The Use of Graph Theory in Modelling

Thematic Structure in the Content of Documents

Barbara Astrid Farbey
Doctor of Philosophy
The City University
Department of Systems Science
September. 1984.

CONTENTS

TABLES

ILLUSTRATIONS

ACKNOWLEDGEMENTS

This thesis grew out of a research project funded by the Social Science Research Council (Grant No HR5958). The work was carried out in the Department of Systems Science at the City University and I should like to thank my colleagues in the Department for their support and interest.

I am particularly grateful to [redacted] for guidance through the substantive area of the research and to [redacted] for their unwearying criticism and help with the manuscript and throughout the research.

ABSTRACT

In this thesis a graph theoretic model of thematic structure in the content of written documents has been examined.

The graph model is derived from the documents using a form of Content Analysis to define the major themes contained in the documents and a relation between those themes. The themes are identified with the vertices of a graph and the relation defines the edges.

The resulting graphs can be studied in terms of their invariants to provide quantitative measures of thematic structure. These measures are used to describe and compare the structure of ideas in groups of documents.

The model has been tested empirically using a substantial body of documents of the Scottish National Party over forty years and a smaller number of Scottish Conservative Party documents.

In this analysis the model proved to be direct and substantively plausible. It was found to be immediately useful in generating a new kind of structural hypothesis and in forcing clarification of concepts such as 'centrality' and 'complexity'. Individual measures, such as eccentricity were sensitive, others, such as rank degree, more robust. The usefulness of the measures was limited by a lack of substantive theory.

The assignment of undated documents was attempted. Longer documents could be accurately placed in the correct period, shorter ones were difficult to assign.

Extensions to the model towards a multi-level, full relational structure have been investigated. These would not be practicable without further computerisation. An outline architecture for a computerised model has been proposed.

The theoretical implications of a structural model for a system of ideas has been examined and the relation of this work to similar work on manifest idea-systems in cognate disciplines has been noted. The model is likely to be most useful in the empirical study of such systems.

For David and our daughters

1. The problem and its context

1.1 Introduction

The intention of this thesis has been to provide a structural model of the content of written documents which would enable quantitative statements to be made about that content. Such statements are necessary in order

i) to be able to describe the structure of the content of documents precisely and concisely

ii) to be able to compare the structure of such content from different sources in a systematic way

iii) to be able to describe structural change in a series of documents

These abilities, in their turn, provide a basis for the effective characterisation of communication which can be used, say, in the recognition and assignment of documents from an unknown source, or in the construction and use of a knowledge base. They also provide a basis for structural hypotheses, that is hypotheses about the nature of structural change in systems of ideas, in so far as (and only so far as) these systems are articulated in collections of documents.

1.2 The model

The model which has been studied is a graph (Harary 1972). Themes abstracted from the content of a group of documents were identified with the vertices of a graph and a single relation between the themes (co-mention in a paragraph) defined the edges. The use of a graph theoretic model was directly evoked by Converse's concept of an ideology as being "what goes with what" (Converse 1964) and has echoes in the Semantic Networks

of Artificial Intelligence and Psychology (Simmons 1973, Abelson 1973, Schank 1973, Waterman and Hayes-Roth 1978 and Hofstadter 1979 among others).

One approach is essentially similar to that presented here. McNeill's work (1979), encompasses both the idea of modelling a conceptual structure as a graph and of using graph-theoretic measures to describe conceptual structure.

For McNeill, a conceptual structure is "meant to be a model of the organisation of concepts that the speech user manipulates". He continues

"Built into the definition of conceptual structure is the assumption that the speech user functions on a single complex level of representation which we will say is the conceptual structure; components of this structure, which we may analyse as belonging to different levels (for example, different types of sign), converge onto a single mode of representation, the conceptual structure level".

He goes on to use distances on a graph to calculate, for example, the "semantic focus" of an utterance.

Two points separate the work presented here from McNeill's. (It should be noted that the original formulation and results presented here antedate McNeill). First, the level of concept is more general here, as is the level of analysis. McNeill is analysing utterances or sentences. In this work the unit is the paragraph. Second, McNeill, whilst explicitly including different levels, nevertheless makes the assumption that distances across levels are to be regarded as the same as distances between levels. In other words the relation that defines the edges is not strictly

uniform, i.e. not a graph. In this study, the relation that defines the edges is either uniform, or, as proposed in the extended model of Chapter 7, explicitly shown as non-uniform.

1.3 Aims of the work

The primary aim of the work described here has been to explore and evaluate the use of a graph model. To achieve this aim, the work was divided into seven sub-goals or research tasks:

i) to put forward a structural model for a system of ideas
ii) to provide a procedure for deriving the model from the data
iii) to illustrate the use of the model by applying it to a substantial body of data
iv) to describe the results obtained
v) to evaluate the model in the light of the results
vi) to suggest improvements
vii) to assess the significance of the model

These tasks are discussed in turn below.

1.3.1 A structural model

The distinguishing feature of the model has been its rigorous definition as a graph. This has served to position it within a well established mathematical theory. The advantages of a graph model were two-fold: firstly, the visual impact of a graph is direct and immediate. The model builder is given an instant picture of the way in which the ideas in the system fit together. Secondly, the invariants associated with a graph (invariants are numbers which do not change for any graph isomorphic to the original) have been used

to provide quantitative descriptors of the structure of the idea system and this allows for the comparison of different idea systems.

1.3.2 Building the model

In order to define a graph it is necessary to define the vertices and edges of the graph, that is, to decide what conceptual categories are of interest and to provide rules for assigning units of text to specific categories, together with a uniform relation which will define the edges. In practice this amounts to performing a Content Analysis on the documents constituting the data.

1.3.3 Testing the model

The practical requirements of the research project, of which this work was a part, (see Section 1.4), made it imperative not only to suggest the method, but to test it operationally and to use it to provide results of substantive interest. Accordingly, the model has been tested using a corpus of 126 documents. All the documents were party political documents, most of them were Scottish National Party documents.

1.3.4 Assessment of the model

An assessment of the model must cover the substantive results (in this case on Scottish Nationalism), methodological issues and the potential significance of the method. The methodological issues included those concerned with the construction of the model from the data, both the reliability and the validity of the procedures actually used and the validity of using any form of Content Analysis in the study of political idea-systems. The major methodological issues,

however, were those concerned with the model qua model and included the form of the model, the robustness of the model and the associated measures, the use of the model as a measurement tool and the practicality of using the model in its present form.

1.3.5 Extensions to the model

The experience gained in analysing a large body of data has suggested a number of directions in which the model and the procedures could be improved and extended. In particular a much greater degree of computerisation is required.

1.3.6 The significance of the model

The construction and use of the model is evidently not confined to political systems. In order to make a comprehensive assessment of the work it was necessary to go beyond the immediate context of the research and to establish what kind of data would be amenable to analysis using the model. That done, it has been possible to indicate, briefly, the significance of the model for related fields, most notably in Cognitive Science and Cognitive Anthropology, and to consider some theoretical implications for substantive theory, of using a structural model in the analysis of text.

1.4 Background

It will be helpful in understanding the particular choice of priorities and research tasks to know something of the background to the original project.

This project had its roots in the research interests of four different researchers working together in a multi-disciplinary team. These interests were an interest in the growth and development of Scottish Nationalism (Webb 1978), an interest

in the use of numerical measures of ideological change as input to a Conflict Model (Bowers et al 1979), an interest in the extension of measurement into soft systems (Finkelstein 1975, 1982) and an interest in applied Graph Theory (Farbey 1977). Their common interest lay in the quantitative description of system structure, especially in the notion of structure when applied to very soft systems.

The purposes of the project as a whole were thus

i) to discover what the model could tell about the idea systems of the S.N.P.

ii) to extend the concepts and techniques of measurement

iii) to extend Content Analysis to include structural measures for input into the Conflict Model

iv) to provide a methodological analysis of the use of graph theory as a structural model for very soft systems in order to understand the dynamics of such systems.

It is the last of these, the methodological study, which is the subject of this report.

1.5 Plan of the thesis

The plan of the thesis tracks that of the research tasks set out in section 1.3. Chapter 2 describes the model in detail, discussing graphs as pictures and as sources of measures and demonstrating the use of the model on a single document which can be readily comprehended. Chapter 3 is a brief history of the Scottish National Party, included for those readers not already familiar with the substantive background.

Chapters 4 and 5 describe the application: in Chapter 4 the data set and the method of Content Analysis actually used

to derive the model from the raw documentary data are described. Chapter 5 contains the results of the analysis.

The evaluation of the model is made in Chapter 6 and in Chapter 7 an extended model for interactive coding and analysis is proposed. In Chapter 8 the work is related to similar work in other disciplines. Chapter 9 summarises the thesis and presents the conclusions.

2. The Model

2.1 Introduction

In this chapter the graph model is described in detail: graphs as pictures and as sources of measures are discussed and the concepts involved in using the model are demonstrated for a single document.

2.2 Relational Structures

The structural model proposed is, in its most general form, a relational structure as defined in Bridge (1977).

$$\mathcal{U} = \{A, \{R_i\}_{i\in I}, \{f_j\}_{j\in J}, \{c_k\}_{k\in K}\}$$

with associated functions λ: $I\in N^+$

μ: $J\in N^+$

such that i) A, the domain of $\mathcal{U}$, is a non-empty set

ii) I is a (possibly empty) set such that for each $i\in I$, $\lambda(i)$ is a positive integer and R_i is a relation on A

iii) J is a (possibly empty) set such that for each $j\in J$, $\mu(j)$ is a positive integer and f_j is a function on A

iv) K is a (possibly empty) set such that for each $k\in K$, c_k, a distinguished element or constant is an element of A.

In simple terms a relational structure consists of a set of elements which need not, in principle, be all of the same kind, on which one or more relations or functions have been defined. It is possible to regard a graph as a relational structure in which all the elements are of the same kind, in this case, ideas, and on which a single uniform relation has been defined - for example, co-mention in a paragraph. To a large extent it is the uniformity of the relation which characterises the model used. It has been considered important to preserve this uniformity in

order to keep the interpretation of the results, in particular the interpretation of the invariants, as clear as possible. For example, the interpretation of the 'distance' between two ideas, and hence all the measures based on distance such as 'centrality', become obscure if the edges connecting the ideas are not all of the same type. It is as though a rail journey were to be compared with a bus journey solely in terms of the number of changes required, without making due allowance for the difference in speed, or the difficulty of changing. Moreover, any generalisation of the graph model in the direction of a more complex relational structure would necessitate an explicit recognition of the problem of interpretation. That is, the model user would be forced to make a conscious decision about the notion of distance for each relation and justify it, even if it were an Ockhamist decision to regard them as all the same.

The requirement that the results be interpretable in the substantive field from which the documents were drawn reflects the purposes of the original project and sets it apart from the work on Semantic Networks found in Computational Linguistics and Artificial Intelligence generally. The model was not intended as an efficient data-structure for storing, retrieving or manipulating content. Rather it was meant as a measuring tool. The model had to say something more, more succinctly, and preferably in numbers, than would a straight-forward reading of the documents concerned.

2.3 Graphs as pictures

2.3.1 Appendix 1 is a reproduction of a complete Scottish National Party document (Doc. 119,1974). The paragraphs have

been numbered: there are ten in all, including headings. Each paragraph has been assigned a category code, corresponding to the main themes of the paragraph, as shown in Figure A1. The precise coding procedure and the categories used are those defined in Chapter 4. The graph of the document is shown in Figure A2. Each vertex corresponds to a theme mentioned in the document and each line to co-mention in the same paragraph. For example, paragraph 4 contains references to the theme 'Income, Wages and Salaries' and to the theme 'Organisation/Structure of communally based political organisation or party: e.g. SNP' (Codebook 1978) and this accounts for the connection between the two. Similarly paragraph 6 contains three themes 'Critique of London government', 'Income, Wages and Salaries' and 'Referencing', and this gives rise to the triangle on the graph joining the appropriate themes.

