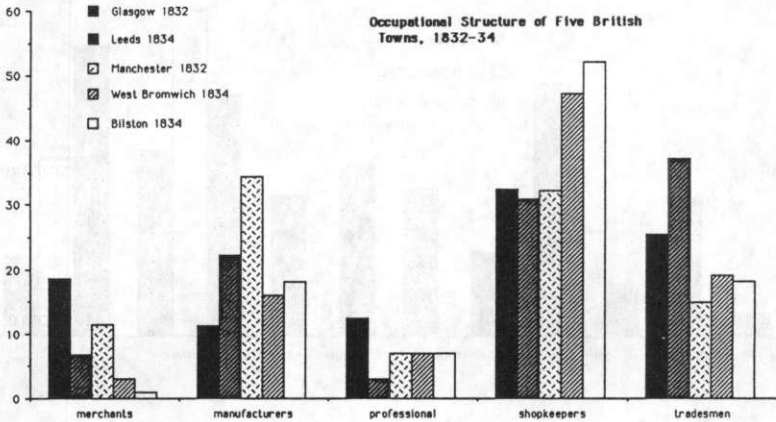


- comparability

These figures were derived from four thesis studies made in Britain over the past 15 years which used context sensitive purpose directed coding to study various aspects of middle class and elite behaviour. (32) The comparison is enough to show broad orders of magnitude and to make comparisons between the three regional 'capitals' and the two manufacturing towns. A full comparison would involve a critique of the coding decisions and document sources of each of the four studies. These would probably not alter the basic conclusions possible from this well based coding.

- The code could be tested against other measures of the concepts it was believed to indicate. In this case status was measured by the mean value of subscriptions (in £s decimal) to a fund for the relief of the poor and showed clearly that distribution and commerce on the one hand and manufacturing and crafts (or tradesmen) were discriminating successfully.

distribution	1.89
commerce	5.06
manufacturing	3.19
craft	2.07
professions	4.06
independent means	4.82



These tests of internal consistency showed clearly that subscriptions and the status based occupational code would measuring related concepts.

Conclusions

There are two major conclusions to this paper. First, occupational titles, codes and classifications were and are used for a purpose. This purpose affected the titles attributed by the variety of documents which historians may use. Thus an individual may be attributed several titles during an investigation as information is gleaned from different documents. Each title is a valid source of information about that individual. Likewise the codes and classifications used, both in the past and by historians reveal purpose and must be chosen in the light of purpose. From this follows the second set of conclusions, coding must be multidimensional and flexible so that the resulting classifications can respond to the historians variety of needs during the course of an historical enquiry.

Appendix

A two dimensional code was selected for the Leeds study. The major headings are listed here. Occupational Status Although this code was based upon the principles of the OPCS Occupational Classification no exact equivalence was possible because of the quality of information available in each source and because of the greater imperfection in the division of labour existing in the economy of 1830. The development of these principles resulted in a code which reflected the type of work being done, but also, as this was a study and a source which emphasized middle class situations, it was a code which reflected the relationship to resources brought into production such as property and professional skills. Thus this code reflected type of work, resources brought to the market and economic status. In the form presented here it was especially suited to sources and problems which gave attention to the middle class ends of the social hierarchy.

10. Agriculture

This combined all productive activities in which land was the major factor of production.

11. Gardener

Land a major factor but intensive use.

20. Quarries

All extractive processes.

30. Distribution and Processing

This takes into account the imperfect division of labour for those who combined retailing and production, or as in many food trades the processing of semi finished goods. (33)

31. Dealer

All those in distribution where the title implied no processing. 'Dealer' and 'Factor' were keywords here. In practice this was hard to separate from 30 and 30 and 31 were almost always amalgamated in analysis.

40. Transport

45. Commerce

This group involved large units of circulating or finance capital; 'merchant' was the main keyword here, as was 'warehouse'. Care was taken to exclude the 'rag and bone merchants'; the slavish use of keywords is never wise.

46. Bankers

Involved in finance rather than commodity trading; almost always amalgamated with 45.

47. Agents and Travellers

Selling on behalf of others.

48. Clerks and Bookkeepers
Involved in the administration of the capital of others; in the early 19th century clerk often implied 'manager' rather than white collar worker in the twentieth century sense.
50. Manufacturing
Using substantial fixed and circulating capital to produce goods. Direct employers of labour.
55. Craft
Small units of capital used in the production of goods; self employed or employers of small amounts of labour; often involved in an apprentice, journeymen, master life cycle pattern of social mobility.
59. Manufacturing and Commerce
A double category to take in the imperfect division of labour.
60. Professions
The older professions were central to this group. They involved the provision of knowledge and skill based services after long periods of training and heavy entry costs which were the basis of their social prestige. Most had some form of legally sanctioned organization which controlled entry, working practices and standards. New occupations which shared these characteristics were added to this group.
61. Medical
62. Legal
63. Religion
These three (61-63) were usually combined with 60.
65. Miscellaneous Services
Provider of a variety of services with low prestige or entry barriers. It may well be that this group should be sub divided to separate education.
70. Construction
There were a small group of building entrepreneurs. Note that 'craft' takes precedence over this division for people like glaziers and masons.
90. Independent Income
95. National Government
96. Local Government
97. Defence
98. Foreign Government
There was an American Consul in Leeds.
99. **No Occupational Title**