2.3.2 Considering the graph just as a picture, it will be seen to be in two pieces - a single vertex not connected with the main body and a heavily triangulated, wheel-like structure.

2.3.2.1 The appearance of the isolated vertex in this particular graph is inevitable - the category 'No Content', cannot, by definition, connect with any of the other categories. Many of the graphs that have been drawn include this vertex. It is also possible for a graph to include categories with content as isolated vertices, for example, the theme 'Local Government' was isolated throughout all the Conservative party documents that were analysed, and this is significant, suggesting either an incoherence in the document or, as happened frequently, an inclination to discuss the particular topic in its own terms.

2.3.2.2 The triangular structure of the main block must in part, be due to the rules of coding (these are set out in Chapter 4), which allowed for up to three themes to be coded for a paragraph. Had, say, four themes been allowed, there might, in general, have been a greater number of 4-complete subgraphs. However, not all the graphs were as strongly triangulated as this one, in which four out of ten paragraphs are triangles directly coded for a single paragraph, and two emerge from building up the graph for the whole document.

2.3.2.3 The wheel like shape is in no way implied by the coding. This is a feature of the document itself and suggests a central 'critical' theme, which is itself worked out and related to the surrounding themes in the document. One would expect 'Critique...' to be in some sense a central theme of the document and it is, being mentioned in every paragraph of content, except paragraph 4 (where it is implied) and paragraph 8 (which is an SNP proposal).

2.3.2.4 The position of the circles and lines on the page is arbitrary. This particular graph has been drawn as a planar graph, i.e. with no lines crossing, and this does impose a limited kind of positioning on the picture. For example, the vertex 'Devolution' could not be moved outside the quadrilateral formed by 'Income..', 'Critique..', 'Social Services' and 'Organisation/structure..', without destroying the planarity of the graph. By contrast, it could be moved into the triangle 'Critique..', 'Income..' and 'Organisation/structure..' leaving the graph still planar. In other words the position of the vertex 'Devolution' near 'Social Services' is an accident of the

drawing, its position in the quadrilateral mentioned above could be a real feature of the document.

2.3.3 To summarise, the picture of the graph gives of this fragment of the SNP idea-system, is of a coherent, closely argued system, with criticism of the central government as the central concept: part of the document being propagandising or advertising and the rest being presented in a simple way. In order to make this description both concise and precise, it is necessary to consider the other aspect of the graph as model - the graph theoretic measures.

2.4 Graph measures and their interpretations.

There are a considerable number of invariants of a graph which could, in principle, be used to describe the structure Not all of these possible measures admit of a sensible interpretation in political or idea-system terms. Of those which have been examined three general classes have proved to be useful:

i) measures of centrality

ii) measures of complexity

iii) measures of flexibility

These are discussed in turn below.

2.4.1 Centrality

These measures can themselves be grouped into three kinds - those concerned with the centrality of the document as a whole, those concerned with the centrality of particular themes and those concerned with the centrality of particular relations. These 'centralities' are evidently not of the same kind and the first result of using the model has been to clarify the different

notions of centrality that arise in the discussion of an idea system. A concept can be central to an idea-system in a number of ways which are not necessarily correlated.

Considering the document as a whole first, the measures which have been used to determine whether or not a document or group of documents were centralised were

i) the circumference of a graph

ii) the radius

iii) the diameter

and iv) the ratio of the radius to the diameter

2.4.1.1 The circumference of a graph is defined as the length of the longest cycle. (The definitions are all based on Harary (1972) unless otherwise stated). A graph with a small circumference is likely to be more compact than a graph with a large one. The circumference of the main component of graph 119 is 7. It is compact when compared with the other documents.

2.4.1.2 The radius of a graph is the eccentricity of the centre. It measures the distance from the central themes to the peripheral themes. Again, if this is small, the graph, and hence the document (and the idea-system) is compact and centralised. The radius of graph 119 is 1 i.e. there is no theme which is not connected directly to the central 'Critique..' theme. The graph is very compact.

2.4.1.3 The diameter of the graph is the value of the largest eccentricity, i.e. the longest of the shortest paths between two themes on the graph. A large diameter would suggest a diffuse system in which some themes were only distantly related

to others. The graph of 119 has a diameter of 2, again emphasising the compact nature of the document.

All three measures then, reinforce the impression of the picture, of a closely argued document, revolving around a single theme.

2.4.1.4 The ratio of the diameter to the radius has, on occasion, proved an interesting measure of the centrality of the whole graph. A graph with a high ratio of diameter to radius would imply peripheral themes 'on either side' of the centre. Graph 119 has a diameter to radius ratio of 2. In this case the ratio does not tell us anything new - a graph of radius 1 is either a star graph, or has a diameter of 2 by definition.

2.4.2 Turning to the centrality of particular themes, the invariants which have been examined are

i) eccentricity

ii) centre

iii) degree

iv) cutpoints

2.4.2.1 The eccentricity of a vertex is the longest of the shortest distances on the graph from that vertex to any other. In graph 119, the theme 'Critique..' is just 1 step away from every other and the graph has eccentricity 1. All other themes of the main component have eccentricity 2. In general, a theme of high eccentricity is a peripheral theme, a low eccentricity argues a central position in the system.

2.4.2.2 The theme, or set of themes with minimum eccentricity in a graph, is the centre of the graph. Like eccentricity, the centre of the graph is a positional feature. In graph 119,

'Critique..' is at the centre.

2.4.2.3 The degree of a vertex is equal to the number of edges incident to it.

Degree is in part a measure of centrality, in the sense that a theme of high degree is central by virtue of being related to a large number of other themes. It is a measure of the penetration of a particular theme through the idea-system.

In graph 119, the theme of highest degree is, not surprisingly, 'Critique..', with degeee 6. 'Devolution' has degree 2. It has not, in this document, been systematically related to the other themes. 'Critique..' has.

2.4.2.4 A further extension to the notion of centrality can be derived from the concept of a cut-point. A vertex is a cut-point if its removal increases the number of components in the graph. In idea-system terms, if the removal of a theme would result in the system breaking up into two (or more) unconnected pieces, that theme is central to the coherence of the system.

In graph 119, removing a single theme from the graph will not disconnect it. Even without 'Critique..' all the other themes would remain connected. To disconnect the graph, both 'Critique..' and 'Organisation/structure..' would have to be removed. 'Devolution' would then be disconnected from the others.

A very similar concept is that of a bridge. A bridge is an edge whose removal would increase the number of components. The bridge is a concept of centrality which involves the centrality of a relation rather than a theme.

There are no bridges in graph 119. At least two edges would have to be removed to increase the number of components (those incident to 'Devolution').

2.4.2.5 These structural notions of centrality can be contrasted with the common notion of centrality as frequency of mention or 'salience'. For example, in document 119, the theme 'Organisation/structure..' is the most frequently mentioned after 'Critique..'. It has eccentricity 2. By contrast, 'Devolution', which is mentioned only once is structurally just as central, with eccentricity 2.

2.4.3. Centrality then, is not a single concept. There are a number of ways in which a theme can be central to an idea-system. Themes can be much talked about, i.e. salient, without being structurally central and vice-versa. The use of the model has served both to clarify individual notions of centrality and to provide measures for these notions.

2.4.4 Complexity

Complexity too is not a single notion. The purpose in trying to provide a measure or series of measures of complexity is to provide a basis for the study of the development of an idea-system over time. It would, for example, be useful to be able to say that "in its early years a system was complexity 4 and that at its peak it reached complexity 6, but that as time went on it decayed to complexity 2". It might then be possible to observe patterns in development across different idea-systems (Farbey, Mitchell and Webb 1979).

However, even in the everyday meaning of the word there are two distinguishable senses of the term 'complexity'. There is

'complex' meaning 'consisting of parts, composite' (Oxford English Dictionary) and there is 'complex' meaning 'intricate, having a complicated structure' (ibid).

A further notion of complexity arises from the model in that an idea-system can be complex because particular themes are themselves complex, i.e. have been worked out in great detail without necessarily much reference to other themes. This kind of thematic complexity reflects increasing specialisation.

The graph invariants which have been used to capture this range of concepts, with a varying degree of success, were

i) the number of vertices
ii) the number of components
iii) average degree
iv) point and line connectivity
v) coarseness, thickness, genus and crossing number

Of these, the first two relate to the 'composite' idea of complexity, the last three to 'intricacy or structural' complexity. Average degree, while largely a measure of intricacy also forms a link with thematic complexity.

These measures are discussed below.

2.4.4.1 The number of vertices is a gross measure of composite complexity. A document with 100 themes is very likely to be more complex than one with 5, but a document with 8 themes, such as 119, might very well be simpler to grasp than a single theme document in which that theme had been very carefully worked out. It is partly a matter of style and purpose, especially when considering a single document, and partly a matter of level.

If a large number of documents are considered together, so that style and purpose were effectively ironed out, and if all the themes could be shown to be at a similar level of thematic complexity (for example, if one could code 'cat', 'dog', 'rabbit' as opposed to 'cat', 'Alsatian', 'Doberman', 'rodent', 'rabbit') then the number of themes would indicate composite complexity. In this study, the measure is ineffective for this reason.

2.4.4.2 A graph with a large number of components has greater composite complexity than one with few. It consists of lots of little bits. As a measure of complexity, however, the number of components suffers from the same disadvantages as the number of vertices. Over a single document the number of components depends heavily on the style: over an entire idea-system a graph with only one highly connected component, such as the main component of 119, may well be more complicated than a system with several simple components.

As a measure of intricacy, the number of components tends to work in the opposite direction to the measure for composite complexity. That is, the fewer the connections made between themes, the simpler it is to understand. For example, if a connection is made between two themes which are in different components, the number of components decreases although the new system is more complex, having an extra relation.

As with the number of vertices, the number of components, therefore, has not proved very useful.

2.4.4.3 The average degree of a vertex is much more straightforward. As more and more themes are related to each

other, the idea-system becomes more intricate and the structural complexity increases. The average degree of 119 is 3.4, not large when compared with the number of vertices.

2.4.4.4 The point connectivity of the graph is the minimum number of points whose removal disconnects the graph. Analogously the line connectivity is the minimum number of edges which must be removed to disconnect the graph.

As the number of themes which have to be removed (before the graph becomes disconnected) increases, the idea-system which the graph models must be increasingly intricate.

In graph 119, the point connectivity is 2 - both 'Critique..' and 'Organisation/structure..' must be removed before 'Devolution' becomes disconnected from the main body. The line connectivity is 2 - the edges connecting 'Devolution' to the graph both have to be removed in order to disconnect 'Devolution'.

2.4.4.5 Four topological invariants have been included in the analysis. These are

i) genus: the number of handles needed on a sphere in order to embed the graph

ii) thickness: the number of planar graphs required to form the graph

iii) coarseness: the maximum number of edge-disjoint non-planar subgraphs

iv) crossing number: the number of crossings there must be when the graph is drawn in the plane.

For graph 119, the genus and the crossing number are 0, because the graph is planar and the thickness is 1. There are no non-planar subgraphs and the coarseness is 0.

As each of these measures increases, so does the complexity of the graph. Perhaps the most important jump in the measure is from 0 to 1 or more, i.e. a graph is single layered or multi-layered. The underlying idea-system is correspondingly single layered and simple, or multi-layered and complex, each layer representing a new level of complexity.

However, it is difficult to see any satisfactory, intuitive interpretation of this dimension of complexity in the context of a political system. It is much easier in, for example, electronic circuitry, where the connection could literally cross.

2.4.5 To summarise, complexity, like centrality, is not a single notion. Two aspects of complexity, composite complexity and intricacy, can be described using graph invariants, although the interpretation of these measures is not always straightforward.

2.4.6 In this section a graph theoretic concept of mobility is discussed. This concept differs from those discussed previously in that additional assumptions about the nature of the system must be made which may be difficult to justify in the substantive context.

The assumptions which have to be made amount to an assumption that the system consists of rigid, fixed length links and of joints which constrain the system to move either in the plane or in 3-dimensional space.

The discussion which follows is therefore of an analogy - 'if the system were to be modelled as a physical system of links and joints then the mobility properties would be...'. The definitions and measures are taken from Graphs, Networks and

Design (Open University 1981)

2.4.6.1 The planar mobility of a graph, M, is defined as

M = 2 x number of vertices - number of edges - 3. A graph of mobility 0 is said to be rigid.

The planar mobility of graph 119 is -1. That is, if the graph is considered as a planar system, it is rigid. In fact, it is overconstrained. One edge, 'Organisation/structure..' to 'Social Services', could be removed and the resulting graph would still be rigid.

2.4.6.2 The spatial mobility of a graph is defined as

M = 3 x number of vertices - number of edges - 3. A graph of mobility 0 is rigid.

The spatial mobility of 119 is 6. There is a good deal of mobility if the graph is considered as a spatial system. For example, the vertex 'Devolution' could move relative to 'Critique..' and 'Organisation/structure..'. It could as it were, swing up above the page or down into it.

2.4.6.3 The pattern of mobility which emerges when the paragraphs of 119 are considered in order, is interesting. This development is shown in Figure A3.

The planar mobility was small everywhere, reinforcing the impression of a closely argued document obtained from the wheel-shaped, planar graph. It is almost as though the author were deliberately tying up the threads of his argument, introducing new subject matter and immediately relating it to what had gone before.

The spatial mobility revealed a slightly different pattern. It peaked about half-way through the content paragraphs,

suggesting that the arguments build up to a certain level of complexity before being wound down again as the loose ends are tied up.

2.4.6.4 It is easy to see the kind of hypothesis which could be put forward, given that the mobility analogy were justifiable in a given context. One might postulate for instance, that there was a tendency towards triangulation, or at least rigidity, in the style of the SNP documents. Alternatively one might suggest that an idea-system would develop/atrophy around flexible areas in the graph - and so on.

Substantive justification could come from a theory of idea-systems development, or, possibly, from an empirical study of a large sample of documents on a number of data points. Neither possibility was available within the project and the measures, however suggestive, must be treated as metaphor.

2.5 Summary

The graph theoretic model has been described in detail, with reference to a particular document. Three groups of measures have been developed: measures of centrality, of complexity and of mobility. Centrality and complexity were seen to have more than one interpretation in the context of the model and these interpretations were not necessarily correlated. The measures of mobility, although interesting as a basis for speculation, require stronger assumptions about the nature of the system than could be justified in a political context.

3. Historical and political accounts of Scottish Nationalism

3.1 Introduction

As indicated in the introduction (section 1.3.3), the model has been tested on a substantial collection of documents. Most of these documents were documents of the Scottish National Party. In order to appreciate the discussion of the results of the analysis, it will be helpful to know something of the background of Scottish Nationalism.

In this Chapter, therefore, a brief account is given of the history of Scottish Nationalism in its modern form and of previous work done in the analysis of Scottish Nationalism. The account is intended only for readers not familiar with the substantive background. The history of Scottish Nationalism has been described at length in Brand (1978), Hanham (1969), Harvie (1977) and Webb (1978). Aspects of the development of the movement are described in Parsler (1980) and Urwin (1982). Histories of Scotland are numerous: Mackie (1972) and The Edinburgh History of Scotland are both authoritative accounts.

3.2. Historical Accounts

A sense of Scottish national identity and Scottish national sentiment can be traced back to the 9th Century Kingdom of Alba and its successor, the Kingdom of Scotland. For 1100 years there has been 'something that it is like to be Scottish' and that 'something' has been different from being British and, even more so, English.

In more recent times this continuing sense of national identity has been buttressed by the distinctively Scottish institutions which survived the Union: the Church, the Scottish

legal system, the Boroughs and the educational system. This identity finds expression in other areas too, for example in sport or in the Scottish regiments.

Modern Scottish Nationalism is not identical with this continuing national sentiment. Rather it represents the transformation of this sentiment into a political mass movement.

3.2.1 Scottish Nationalism in this latter sense can be traced back in a more or less continuous line to the formation of the Scottish Home Rule Association in 1886. Essentially a Liberal creation (Mackie, 1972, p.367) the S.H.R.A. did have some connection with the emerging Labour Party. The S.H.R.A. was committed to attempt to work through the existing parties rather than as a party in its own right. Although the S.H.R.A. could claim some success in influencing public opinion of the day and in getting Home Rule debates in the House of Commons, by 1906 it had flagged and the initiative of the Nationalist movement had passed to the Young Scots Society, an almost exclusively Liberal grouping. This movement was successful enough to force the hand of a Liberal government and a Bill for Home Rule passed its second reading in 1913.

3.2.2 The Bill was lost in the catastrophe of the Great War and the consequent sea-change in British politics. The rise of the Labour Party coupled with the decline of the Liberals meant, too, the decline of the Young Scots Society, which was cut off from the centre of influence at Westminster (Webb, 1978, p.64).

Not long after the war, however, a number of small, radical, nationalist groups began to appear, a revived S.H.R.A. setting

the tone with its founding resolution:

"That this meeting, representing all shades of Scottish opinion and industrial activity, being convinced that the present centralised system of government from London is inefficient and inconsistent with national sentiment, resolves to form itself into a committee for the purpose of organising and focusing the Scottish demand for self-government in respect of Scottish affairs."

It was the post-war S.H.R.A. which, together with the Scots National League (1921), the Scottish National Movement (1926) and the Glasgow University Student Nationalist Association (1927) formed the National Party of Scotland in 1928.

3.2.3 The N.P.S. was undoubtedly a political party formed for the purpose of fighting Parliamentary elections. It was largely socialist in composition (both the Chairman, John McCormick and the secretary, R.E. Muirhead, were I.L.P. members) and was much closer to the demand for complete independence than the S.H.R.A. had ever been, although this demand was tempered by McCormick's pragmatic approach.

At much the same time (1932), a grouping of nationalists on the other end of the political spectrum had been formed - the Scottish Party. The Scottish Party consisted in the main of Liberals and Conservatives, many of the latter being former members of the Cathcart Divisional Unionist Association. The party was devolutionist rather than secessionist in its views.

3.2.4. In 1934 the National Party of Scotland and the Scottish Party merged to form the Scottish National Party. The policy

and personal tensions inherent in the fusion of these disparate groups remained. The degree of self-rule demanded, the best methods of obtaining it, the right of the present party to determine the policy of a future Scottish Parliament and the Class issue have all occasioned major splits in the Party. To a limited extent these splits and the consequent breakaways have solved the problems for those left behind on each occasion.

The most dramatic of these was McCormick's departure in 1942 to found the Scottish Union, later the Scottish Convention. The Convention, at its peak, obtained between two and two and a half million signatures on a new "covenant", which pledged the signatories

"in all loyalty to the Crown, and within the framework of the United Kingdom... to secure for Scotland a Parliament with adequate legislative authority in Scottish affairs" (Webb 1978, p.92) - Home Rule.

Eight years later R.E. Muirhead left the S.N.P. over the issue of dual membership to found the Scottish National Congress, a pacifist group, influenced by Ghandi and the Indian Congress Party.

3.2.5 Both the Convention and the Congress have disappeared, leaving the S.N.P. in the forefront of Nationalist politics: a party committed to full independence, to achieving independence through electoral advance, non-violent and, until very recent times, drawing allegiance from all classes. In so far as it is successful, it has succeeded in turning a 'latent, private sense of ethnicity into a public, manifest one' (Smith 1981, p.20).

3.3 Political Accounts

The explanations which are available for this success, and which form the substantive background to the project, are generally of two sorts: those which are principally concerned with the Scottish phenomenon and those which are principally concerned with the Nationalist phenomenon. These two kinds of explanation are described briefly below, again principally for the reader unfamiliar with the substantive context.

3.3.1 It should be noted that the terms 'ideology' and 'idea-system' are used interchangeably throughout the discussion of the S.N.P. documents. This is in effect, a stipulative definition of the term 'ideology' which is thereby identified with and idea-system as manifested in the public output of a group or organisation. Both the identification and the justification for it are due to Webb (Farbey, Mitchell and Webb, 1979) and are discussed in that paper.

3.3.2 The analysis of Scottish Nationalism as a Scottish political phenomenon has had as its main object the explanation of the rise of the S.N.P. as a political force, in particular the exponential rise in the party's vote which occurred in the 1960's and 1970's. These analyses are largely historical and socio-economic. Harvie (1977), for example, claims that

"no understanding of the forces making for a renegotiation of the Union is possible which omits the historical factors which have kept the Union in being"

Brand (1978) '..identifies Scottish Nationalism as a modernising movement, not as an attempt to re-create a Scottish

Golden Age but as a determination to work for a technologically advanced, prosperous, modern, small nation-state..." and, just before,"The vote for nationalism was a vote to establish a modern, prosperous society where there were problems of unemployment and decay." (p.23). Neither Brand nor Harvie are particularly concerned with the development of S.N.P. ideology per se.

Webb (1978, p.319), by contrast, claims that

"...the form of the S.N.P. is not determined only by the social and economic characteristics of Scotland, nor even by the historical and cultural conditions in the country, but also by the beliefs and visions of the nationalists themselves: how they structure their political world...".

This study, in its substantive aspect, is an attempt to capture that structure.

3.3.3 The second source of explanation and ideas has been the general literature of Nationalism, of which Scottish Nationalism is a case in point. This includes work on regional and ethnic nationalisms (Urwin 1982, Nairn 1981 and Hechter 1975) as well as theoretical accounts of the Nationalist phenomenon (Smith 1971 and 1981, Kamenka 1973, Kedourie 1960, Kohn 1962 and Deutsch 1966).

3.3.3.1 There were, in the more contemporary accounts of nationalism, fewer points of contact than might naively have been expected. This would seem to be because the later accounts, like the specifically Scottish accounts, de-emphasise or even explicitly criticise ideology as a basis for the taxonomies of modern nationalism which are their primary concern.

Smith (1971), for example, writes:

"..we can clearly see the European ethno-centrism of all historical schemes, and the need to dismiss all extra-European cases as 'borrowings', which results from paying too much attention to ideology at the expense of the movement. The tracing of ideological pedigrees on a world-wide scale reveals the inherent implausibility of the whole undertaking, even if the concentration on Europe conceals it."

Smith finds a sociological classification more relevant, contrasting 'primitive' with 'developed' (i.e. organised and persistent) nationalism, as well as 'territorial' and 'ethnic' nationalisms.

Gellner, too, is dismissive of ideology:

"Their precise doctrines are hardly worth analysis. We seem to be in the presence of a phenomenon which springs directly from basic changes in our shared social condition, from changes in the overall relation between society, culture and policy. The precise appearance and local form of this phenomenon no doubt depends a very great deal on local circumstances which deserve study: but I doubt whether the nuances of nationalist doctrine played much part in modifying these circumstances." (Gellner 1983).

The study of nationalist ideology, then, is not very fashionable.

3.3.3.2 Nevertheless, there were some points of contact. Smith, in particular, in his later work (1981), classifies Scottish Nationalism as 'ethnic' (and Gellner notes that Scottish Nationalism is the sole version which does not fit

into his classification!).