Production

This code included 28 categories. Not all of which were usable because of small number problems. Although it was based upon the principles of the Standard Industrial Classification of 1968, the industrial classification of 1931 proved a better starting point as it contained fewer anachronisms and had been used in studies of Leeds and Hull. (34)

Production, 1834	Ind. Classification, 1931	SIC, 1968
1. Farming	1. Agriculture	1. Agriculture, forestry, and fishing
2. Food, Drink and Tobacco		2. Fishing
3. Mines and Quarries	3. Mines and Quarries	2 Mines and Quarries
4. Glass and Pottery	4. Bricks, Glass and Pottery	14. Bricks, Glass and Pottery
5. Chemicals	5. Chemicals and Oils	5. Chemicals and allied Industries
6. Leather and Leather Goods	9. Skins and Leather	4. Coal and Petrol
7. Metals and Metal Goods	6. Engineering	14. Leather, Leather Goods and Furs
	7. Precious Metals	6. Metal Manufacture
		7. Mechanical Engineering
		8. Instrument Engineering
		9. Electrical Engineering
		10. Shipbuilding.
		11. Vehicles
		12. Other Metal Goods
8. Non Specific Manufacturing		
9. Textiles	8. Textiles	13. Textiles
10. Clothing and Footware (see 1.)	10. Dress	15. Clothing and Footware
11 Timber	11. Food, Drink, Tobacco & Lodging	11. Food, Drink, Tobacco & Lodging
	12. Woodworking	17. Timber and Furniture
12. Paper, printing and Publishing	13. Paper, Books and Printing	18. Paper, Printing and Publishing.

13. Construction and Repairs	14. Building	20. Construction
14. Transport	16. Transport	22. Transport and Communications
15. Business and Financial Services including Management	17. Commerce	23. Distributive Trades
16. Public Admin., Public Utilities and Defence	18. Local and National Government	24. Insurance, Banking, Finance and other Business Services.
19. Defence	19. Defence	27. Public Admin and Defence
20. to 23 Professional and Scientific Services	15. Gas, Water, Electricity, Sanitary	21. Gas, Water, and Electricity
23. and 24. Misc. Services including education	20. Professional Services	25. Professional and Scientific Services
26. General Distribution	21. Domestic and other Services	26. Misc Services
998. Independent Income		
999. No Occupational Title		

Although some categories have identical name this does not mean that they included the same titles for both types of analysis. Thus, the 'clerk in a bank' would be 48 (clerks and bookkeepers) for organization of work but 15 (Business and Financial Services) for production, and the 'army surgeon' would be 61 (medical profession) for organization of work and 18 (Defence) for production.

References

- (1) This work was supported by a personal research grant from the ESRC/SSRC. See R J Morris, *The Leeds Middle Class, 1820-1850*, End of Grant report to the SSRC, Grant number B/00/24/0003/1 now deposited in the British Library Lending Division, Boston Spa.
- (2) Information kindly supplied by Professor M Anderson, Dept of Economic History, Edinburgh University.
- (3) P J Corfield, *Computerizing Urban Occupations*, in Peter Denley and Deian Hopkin (eds.), *History and Computing*, Manchester 1987, pp.68-72.
- (4) These comments on coding are set in their context of nominal record linkage in my contributing to the Proceedings of the Amsterdam conference on History and Computing held in 1989. These will short-