The first point of contact concerns the repertoire of responses available to ethnic groups with respect to the majority community. Smith's list of strategies ranges from isolation through accommodation, communalism, autonomism, separatism and irredentism. The history of Scottish National movement, as described in the first part of this chapter, would appear to realise the conflict between the different strategies. At any one time a particular strategy dominates, currently the S.N.P. and separatism, whilst the other strategies are recessive, the property of certain groups or individuals, but which may return with a change in circumstance.

The second point of contact is Smith's contention that

"...nationalism extends the scope of the ethnic communities from purely cultural and social to economic and political spheres: from predominantly private to public sectors. To make any headway in the modern world, ethnic movements must stake their claims in political and economic terms as well as cultural ones, and evolve economic and political programmes. They must organise in the political market-place... ethnicity is a total phenomenon, covering all aspects of social life."

This shift from the cultural to the political can be detected using the model and it was the first hypothesis to be tested. The immediate source of the hypothesis though was from the work of Webb and the particular form it took was proposed by him.

A third point of contact between this work and established nationalist theory derives from Deutsch (1966) and stems from a similarity in approach.

Deutsch, although he uses primarily economic and demographic variables is concerned to provide a functional definition of Nationality which is not "merely subjective... [but] ... At any moment exists as an objective fact, measurable by performance tests."

This work, too, is an attempt to provide a model and a measurement procedure for manifest behaviour, an operationalisation of ideological variables which can be used with other types of variable in the explanation of the Nationalist phenomenon.

3.4 Summary

In this Chapter an account of the substantive background to the project has been presented. In the first section, the historical events which form the context to the written documents have been briefly set out, in particular those pertaining to the development of the S.N.P. as a political force.

In the second section, previous work done in the political account of the development of Scottish Nationalism, in both its Scottish and its Nationalist aspects has been noted and the salient points of difference and of contact set out.

4. Building the model - The Data Set and the Content Analysis

4.1 Introduction

In order to derive the graph model from the raw data, as represented by the documents, it was necessary to use a form of Content Analysis. This analysis generated both a coding of the documents into themes, i.e. the vertices of the graph, and the relation "co-mention in a paragraph", which constituted the edges. In this Chapter the set of documents available for analysis is described first. This is followed by a full description of the method of Content Analysis actually used, including the categorisation of the concepts to form themes, the selection of coding units and the assignment of units to categories.

In addition an indication of the reliability of the coding thus obtained is given.

4.1.1 Content Analysis

(de Sola Pool (1959), Stone et al (1966), Gerbner et al (1969), Carney (1972) and Krippendorf (1980), North et al (1963) and Osgood et al (1957)) is a collection of techniques for the systematic, (and, as far as we are concerned here) quantitative analysis of communication content. These techniques are concerned with 'establishing categories, context and counting units, measurements of data which these produce and the sample.' (Carney, 1972, p.280).

Typically a content analysis would proceed by defining the categories of interest, assigning the units of text to the appropriate category, observing the resulting frequencies and drawing inferences from these results. All the stages need to

be validated, the first and last must ultimately be valid substantively, the second and third stages can be analysed statistically.

4.1.2 Classical content analysis has been very much concerned with describing the manifest content of communication in terms of the frequency of particular events - the occurrence of a particular category, say, or the number of column inches devoted to a particular topic. Modern Content Analysis continues this work, but also makes use of the techniques of multivariate analysis such as principal component analysis or discriminant analysis (Osgood et al, 1957; North et al, 1963; Krippendorf, 1980).

By and large, frequency and multivariate content analysis does not aim at eliciting a categorial structure from a pattern of co-occurrences as has been done here. Work of this kind is known in Content Analysis as Contingency Analysis (Osgood, 1957).

A contingency analysis

"...starts with a set of recording units, each of which is characterised by a set of attributes which are either present or absent.....In a second step the possible co-occurrences of attributes are counted and entered as properties. In a third step the statistical significances of these co-occurrences must be tested" (Krippendorf p.115).

According to the significance or otherwise of co-occurrences a distance formula can be computed for every pair of categories. A spatial model can be derived from this distance matrix which looks like a graph but is tied to a metric other than distance-

on-the-graph. The method is, thus, not strictly the same as that discussed here, although similar in purpose and in the visual impact that it produces.

4.1.3 Krippendorf (1980 p.156) provides a useful typology of the problems arising in validating a content analysis. There are two groups of problems. The first concerns reliability and includes questions about the accuracy, stability and reproducibility of the results. The second group which he calls 'validity proper' includes questions about the semantic, sampling, categorial, predictive and correlational validity of a particular study. Over and above these technical problems is the question of Content Analysis as a valid form of enquiry.

Since these problems are common to all Content Analyses and are discussed at length in all the references above, there is little more to do here than to note their existence and to use them as a framework for assessing the validity of the particular analysis used in this study. (Chapter 6).

4.2 The Data

In this section the documents available for analysis are described. The documents were predominantly those of the SNP, with a small number of documents from other parties. This reflected the immediate substantive aim which was to discover structure in the idea-system of the SNP and to describe changes in that structure over a long period. The documents from other parties were only used for very broad comparisons.

4.2.1 There were two major groupings of the documents. The first was based on the documents available at the start of the project and the second based on a slightly larger collection

which became available later in the project.

4.2.1.1 The number of documents from each party in the original data set is given in Table 4.1.

PARTY	NUMBER OF DOCUMENTS
SNP	90
Conservative	19

Table 4.1 Breakdown of the first data set, by Party.

The SNP documents were further divided into three groups, those dated before 1960, called 'early' SNP, those dated after 1960, called 'late SNP', and those which were not positively dated. The numbers in each group are shown in Table 4.2.

GROUP	NUMBER OF DOCUMENTS
Early SNP	22
Late SNP	58
Undated	10

Table 4.2 Breakdown of SNP documents, first data set, by date.

4.2.1.2 Late in the project more SNP documents became available, as a result of which it was possible to re-divide the SNP documents into four, consecutive groups:

i) SNP documents dated before 1940 (SNP 1)

ii) SNP documents dated after 1940 but before 1960 including documents for which there was no positive date, but which could be reasonably assigned on the basis of their content (SNP 2) (for example, by reference to an event which could be dated) or, on the basis of the analysis

of the original data-set.

iii) SNP documents dated after 1960 but before 1974 again including previously 'undated' documents which had been assigned to this group as in (ii). (SNP 3)

iv) the new SNP documents dated between 1974 and 1979. (SNP 4)

This gave four consecutive data points. The number in each group is given in Table 4.3.

GROUP	NUMBER OF DOCUMENTS
SNP 1	4
SNP 2	26
SNP 3	60
SNP 4	17

Table 4.3 Breakdown of the later data set -

the numbers correspond to the descriptions above.

4.2.2 The analysis thus took place in two phases. The first contrasted 'early' and 'late' SNP and each of these with the Conservative documents. On the basis of these results, and whatever other information could be gleaned from the content (and, indeed the typography) of the documents, the undated documents were assigned as being either early or late. The enlarged set was then used to contrast the groups of documents from the same party, over time.

4.2.3 The data set was not complete. It did not, and could not, include every leaflet, manifesto, press release or speech transcript of the SNP over the period covered. Nor was

it possible, within the resources of the project to arrange for a statistically sound sampling frame, with the result that, for instance, the pre-1940 group is very small. Nevertheless, it is a substantial sample, considered adequate by the political scientists for whom the work was done, and more than adequate as a demonstration of what might be achieved using the graph-theoretic model.

4.3 Coding and the Code-book

4.3.1 The procedure used to derive the model from the data was a form of Content Analysis. In this analysis, the themes of interest were identified, together with a collection of supporting sub-themes which would aid classification. Each paragraph of a document was coded as containing one or more themes - a theme appearing in a paragraph of a document appeared as a vertex on the graph of that document. If more than one theme appeared in a paragraph there would be an edge joining the two themes on the graph, denoting the relation 'co-mention in a paragraph'. This was the only relation considered in this study. The coding procedure is described in full below.

4.3.2 The first step in the analysis was to code the data. Four processes can be distinguished in the coding procedure

i) the categorization of concepts to form themes
ii) the selection of coding units to be categorised
iii) the assignment of units to categories
iv) the definition of a uniform relation on the categories/themes.

In this section these processes are discussed, together

with the content of the theme-list and the problems of achieving a reasonable replicability in coding.

4.3.2.1 The categorisation of concepts to form themes

The problems of devising a code-book, in other words a set of categories, each of which would form a theme, or vertex in the graph, were, first, to devise a set of categories that were relevant to the questions likely to be asked of the model and, second, to provide a sufficiently detailed explanation of these categories to the coders.

This was done in the project in a circular way as shown in Figure 4.1.

The substantive 'field expert' produced an initial set of categories. Associated with each was a set of lower-level ideas or 'sub-themes' which were characteristic of, or could be said to support, the choice of main theme. The coders (initially there were 5 coders, including the field experts) then coded trial documents and the results were compared. The process was repeated until an acceptable level of agreement, 90% or better, was reached and all the concepts encountered in the trial run could be satisfactorily encoded. An example of a theme with its subthemes is given in Figure 4.2 (Codebook 1978). An example of coding is given in Figure 4.3. A list of all the themes considered is given in Figure 4.4.

It will be seen from these figures that the choice of categories was subjective and particular. It is a strategically chosen framework, designed to reflect the researcher's concern with particular aspects of SNP policy. The choice was not, however, arbitrary.

The classification had first to be subjectively adequate, that is, 'not do violence to (the researcher's) own feel..... for the structure of what is described' (Goodenough 1969).

More importantly, it was objectively adequate in that it provided the coder with the knowledge needed to classify a document in a way that the researcher would accept as corresponding to his own classification.

An alternative to this kind of categorisation might have been to produce the categories automatically, based, say on a frequency count of key-words. (Stone et al 1966). This was not a feasible prospect in view of the time-scale and the budget of the project.

Moreover, in the context of a political idea-system, a system in which the language is partly technical and partly natural (and may even be designed to be ambiguous) the dangers of jettisoning ordinary human understanding are considerable. An automatic classification, given the present state of the art, could be naive or, at worst, produce a 'quasi-precise knowledge of less than we already know' (Crick 1976). In a more technical environment the reverse situation may obtain.

The second problem in devising a code-book was to provide the coders with a sufficient knowledge to achieve objective adequacy. This was accomplished in the project largely by providing, for each theme, a set of lower level concepts which served to explain the meaning of each theme name and to support particular choices of theme. These lower level themes, 'subthemes', effectively defined the themes.

4.3.2.2 The selection of coding units

The theme and subtheme coding was done at paragraph level; each paragraph in each document was coded for up to three themes.

In the later stages of the project, a finer grained coding at sentence level was attempted. This would have coded each sentence as containing a number of 'idea-elements'. These idea-elements were to form the basis for operationalising 'attitudinal variables relevant to the Conflict Model.... to be classified and coded.... as being about indignation, or containing an element of hostility, or revealing a sense of impotence, or advantage in political power'. The substantive aim of this part of the project was to have been '....to obtain data on a number of dimensions of community and/or elite psychological attributes' (Codebook 1978). An example of idea-element coding is given in Figure 4.5.

In practice only a limited amount of idea-element coding was done. This was partly because the volume of data generated would have been too large to handle with the resources available (each sentence may generate several idea-elements) and partly because of an increasing difficulty in getting agreement at very low-level coding. These points will be taken up again in Chapter 7.

4.3.2.3. The assignment of units to categories

The final codebook contained a list of 30 themes and 685 subthemes. Each paragraph was coded for up to three main themes and each theme could be supported by up to two subthemes.

4.3.2.4 The definition of a uniform relation on the categories

The relation considered was co-mention in a paragraph. This had the advantage of being extremely simple to operate. A possible disadvantage would be that, in a small sample, the relation would depend on the paragraphing ability and style of the author. In a large sample this did not emerge as a problem. Informally, if two ideas were associated, they were bound to turn up in the same paragraph particularly in the more prolix prose of the early Scottish National Party.

Other relations which might prove useful would be linguistic relations such as polysemy, hyponymy, inclusion etc. These relations might also form part of a multi-level study (McNeill 1979, and see also Chapter 8).

The importance of a uniform relation lies in the interpretation of the graphs. It is particularly important when distance measures are to be interpreted. McNeill (1979 p.107) argues in his analysis of hierarchical graphs that 'the inclusion of one point by another is the equivalent of a reciprocal connection between the higher and lower points of one step' which leads him to combine 'vertical' (inclusion) and 'horizontal' (association) distances. This is a reasonable first step, but should really be justified in the long run empirically.

4.3.3 The content of the theme and subtheme lists

The list of themes in Figure 4.4 is a list of 'policy' themes rather than 'pure' themes. This simply reflected the political scientist's concern with SNP policy, rather than SNP rhetoric, and with the eventual problem of linking the

ideological study with other, behavioural variables.

Much of the rhetoric and non-policy paragraphs were contained in the code 'No Content'. This category, defined as 'matter that does not fit any other coding category and for which another category creation is not desirable' (Codebook 1978), should be distinguished from the categories 'Other' and 'Uncodable'. 'Other' was predominantly a category used in the early stages by the coders to signal a gap in the theme list and virtually disappeared as the list extended. 'Uncodable' took care of the occasional meaningless remark. It too was a negligible coding in the final coding.

Paragraphs, mostly headings, with some meaning, such as 'Scotland lives' were classified 'No Content'.

Whilst all the themes in the list of Figure 4.4 are policy themes, it will be seen that some are very much more detailed than others. For example, economic themes, as well as having a general 'economy' category, are also spelled out in the themes 'Agriculture' 'Employment' 'Fishing' 'Foreign Trade' 'Income..' 'Industry..' 'Natural resources..' and 'Transport'. This leads to a certain imbalance in the levels. It is as though one were looking at a map with a magnifying glass over one portion of it. It also leads to the possibility of 'a priori' connections, that is two themes being associated not as a peculiarity of the idea-system, but for linguistic reasons (Finkelstein-private communication).

The imbalance in level is another manifestation of the problem of subjective (albeit replicable and adequate)

categorisation versus some other method, a point which has already been raised. (The most useful sections of a map to highlight depend on where you want to go).

The second problem, the problem of a priori connection is much more difficult. In this particular study the overlap between themes was not very great. Each theme was carefully defined so that, although two themes might both be included in another, hyponomous, theme, two different aspects of the higher level theme would have been described. To make the connection, an ideological, rather than a semantic connection had to be made. For example, 'Employment' and 'Industry' are not linked in all the groups of documents, although both would seem to be a part of the theme 'Economics'.

Nevertheless, the themes arrived at by this method are not completely orthogonal. Possible ways of overcoming this problem might be to use a componential analysis (as in Leech 1981) or a canonical analysis using the already coded data and the frequency of connection to suggest possible 'super-themes' (along the lines of the semantic differential work of Osgood et al 1957).

It is worth remembering in this context, that the themes are by definition related to each other at some level, even if only the very general one of 'has been mentioned in the same set of documents as'. The interesting question, from the analyst's point of view is 'do they form sub-clusters that could emerge as super-themes from those that have been coded?' (Mitchell - private communication).

4.4 Reproducibility

Reproducibility in a content analysis is the degree to which different coders are able to produce the same coding of a document. As has been mentioned (4.2.1) the coding of the documents took place in two batches. A total of seven coders was involved, two 'outside' coders and three people from the project for the first batch of documents, two more 'outsiders' for the second.

4.4.1 Initial reproducibility, especially with respect to the outside coders, was low. A trial coding of 60 paragraphs produced 41 disagreements. A detailed examination of the coding showed that on most occasions when there was no agreement, one or other coder was using a 'no content' code or 'other'. Further, a difference in coding could be magnified by a run of paragraphs on the same theme. Accordingly, the code-bookwas extended to cover those themes previously coded 'other' and existing themes were defined more carefully, using the sub-themes. After several modifications, the code-book was sufficiently detailed to allow a 90% agreement. The bulk of the coding was then carried out by the two 'outside' coders.

4.4.2 For the second batch of data, two more outside coders and one of the original outside coders were employed. Again, initial reproducibility was shaky. All three coders agreed on only seven out of 25 themes on a trial run. The original coder was able to produce his original coding almost exactly, but the two newcomers, although agreeing with each other quite well (19/25), were coding differently. In this case, the

code-book was not modified. Instead the coders were talked through the coding for several documents (in effect they were trained) until they could reproduce the required coding with 90% reproducibility.

4.4.3 The disagreement at the subtheme level was more pronounced and did not improve substantially even with training or modification of the code-book. A more detailed, subtheme level, analysis, which had been planned, had to be abandoned. Evidently it was easier to agree that a particular paragraph was about 'housing' and 'economics' than to agree why. It may have been possible, by improving the code-book, to achieve a better level of agreement, probably by making the categories even more detailed. However, this would merely have had the effect of pushing the level of disagreement one level lower. In a very soft system, such as an idea-system, the categorisation may always be part personal, part shared.

4.5 Summary

The input to the graph model, then, consisted of a code version of 126 documents. The documents were coded using a form of Content Analysis to give the themes, which formed the vertices of the graph and the relation, co-mention in a paragraph . The coding was accomplished by a number of different coders, in two main batches. After a period of training, an agreement of 90% was achieved between coders at theme level. At the level of the supporting 'subthemes', the level of agreement was not adequate as between coders and this did not improve with training. A more detailed level of model was therefore abandoned.

5. The Result

5.1 Introduction

This Chapter contains the results of applying the methodology to the data described in Chapter 4. The discussion includes a comparison of the results obtained for three different groupings of the documents. The first grouping compares 'early' and 'late' Scottish National Party documents. The second compares the groups of documents at the four points of the enlarged data set and the third compares Scottish National data with Scottish Conservative documents.

This is followed by a description of work carried out on the recognition of undated, or otherwise unassigned documents. Finally there is a brief discussion on the dynamics of idea-systems.

5.2 The data sets to be compared

The effect of the Content Analysis described in the previous chapter is to crystallise a body of text into a set of themes and a relation on those themes, in other words, a graph. The discussion in Chapter 2 centered on the graph of a single document and described in detail the measures which could be obtained on the graph. In the following sections the graphs which were obtained by grouping the documents, that is, by merging the graphs for each document to form a single graph are described and compared.

5.2.1 Three sets of graphs were compared. The graphs compared were 'early' and 'late' SNP, SNP 1, SNP 2, SNP 3 and SNP 4; SNP and Scottish Conservative documents. (See also 4.2).

5.2.2 Two rather different kinds of comparison were made. Where possible, the comparisons were made in the light of specific hypotheses derived either from political science or cognate disciplines. Where this was not appropriate the results obtained in the study were used to generate questions and hypotheses which will require explanation from the theory.

5.2.3 These hypotheses were of differing degrees of generality. The more specific derive from (political) theory, and in particular from discussions of the SNP as a peripheral, nationalist party. The hypotheses at this end of the spectrum are referred to as 'substantive'; they are concerned with the world of Scottish Nationalism and demand political explanation.

At the other end of the spectrum, the more general questions, which are referred to as 'structural' are concerned with the nature of an idea-system considered as a relational structure and the way in which such systems change over time; Scottish Nationalism is here thought of as an examplar of a kind of system, to be compared with other nationalist idea-systems, other political systems or, taking a structuralist view, with other social and cultural systems.

5.3 SNP 'early' v. 'late' - the full graphs

The first group of documents to be compared with those of the SNP 'early' and 'late' groups. The graphs were first compared in the light of three hypotheses concerning the relationship between the development of a nationalist party and the structure of its ideology. For this comparison the full graph was used, that is, every co-mention, regardless of the frequency with which it occurred, was included and given equal weight.

The three preliminary hypotheses were

i) as a nationalist party becomes a mass party, the role of cultural factors in its ideology will become less pronounced and economic and material factors more pronounced

ii) as a nationalist party grows in size its ideology becomes more diluted, especially in a democratic context

iii) as a nationalist party develops major party status in its impact on an electoral system, its ideology becomes more complex

These hypotheses, which were first put forward by Webb (Farbey, Mitchell and Webb 1979) arose from earlier work by him on Ideology and on Scottish Nationalism.

5.3.1 The role of cultural factors

Specifically Webb states

"Many of the early nationalists, although perturbed by the economic state of Scotland, were in the main motivated by a fear for Scottish culture which they saw as being increasingly anglicised..... It is noticeable that since its early days as a political party the cultural content of the SNP programme has become less obvious and the emphasis upon economic and social matters greater. The nationalist political party, which began by stressing on every occasion the unique cultural identity of the Scottish people, now assumes it. No longer does it need to be stressed, but merely mobilised from time to time." (Webb 1978)

Smith, in the passage already quoted in 3.3.3.2, makes

much the same point.

5.3.1.1 The themes which most clearly reflect cultural factors in the graphs are 'Cultural Protection', 'Historical Justification' and 'Symbolic Injustice'. If the hypothesis were correct, the measures of centrality for these themes should show a decrease as between the early and the late documents. The relevant measures are shown in Table 5.1. (p.80).

Clearly the first two themes have declined on all the measures (except that the degree of 'Historical Justification' has marginally increased. This is offset by the clear drop in rank degree). In both cases the rank degree is always higher than the rank salience, suggesting that, even given the decline, the penetration of the theme is greater than its salience would suggest.

The third theme, Cultural Protection, behaved differently. The theme had become more central. For this theme the hypothesis fails and requires modification; either the definition of cultural factors has to be tightened or the existence of different types of cultural factor has to be acknowledged, which respond differently to changes in party status.

It should be noted that in neither group was this theme particularly central. What has happened is that a peripheral theme has become slightly less peripheral, a result which could be due to a change in cultural emphasis, as suggested above, or could simply be a blurring round the edges.

5.3.2 Dilution

In a previous paper (Farbey, Mitchell and Webb 1979) the

hypothesis was proposed that the ideology of a party would become more dilute as it grows in size. Dilution is defined here as weakening of a few principles which structure the nature of responses to other issues and problems.

This too, can be interpreted as a problem in the centrality of particular themes or groups of themes. The hypothesis suggests the persistence of a small number of themes at the centre of the ideology which have low eccentricity and rank degree.

5.3.2.1 Looking first at the individual themes in the two groups, the most central themes on rank degree and also on rank salience are shown in Table 5.2.

Certain themes appear in all four lists at rank greater than 5. These are 'Critique' and 'Devolution'. Clearly there has been no dilution with respect to these two themes. It is interesting to note that these are abstract themes, as opposed to, say, 'Foreign Trade'. A third theme, of a similarly abstract nature is 'Referencing' which is again central on rank degree in both the early and late sets and only slightly less so on rank salience.

(note: it may be that referencing and criticism have, suis generis, high degree, in that referencing, i.e. comparing oneself to others, and criticising are usually done in relation to something concrete, wage levels for example. In order to be certain that the observed high degree was peculiar to the SNP, a comparison has to be made with documents from another party. This has been done to a limited extent in section 5.7).

These three, non-specific themes 'Critique..' 'Devolution' and 'Referencing' then, do persist as between the early and the late documents. They have not been diluted.