- ly be published edited by Roel van de Voort and other.
- (5) Other examples of this problem are discussed in my contribution to the Proceedings of the 1989 Amsterdam conference. See above.
 - (6) *Census of England and Wales, 1881. Instructions to the Clerks Employed in Classifying the Occupations and Ages of the People*, London 1881, pp.1 and 4. I am grateful to Professor M Anderson for showing me this document.
 - (7) F Crouzet, *The First Industrialists. The Problem of Origins*, Cambridge 1985
 - (8) G Stedman Jones, *Outcaste London. A Study of the Relationship between Classes in Victorian Society*, Oxford 1971
 - (9) Jane E Norton, *Guide to the National and provincial Directories of England and Wales ... before 1856*, London 1950, p.8 gives the example of Directories being used by commercial travellers.
 - (10) G Crossick, *An Artizan Elite in Victorian Society, Kentish London, 1840-1880*, London 1978 ; R Gray, *The Labour Aristocracy in Victorian Edinburgh*, Oxford. 1976
 - (11) Asher Tropp, *The School Teachers: the growth of the teaching profession in England and Wales*, London 1957.
 - (12) W A Armstrong, *The use of Information about occupation. Part Two. An Industrial Classification, 1841- 1891*, in E A Wrigley (ed.), *Nineteenth Century Society*, Cambridge 1972 pp.226-310.
 - (13) R J Morris, Property Titles and the use of British urban poll books for social analysis, *Urban History Yearbook*, 1983, pp.29-39; R J Morris, *The Leeds Middle Class, 1820-1850*, End of Grant Report.
 - (14) R G Wilson, *Gentlemen Merchants. The merchant community in Leeds, 1700-1830*, Manchester 1971, gave especial emphasis to the importance of land in merchant assets; Julian Hoppit, *Risk and Failure in English Business, 1700-1800*, Cambridge 1987, found the distinction between inland and overseas traders especially important.
 - (15) G S Holmes, Gregory King and the Social Structure of Pre Industrial England, *Transactions of the Royal Historical Society*, vol.27, 1977, pp.41-68; P J Corfield, Class by Name and Number in Eighteenth Century Britain, *History*, vol.72, 1987, pp. 38-61 surveys the development of socio-economic classifications over the 18th century.
 - (16) *Report of the General Board of Health on the Epidemic Cholera of 1848 and 1849*, appendix B, pp.182- 6, PP, 1850, vol.21.
 - (17) R S Neale, *Class and Ideology in the Nineteenth Century*, London 1972 and R J Morris, *Class and class consciousness in the Industrial Revolution, 1780-1850*, London 1979.
 - (18) J A Banks, *Victorian Values. Secularism and the size of families*. London 1981, p.41-42.
 - (19) Central Statistical Office, *Social Commentary: Social Class, Social*

- Trends, vol.6, 1975
- (20) Studies of origins of manufacturers are often limited in their analytical power by restricting themselves to this scale; Katrina Honeyman, *Origins of Enterprise. Business Leadership in the Industrial Revolution*, Manchester 1982; F Crouzet, *First Industrialists*. are both restricted by this choice.
 - (21) S Szepter, The genesis of the registrar General's Social Classification of Occupations, *British Journal of Sociology*, vol.35, 1984.
 - (22) E Higgs, Domestic Servants and households in Victorian England, *Social History*, vol.8, 1983, pp.203- 210.
 - (23) *The Crisis*, 17 May 1834, quoted by P Hollis, *Class and Class Conflict in 19th century England*, London 1973, p.6.
 - (24) Henry Mayhew, *London Labour and the London Poor*, 4 vols, 1861-62, reprinted New York 1968.
 - (25) E P Thompson and Eileen Yeo, *The Unknown Mayhew, selections from the Morning Chronicle, 1849 - 1850*, London 1971, introduction.
 - (26) Charles Booth, *Life and Labour of the People of London*, London 9 vols, 1892-97, vol. one, pp.32-51.
 - (27) *Office of Population Censuses and Surveys, Classification of Occupations, 1970*, HMSO London, 1970, p.v; W A Armstrong, op.cit.
 - (28) M Anderson, B. Collins and C Stott, National Sample from the 1851 Census of Great Britain, Background paper VII: Data Set Documentation, Edinburgh University, July 1977; at the same time Theodore Hersberg was also using a multi dimensional code for US census tracts, see T Hersberg (ed.), *Philadelphia: Work, Space, Family and Groups Experience in the Nineteenth Century. Essays towards an Interdisciplinary History of the City*, New York and Oxford 1971 and Theodore Hersberg, The Philadelphia Social History Project, a methodological history, Stanford PhD, 1973.
 - (29) A Stewart, K Prandy and R M Blackburn, *Social Stratification and Occupations*, London 1986; J H Goldthorpe and Keith Hope, *The Social Grading of Occupations*, Oxford 1974.
 - (30) Office of Population Censuses and Surveys, op.cit.
 - (31) Peter J Aspinall, The Size Structure of the House Building Industry in Victorian Sheffield, Centre for urban and Regional Studies, University of Birmingham, Working Paper 49, March 1977.
 - (32) R H Trainor, Authority and Social Structure in an Industrialized Area: a study of three Black Country Towns, 1840-1890, Oxford D Phil 1981; Stana Nenadic, The Structure, Values and Influence of the Scottish Urban Middle Class: Glasgow, 1800-1870, Glasgow PhD 1986; V A C Gatrell, Incorporation and the pursuit of Liberal Hegemony in Manchester, 1790-1839, in Derek Fraser (éd.), *Municipal*

Reform and the Industrial City, Leicester 1982.

- (33) D Alexander, *Retailing in England during the Industrial Revolution*, London 1970.
- (34) W G Rimmer, *Occupations in Leeds, 1841-1951*, *Publications of the Thoresby Society*, no. 113, 1967; Joyce Bellamy, *Occupations in Kingston upon Hull, 1841-1948*, *Yorkshire Bulletin of Economic and Social Research*, vol.4, 1952, pp.33-50.