5.3.2.2 The more specific themes have, in general, not persisted as strongly. Some, such as 'Foreign trade' have been central in the early documents only to lose prominence in the later ones. Others, such as 'Oil/energy..' and 'Quality of life' show the opposite tendency.

5.3.2.3 One specific theme, however, shows no sign of dilution. This is the theme 'Organisational structure of a communally based organisation'. It is the theme concerned with promoting the party itself and the way in which the party is structured and is an organisational rather than a political 'principle'. Its continued prominence suggests that a part, an increasing part, of the SNP response is organisational and propagandist. This observation ties in well with Webb's (1978) and other comments on the importance of organisation in the growth of the SNP.

5.3.3. Summary

In summary, the central themes of early and late SNP documents appear to be of two kinds. There are the abstract, non-specific themes of 'Critique..' 'Referencing' and 'Devolution' and the 'organisational' principle reflected in the theme 'Organisation/structure..', which persist, forming the kernel of SNP ideology. There are too, more specific themes, which move in and out of the centre. If an identification is made between the former and the unchanging principles required by the definition of dilution, then the hypothesis is refuted, and,

on the basis of these results one would expect these themes to persist through the four consecutive data points which will be discussed in section 5.5.

5.3.4. Complexity

Both the early and the late documents included all the themes. In the later documents one theme, 'English reaction' was mentioned only once, and that on its own. The graph therefore had, formally, three components: one consisting of 26 (out of 27) of the content themes, all closely interrelated (diameter 1, radius 2) and 2 isolated vertices, 'English reaction' and 'No Content'. The early graph consisted of just 2 components: the 27 content themes (radius 2, diameter 2) and 'No Content'. Neither the number of themes or the number of components could in this case be used to test complexity.

The average degree increased from 11.5 to 15.9 as between early and late SNP. The ideology as a whole became more intricate over the period. On average, each theme in the late group had become associated with about 4 more themes than in the early. Figure 5.1 shows the change for each theme. From the Figure it will be seen that the change was not uniform, some themes such as 'Oil/energy' showing a marked increase.

The other measures of complexity must all point to an increase in complexity (this follows from an increase in the number of edges on a connected graph). There is therefore confirmatory evidence for the hypothesis.

5.4 SNP "early" v. "late" - the reduced graphs

The full graphs of the previous section contain too many vertices and edges to be sensibly compared as pictures. For

this reason, and also to give some weight to the very different frequencies associated with each edge, the graphs were reduced. In the reduced graphs, only those edges appear for which the frequency is greater than or equal to the mean frequency plus three standard deviations.

The resulting graphs emphasised only the most significant connections and were not used directly to test the hypotheses. They are described below.

5.4.1 "Early" SNP

The graph in Figure 5.2 has been drawn to emphasise the different positions of cultural and economic themes in the reduced structure. Broadly, the economic themes lie on the left-hand side and are interconnected, the cultural themes lie to the right. A path from one to the other necessarily involves the themes 'Devolution' 'Referencing' or 'Critique..'. These themes are central on all the measures: they are at the centre of the graph (eccentricity 2 except for 'Devolution' 3) they have high degree and they form a cut-set. Removing these vertices would produce 5 components, a markedly 'economic' component consisting of 'Agriculture', 'Foreign Trade', 'Economy', 'Industry' and 'Employment'; a cultural one consisting of 'Symbolic Injustice' and 'Historical Justification' as well as three isolated vertices 'Organisation/structure', 'Housing and Local Government'.

A similar effect is obtained by removing the cut set 'Economy', 'Referencing' and 'Critique..'.

Taken together these results confirm those of section 5.4.1.1, that 'Critique..' and 'Referencing' are critical to

the early SNP ideology and with 'Economy', 'Devolution' and 'Organisation/structure' form the kernel of that ideology. They are, too, very like the founding resolution of the S.H.R.A. (section 3.2.2) in which all these themes can be discovered.

5.4.2 "Late" SNP

The graph for the later documents, Figure 5.3, is very different. The themes appear to radiate from the two themes 'Critique..' and 'Devolution' together with 'Organisation/ structure..' and 'Oil/energy'. Removing 'Critique..' and 'Oil/energy' from the graph would disconnect it, although no very obvious subject matter connects the vertices of the resulting components. The cultural themes have, except for 'Symbolic Injustice', disappeared, which reinforces the results on the full graph in that cultural themes appear to have declined in importance and, possibly, that there may be different types of cultural theme.

The emergence of 'Oil/energy' as a critical theme is interesting in so far as it demonstrates the contingent nature of an idea-system. There was nothing in the early idea-system that could have indicated that a specific theme, such as 'Oil/energy', would move dramatically into the centre of the system. This movement is a response to an external event - the discovery of Oil in the North Sea.

North Sea Oil, or rather the possibility of economic self-sufficiency which it promised, provides an alternative 'raison d'etre' to impatience with the government at Westminster for the SNP. If either is removed from the graph in Figure 5.3,

the graph remains connected. If both are removed the internal logic of the SNP idea-system falls apart.

5.4.3 The "early" and "late" adjacency matrices

The adjacency matrices of the two graphs are shown in Figures 5.4 and 5.5, together with the eccentricity.

The centre of the early graph consisted of the vertices 'Critique..' 'Referencing' and 'Economy'. The vertices rank first, second and seventh on rank degree and also form a cutset. There was a second cutset, 'Critique..' 'Referencing' and 'Devolution' of rank degree 1,2 and 3 respectively.

The second graph had a larger centre, containing 8 themes of eccentricity 2, including the four just mentioned, together with 'Agriculture' 'Foreign Relations' 'Oil/energy' and 'Organisation/Structure'. On rank degree 'Critique..' dominated, followed by 'Devolution' and 'Organisation/structure'.

The second graph is more complex than the first, 29 edges on 13 vertices as opposed to 25 edges on 13 vertices.

These measures reinforce the findings from the full graph and the impression gained from the pictures.

5.4.4 Flexibility

The third aspect of the graph theoretic measures described in 2.4.6 was that of flexibility. Because the reduced graphs are not overconstrained as are the full graphs, it is possible to demonstrate the use of the concepts on these graphs.

Table 5.3 shows the spatial and planar mobilities for each graph. The early SNP is more mobile on both measures, although both graphs are, according to the measures, overconstrained in

the plane. Looking at the graphs, however, it would seem that the measures are misleading since each of the leaves, 'Housing' in the early graph and 'Education' in the late, can move, even in the plane. The overall mobility measure has been affected by the high degree of overconstraint elsewhere in the graph. Mobility, as it is defined, appears to give a global measure, where, perhaps, a more local measure would be more instructive.

Removing the leaves, for the sake of demonstration, leads to planar mobilities of -6 and -8 in the early and late data respectively. This shows the remaining graphs to be overconstrained, i.e. several edges could be dropped without affecting the rigidity of the structure.

Removing some of the highly connected vertices, again purely for demonstration, has an interesting interpretation. For example, if the vertex 'Critique..' is removed from the early graph, Figure 5.6, the resulting graph has planar mobility 5, an increase of 7 on the original and one which is not only due to the presence of the leaves and the (now) isolated vertices. There is a non-rigid cycle 'Economy-Devolution-Referencing-Foreign Trade-Economy'.

By contrast, removing the theme 'Devolution' leaves the rest of the graph still triangulated, Figure 5.7, and rigid. Now, although the mobility does increase by 4, this is accounted for by the (now) isolated vertex and the leaves.

Thus it might be argued that 'Critique..' has a more fundamental role to play in underpinning the idea-system than does 'Devolution'.

The same effect can be observed in the late graph,

Figures 5.8 and 5.9. It is tempting to conclude that in these two periods at least the idea-system of the SNP was held stable more by the negative aspect of 'Critique..' rather than the positive one of 'Devolution'. This would certainly be consistent with the numerical measures. However it must be stressed that, of all the measures, mobility requires a number of additional assumptions, or analogies, for its interpretation (c/f section 2.4.6) and, as we have seen, reflects a global rather than a local mobility.

As things stand, mobility can only provide a starting point for speculation.

5.5. The four consecutive SNP groups - the full graphs

The next collection of documents to be examined were the four consecutive groups of SNP documents described in Section 4.2.1.2. These were SNP 1, the very earliest documents available, ranging from 1896 to 1940, and had been abstracted from the 'early' set. The second group, SNP 2, consisted of the remainder of the 'early' set plus those which had been assigned to this period. The third set, SNP 3, ranged from 1960 to 1974 and consisted mainly of the 'late' set plus some assigned documents. The last group, SNP 4, consisted of documents from 1974 to 1979.

The analysis followed the same pattern as that for the 'early' and 'late' documents, each of the preliminary hypotheses being considered again for the new and re-organised data.

5.5.1 The role of cultural factors

Table 5.4 shows the analysis of centrality for the

cultural themes.

For the first two themes 'Historical Justification' and 'Symbolic Injustice' the results showed the same pattern as before. The themes have declined steadily on both rank degree and rank salience. 'Historical Justification' has declined more sharply than 'Symbolic Injustice' - from 4th or 5th on both measures to virtually last. The decline in the theme has been steady on rank degree, rank salience and salience, although, as with the 'early' v. 'late' groups, rank degree is always higher than rank salience. The drop in rank salience comes earlier than in rank degree, suggesting that the change in structure has taken longer to accomplish than the surface change, although following the same general pattern.

'Symbolic Injustice' behaves in the same way, although the change is not as marked. In both cases the absolute degree varies erratically; some of this must be due to sample size, but this is unlikely to be the only explanation as the variation is not consistent as between the two themes.

The theme 'Cultural Protection' behaves differently from the other two as in the previous grouping. The increased centrality in the theme, noted in the earlier analysis, is not consistent over the four points. Cultural Protection is more central in SNP 1 and SNP 3 than in SNP 2 and SNP 4. (Figure 5.10) Nor is rank degree everywhere higher than rank salience, the latter is lower (i.e. the theme is more prominent) at SNP 3. Again the two measures vary together; in this case the rank degree varies less, as opposed to more slowly, than

rank salience, but the impression is still that structural change is more difficult than surface change.

5.5.2 Dilution

In Section 5.3.3, it was suggested that the ideology of the Scottish Nationalist Party could be seen as consisting of a persistent core of abstract, non-specific themes, namely 'Critique..', 'Devolution', 'Referencing' and 'Oranisational Structure', together with a number of specific themes which moved in and out of the centre.

When the longer time scale is considered, the pattern is not dissimilar. Table 5.5 shows the league table of degree and salience for the four data points.

Clearly 'Devolution' persists at the centre of the ideology, both on degree and salience. That is, 'Devolution' is always prominent both because of the number of times it is mentioned and in the degree of penetration through the whole ideology. Almost as central is 'Critique..' which is in the top five themes on both measures, with the notable exception of salience in SNP 4. The sudden disappearance of the theme in the later documents does require explanation, especially as it is not echoed in the SNP 4 rank degree. Several explanations are possible, the first being that it is simply a statistical aberration, a chance fluctuation. This seems unlikely in the face of an absolute drop from 12.4 to 2.6% in frequency of mention. Alternatively there may have been a conscious decision to play down the critical aspect, perhaps for stylistic or propaganda reasons, or because the party has genuinely become more concerned about its own policies and programme than about

slamming Westminster, or even because after the 1974 elections the SNP was itself a force in Westminster. After 1974, for a significant section of articulators, both their perspective and the problems to which they had to respond had changed.

However, without a further data point, and in the absence of a theoretical explanation, there is no conclusive evidence for any of these possible explanations. From the point of view of the methodology the interesting point is the contrast between salience and degree for an important theme.

'Referencing' can be seen in Table 5.5 to be a declining theme, on both counts. As with 'Historical Justification' and 'Symbolic Injustice' in the previous section, the structural change is smaller than the surface change and lags behind it. 'Referencing' also ranks higher on degree than on salience at all four points. Like the two other themes, 'Referencing' permeates the ideology to an even greater extent than is suggested by its salience.

The theme 'Organisation/Structure', which, with the themes 'Referencing', 'Devolution' and 'Critique..', seemed to form the central core of the SNP idea-system on just the 'early' and 'late' sets, behaves erratically when all four points are considered. There is a marked divergence between salience and degree for this theme; once established as the central theme on rank salience in SNP 2, it continues to be first or second (giving way to No Content) through SNP 3 and SNP 4. On rank degree, by contrast, 'Organisation/Structure' is only just visible in SNP 2 (rank 10), becoming prominent in SNP 3 (equal first with Critique..) as observed in the 'early' v. 'late'

analysis, but then declining in SNP 4 to roughly its previous level. It has been displaced by the more specific themes such as 'Oil' and 'Industry' and does not permeate the ideology to anything like the extent indicated by both salience and degree in SNP 3.

Considering all the four themes together, the evidence for the dilution hypothesis would seem to be mixed. Both 'Critique..' and 'Organisation/Structure' would seem to have been partially diluted over the period, one in salience and one in rank degree. 'Referencing' is clearly declining. 'Devolution', however, has held its place. Thus, if the ideology has in general been diluted, it has not been on the particular issue which defines the party.

5.5.3 Complexity

As with the previous analysis, neither the number of vertices nor the number of components provided any significant information about the complexity of the ideology. There had been an increase in the number of vertices from SNP 1 to SNP 2 (see Table 5.6) but this could well have been attributable to the small sample.

The average degree, too, will have been to some extent influenced by sample size. Nevertheless, it is surprisingly low at SNP 1. The marked increase from SNP 1 to SNP 2 does suggest a real development of the ideology between the two periods. It is a development of the intricacy of the relations between ideas and one which continues from SNP 2 to SNP 3. By SNP 4 there has been a change in direction - the ideology has apparently become simpler, each theme being related to, on

average, four fewer themes. On this measure the hypothesis of an increasing complexity is refuted and needs to be reformulated at the least.

5.6. The four consecutive SNP groups - the reduced graphs

The reduced graphs SNP 1, SNP 2, SNP 3 and SNP 4 are shown in Figure 5.11 (a,b,c,d respectively). These graphs are compared below.

5.6.1 The graphs are different to look at. For example, SNP 1 is simpler than SNP 2, having both fewer vertices and fewer edges. SNP 3 is also simpler than SNP 2, being a star graph with 'Criticism' at the centre. SNP 4 is different again, with four components and a large number of vertices. In SNP 4 too, 'Critique..' has shifted away from the centre of the graph.

5.6.2 The role of cultural factors

The graphs confirm the previous results on this hypothesis. The three cultural factors 'Cultural Protection', 'Historical Justification' and 'Symbolic Injustice' all appear in SNP 1, but only 'Symbolic Injustice' appears in the later reduced graphs. By contrast the economic themes 'Economy' 'Agriculture' and 'Industry' maintain or even improve their position through the four data points.

5.6.3 Dilution

The four themes which persist through all four reduced graphs are 'Devolution', 'Economy', 'Referencing' and 'Criticism'. Of these only 'Devolution' is consistently central (i.e. one step away from the centre), and therefore remains undiluted.

'Economy' shifts from the periphery towards the centre over

the four periods. This result conflicts with both the full graph result and the earlier 'early' versus 'late' comparison. There is, therefore some doubt as to the inclusion of this theme in the kernel of SNP ideology and whether it is, or is not, diluted.

'Referencing' has shifted from the centre towards the periphery and in SNP 4 is not in the major component. This is consistent with the full graph result and shows a gradual dilution.

'Critique..' is at the centre of the first three graphs but not of SNP 4. The same result is obtained if rank degree is considered as the measure of centrality. This result contrasts with the result on the full graph.. In the latter 'Critique..' maintains its centrality as measured by rank degree. However, in the full graph there is a drop in rank salience over the four periods. It would seem that 'Criticism' remains widespread (i.e. connected to many other themes), but thinly spread, only one association being mentioned sufficiently frequently to appear on the reduced graph. The criticism has been muted; no longer hammered home at every possible opportunity. The web is intact, but finely spun. In the sense of dilution as it has been used here, 'Critique..' has not been diluted.

Only one edge persists through all four reduced graphs - the edge 'Devolution' - 'Critique..'. There has been no dilution of this particular link. It has been the key connection in all the document sets, reflecting again (c/f 5.4.1) the sentiment of the founding meeting of the S.H.R.A. that 'government from London is inefficient'. (3.2.2).

One edge appears in three of the graphs, 'Critique..' - 'Referencing'. In other words the Scottish situation is seen as being unsatisfactory particularly by comparison with other nations. This link is lost in the last document accentuating both the decline of 'Referencing' and the thinning of 'Critique..'.

5.6.4 Complexity

On the average degree, the reduced graphs show the similar pattern as the full graphs, i.e. the earliest and latest graphs are simpler than the middle two, although the variation is smaller. (Table 5.7). That apart, the results are not consistent and show no pattern.

5.6.5 'Early' v. SNP 2

A comparison of the 'early' data with SNP 2 is interesting in that the two sets of documents are subsantially the same so that the comparison will serve as an indication of the robustness of the model to a small change in the data-base. Since SNP 1 is a sub-set of the 'early' set, comprising the earliest documents of the 'early' set, the comparison will also highlight possible differences between the pre-war SNP 1 and later SNP 2.

Looking first at the two reduced graphs, they are not unlike (Figures 5.2 and 5.11b). SNP 2 is a sub-graph of 'early', i.e. the addition of new documents has not added any significant new links, but the removal of earlier documents may have resulted in the removal of some significant links. Moreover, the vertices have, in general, retained their place in the structure; for example, 'Critique..' and 'Economy' are central to both (eccentricity). The most notable difference

is the absence of 'Historical Justification' in the SNP 2 data. Coupled with the appearance of 'Historical Justification' in the SNP 1 data, this suggests strongly that this cultural aspect was a feature of the earliest SNP ideology and, unlike 'Symbolic Injustice' it had begun to decline as early as 1940. It is probably significant that the SNP was only formed under that name in 1934 - it is possible that the particular form of cultural reference changed with the composition of the party.

5.7. The comparison of SNP with Conservative documents

A small number of Conservative party documents had been included in the data set to provide a rough test of the power of the methodology to discriminate between documents known, a priori, to be from different sources.

5.7.1 There were 19 Conservative documents in all, ranging in date from 1971 to 1974 and providing a total of 399 paragraphs. 22 out of 28 themes were coded and the resulting graph is shown in Figure 5.12. The missing themes (i.e. those not mentioned at all in this group) were 'Historical Justification','Symbolic Injustice', 'English Reaction', 'Foreign Trade' and 'Post Independence Constitution'.

The complete absence of the two cultural themes contrasts strongly with the SNP documents and, obviously, reflects the Nationalist basis of the SNP. Or perhaps the Scottish Conservative party finds 'Elizabeth the First and Second' too much of a mouthful.

5.7.1.1 The graph consists of 23 vertices and 6 components. The major component has radius 2 and diameter 4; it is less centralised than are the SNP graphs which have radius 1 and

diameter 2. It is nearly planar.

5.7.1.2 The vertex 'Economy' is at the centre of the major component. It is central (uniquely) in eccentricity, it is highest on rank degree and, with 'Critique..' forms a non-trivial cutset - splitting the component 'Referencing' - 'Legal' - 'Defense' and Foreign Relations from the rest of the graph. Strictly 'Social Services' 'Critique..' 'Quality of Life' and 'Economy' are all cut-points in that their removal of any one of these vertices would disconnect the graph. The resulting components are not very interesting.

In his discussion of the political background to the rise of the national movement in Scotland, Brand (1978), writing of the Scottish Conservative party soon after the war, say 'There is little evidence that Scotland's economic situation is even recognised'. Clearly by the early 1970's this was no longer true. A Scottish party is presenting an Economy-based argument to a Scottish electorate. What may be true, and this would not appear on the classification system used, is that the 'Economy' under discussion is a 'British' rather than a specifically Scottish one. In the documents we have, there are references to both.

5.7.1.3 In the graph of the Conservative party documents, the vertex 'Devolution' is a leaf, attached to the rest of the ideology only through its connection with 'Economy' and achieving what centrality it has through that connection. It plays little or no part in the structure - removing the vertex would leave the rest of the structure intact.

As one would expect, this is very different from the role

of 'Devolution' in the SNP data.

There is a marked difference too in the salience and the other measures for 'Devolution'. It will be seen from Table 5.8 that 'Devolution' ranks 4th on salience. That is, 'Devolution' is much talked about, but largely in its own terms and has not been related to other themes. Reading the documents it is clear that 'Devolution' is considered primarily as a constitutional issue, not immediately related to the rest of the system, but only indirectly through the possible effect on the economy, as in the paragraph:

"Decisions affecting Scotland will be taken in Scotland. A Scottish Budget will be decided by the Secretary of State for Scotland with the Assembly. Scotland's housekeeping will be Scotland's business." (Doc.068)

5.7.1.4 This tendency to work particular themes out in their own terms, rather than by relating them to other themes, is apparent in the larger number of components and the lower average degree of the Scottish Conservative documents. This is a much simpler structure than that of the SNP, reflecting presumably the pragmatic, reactive, non-rationalist view of the Conservative party (Farbey, Mitchell and Webb 1979).

A second explanation, suggested by Mitchell (private communication) might be that the Scottish Conservative documents are fragments of a larger philosophy, being bits of a general philosophy adapted to Scotland.

A third explanation is possible (see 8.4.2). The Scottish Conservative party is a part of a much larger party, appealing to a much wider population. In such parties one would expect

to find a simpler level of presentation (though not necessarily of the full ideology), precisely because of the mass audience.

5.7.1.5 A further feature of the Conservative ideology is the very high proportion of paragraphs coded as 'No Content'. These paragraphs are largely headings, sub-headings and end of document exhortations. 13, (3% of the total) are empty phrases e.g. 'That's worth having' or 'Ahead of course'.

This is partly a question of presentation - the documents are broken up by using more headings and sub-headings and partly a question of the ideology having a greater ratio of non-policy to policy content. For example, the slogan 'For all the people we must win freedom' is empty of policy content but it is not an empty ideological phrase.

It may be that here, too, the need to appeal to a mass audience necessitates a simpler, less content-bound approach. It is particularly noticeable that the SNP, too, as it has adjusted to the requirements of a mass audience, has increased the No Content category as shown in Table 5.9 and, as has been shown, the reduced graph of SNP 4 is simpler and has more components than the earlier SNP graphs.

5.7.2 The preceding discussion has described the structure of the Conservative ideology and noted particular features of the structure which differ from the SNP graphs.

In order to compare whole structures the rank correlation coefficients for each pair of idea-systems studied, were calculated with the results shown in Table 5.10 and Table 5.11. Table 5.10 shows the rank correlation co-efficients for degree,

Table 5.11 for salience.

Considering first the contrast between all SNP and the Conservative groups, both tables show much lower co-efficients for each of the SNP groups than that achieved between SNP groups. (i.e. an SNP group is in general more like another SNP group than it is like a Conservative group). The contrast is sharper for rank degree than it is for rank salience, suggesting that rank degree is a better discriminant between the two groups. This latter finding has been consistent through all the comparisons that were made and was used as a basis in the attempt assignment of undated documents to their correct group. (see 5.8.3).

5.7.2.1 In both degree and salience, there has been a steady increase in the rank correlation with the Conservative documents over the four periods of the SNP data (with the exception of the SNP 2 degree). That is, the allocation of attention and cross-referencing in the SNP ideology grows consistently more like that of the Conservative ideology.(1971-1974). This finding is further confirmatory evidence for the hypothesis (5.3) that there has been a shift in the Nationalist idea-system from cultural to social and economic issues; a shift which can now be seen partly as a result of the need to react to a strongly economy-centred party and partly as a result of the need to communicate with a mass audience as the party assumes mass party status (both points are discussed extensively in Farbey, Mitchell and Webb 1979 and 1980). A further structural explanation is discussed in 5.9.4.

5.7.2.2 In addition to the comparison between Conservative

and SNP groups, Tables 5.10 and 5.11 show the rank correlation coefficients which have been used to compare the different SNP groups.

From these tables it will be seen that there is no progression over time within the SNP groups, rather, the highest co-efficients are in the SNP 3 column, indicating that the development of the ideology had reached something of a watershed at that time. SNP 3 contains both the structures of its past and the seeds of its future development. The past and the present, however, are not alike. SNP 4 is much closer to Conservative ideology (in structure, not necessarily in content) than it is like SNP 1.

The pattern on degree and on salience is very similar, although, as with other measures (5.5.1) the salience measure responds more quickly. SNP 4 for example, is still highly correlated with SNP 3 (0.75) on degree, less so (0.67) on salience. In fact, on salience there is very little difference between the co-efficients for SNP 4 - SNP 3 and SNP 4 - Conservative. SNP 4 documents would seem to be as similar in surface content to both SNP 3 and Conservative documents. In terms of their structure, however, they are much more closely related to the SNP 3 group than to the Conservative group. In this case the structural measure has detected a continuity in a way which would not have been possible using only salience measures.

5.8. Recognition of undated documents

Ten documents of the original set were undated and had therefore been excluded from the 'early' versus 'late'

comparison. Of these ten most contained some textual material (references to people or events) that allowed them to be assigned to each period with some degree of confidence. The remainder contained no such evidence.

An attempt was made to assign each document to either the 'early' or 'late' group, using the graph-theoretic features and using the rank correlations between the main body of documents and each of the longer individual documents, on salience and degree.

The attempt was only partly successful. Long documents can usually be correctly assigned on the basis of rank degree, particularly those which covered a range of subjects. Short documents, or single subject documents, were more difficult to assign.

5.8.1 A list of the documents, together with the assignment (if any) based on textual reference, and a table of graph-theoretic features which seemed most likely to discriminate between the groups is given in Table 5.12.

It is immediately apparent that most of these measures could not be used because the vertex in question was not part of the graph. Only average degree was obtainable for all the documents and it does show a marked increase as between the two late documents (049 and 033) and the others. It is interesting to note that doc. 003, which has high average degree, is dated in Hanham (1969) as issued in 1946, but was in fact current until much later. (It is possible that the version of doc 003 is later than Hanhams as it contains a reference to Israel i.e. is presumably written after 1948. Its title is "Policy

of the Scottish National Party" and it may have gone through several updates. Hanham reproduces only a part of the document.)

The only other measure of this set obtainable on all but one of the documents, was rank degree 'Devolution'. Since 'Devolution' is of continuing importance throughout the period it would have been a little surprising if the 'late' document had been substantially different on this measure, and, indeed at rank 6, it is well within the range of the early documents.

5.8.2 For the two longest documents in the set, 003 and 033, it was possible to estimate the rank correlation co-efficients for salience and for degree. Where a vertex was not mentioned the ranking was taken as 28 or 27 (the lowest possible ranking on salience and degree respectively). Table 5.13 shows the results of these calculations.

5.8.2.1 Considering rank salience first. For document 003, the correlation co-efficients were 0.28 and 0.24 for 'early' and 'late' respectively, and 0.27 when compared with the Conservative party documents. None of these is significant and there is not a great deal of difference between them. On this evidence, rank salience is not a good discriminator either between parties or between the same party at different times.

The coefficients for rank degree were 0.66 and 0.64 for 'early' and 'late' documents and 0.31 for Conservative documents. In this case the coefficients showed a clear difference between the two parties but still gave no clear indication of the period.

5.8.2.2 A similar situation obtained for document 033. In this case the rank degree correlation seemed to indicate that the document was almost certainly SNP rather than Conservative, and possibly earlier rather than later. Since 033 was known a priori to be a late document the latter conclusion was false.

On reading the document one possible reason emerged. Unlike most of the other late documents, 033 made very little mention of North Sea Oil. It was a document concerned only with the development of Tayside and not a general statement of policy. When the vertex was dropped from both graphs the resulting coefficients were very similar (with 'early' still slightly larger) again suggesting that rank degree, whilst being an effective discriminant between parties, was not effective within party.

5.8.2.3 Finally, as a check on the findings of the analysis, a document which was known to be 'early' (doc 019) was treated in the same way with the results shown in the last line of Table 5.12. In this case, the proper assignment was clear from the co-efficients, both as to period and as to party and again, this assignment could not have been made on the basis of rank salience.

5.8.3 From this analysis it would be reasonable to conclude that the rank degree correlation coefficient is useful in assigning documents to sources which are themselves markedly different. To do this, however, the documents must be of a sufficient length to (possibly) include most of the vertices and also of sufficient scope to (possibly) include all the vertices. A short document, on a few themes, cannot be

assigned in this way.

5.8.3.1 It is of interest to note that doc. 003, was part of SNP policy through both periods, although it clearly originated in the early years. It may be that the similarity in the two coefficients is real, that is, the document reflects continuing SNP policy and is therefore similar in structure to both groups, although not in the same way. The result for doc 019 would support this view as the co-efficients are much more clearly strung out.

To test this hypothesis would require another, independent set of documents and this was not available.

5.8.3.2 It is also of interest to realise that assigning a document on rank degree implies that it is the number of themes to which a particular theme has been related that is significant - not the actual themes to which it has been related, a structural not a content feature.

5.9. The dynamics of idea-systems

The graphs derived from the documents have been used in the preceding sections to derive substantive results and to test substantive hypotheses. In this section the results are used to derive structural (c/f 5.2.3) results, i.e. results about the system qua system, and to generate structural hypotheses to be tested on other idea-systems.

The structural results which have been observed have been divided into five groups

i) those concerned with the differing roles played by specific as opposed to abstract themes

ii) those concerned with the difference between structure

and salience

iii) those concerned with mechanisms of dilution

iv) those concerned with the changes in complexity

v) those concerned with the development of a single document in terms of flexibility

These are discussed in turn below.

5.9.1 The themes which persisted through the four SNP graphs, were, in general non-specific, high-level themes such as 'Critique..' and 'Referencing'. Specific themes such as 'Oil/energy' were more erratic.

As a starting point in the study of a (further) group of documents it would be reasonable to hypothesise two kinds of theme - high-level, abstract and generally stable themes and specific, low-level, non-stable themes. The former would form the principles (strategic (?) themes of the idea-system, the latter the policy (tactical (?)) themes.

5.9.2 In sections 5.5 and 5.7 it was shown that 'centrality' as measured by salience and 'centrality' in structure were not necessarily the same. Furthermore the structural measures changed more slowly than did the salience ones. This suggests the hypothesis that structural change is more difficult to accomplish than a change in salience.

Certainly the pattern of the SNP documents was that of a slowly-changing central pattern coupled with changes in the periphery of the idea-system and changes due to contingent events, for example, the discovery of North Sea Oil. This may well be a pattern that is repeated in other idea-systems.

5.9.3 In section 5.6.3 the theme 'Critique..' provided an example of a particular manner of decline, in which a theme, although remaining widely connected, appears to be spread thinner and thinner until, in some cases, the links actually snap. This too could prove a model for other systems.

5.9.4 The full graph complexity decreased unexpectedly in SNP 4. An explanation for this phenomenon might be found in the work of Dougherty (1981) and others. This work shows that the most immediate level of recall (the term actually used is 'salience') of different biological categories depended on the relation of the classifier to the domain of classification. Thus, for example, an urban American was more likely to classify plants on a high level ('that is a tree') rather than on a generic level ('that is an ash, oak, sycamore etc.).

By inverting the role of the classifier to that of the presenter it can be hypothesized that, as the SNP presented itself to a wider audience, many of whom do not share the presenter's relation to the domain, the salient level of presentation has to be higher; the ideology has to look simpler, even though the underlying structure may not have changed. This could also explain the apparent simplicity of the Conservative party documents, as was remarked in 5.7.1.4. Much the same phenomenon is described in Hofstadter (1979 p.373) in his attack on the problem of isomorphism between two semantic networks.

5.9.5 Finally, when considering a single document, it was found that (2.4.6.4) it was reasonable to postulate a tendency

towards triangulation, or at least rigidity, in a document and that an idea-system might atrophy or develop around flexible areas in the graph.

These hypotheses, in common with all those mentioned in the section will have to be justified in the substantive domain.

5.10. Summary

In this Chapter the results of applying the method to a substantial body of data have been presented.

The data were divided into three sets for comparison. These were

i) the original SNP data set was divided into 'early' and 'late' groups.

ii) an extended SNP data set was divided into four groups to give four consecutive points

iii) each of these groups was compared with a group of documents from the Scottish Conservative party

5.10.1 It was found that the 'early' and 'late' groups did produce different graphs.

Three hypotheses, concerning the position of cultural themes, the dilution of key themes and the complexity of the idea-system were examined. In these comparisons there was confirmatory evidence for all the hypotheses.

Within this general agreement it was noted that cultural themes may not all behave in the same way and that some tightening of the first hypothesis was required.

The graphs were reduced by considering only the most significant connections. The measures obtained on the reduced

graphs confirmed the findings on the full graphs.

On the reduced graphs it was possible to consider the notion of flexibility of a graph. The measures of flexibility were inherently less plausible than the other measures and less satisfactory in that they provided local rather than global measures on the graph. Nevertheless they proved a useful source of hypotheses about the nature of the development of an idea-system.

5.10.2 The four consecutive data points were examined next. The results on the role of cultural factors confirmed those obtained with the early and late groups. In addition it was noted that change in structure was slower than change in salience.

The evidence for the hypothesis on the dilution of key themes was mixed, 'Devolution' being the only theme to show no sign of dilution through all four periods.

The analysis of complexity gave mixed results. There seemed to be an increasing complexity over the first three periods, followed by a drop in complexity in the last.

Again the reduced graphs echoed the full graphs.

A comparison of the 'early' graph with the SNP 2 graph showed a difference between the cultural factors 'Historical Justification' and 'Symbolic Injustice', with the former showing evidence of a very early decline. In general it seemed that the addition of new documents did not necessarily mean the addition of new links, i.e. the system was relatively robust.

5.10.3 The comparison with the Scottish Conservative document

graphs showed quite clearly that the two idea-systems were very different. The Conservative documents were more fragmented, themes tended to be worked out in their own terms rather than related to other themes and the proportion of the 'No content' theme was high.

5.10.4 The assignment of undated documents was attempted with mixed success. The longer documents could be placed accurately using rank degree correlation. The shorter ones, or single subject ones, were difficult to assign.

5.10.5 Finally the possibility of the model providing an empirical account of the dynamics of idea-systems was examined and five separate areas were identified for further study with other document sets.

Theme: Historical Justification

Measure	Early SNP	Late SNP
Eccentricity	2	2
Degree	15	16
Rank Degree	7.5	14
Salience	2.9	0.9
Rank Salience	13	24
Centre	Yes	No
Cutpoint	No	No

Theme: Symbolic Injustice

Measure	Early SNP	Late SNP
Eccentricity	2	2
Degree	19	19
Rank Degree	4	9
Salience	4.1	2.4
Rank Salience	11	13.5
Centre	Yes	No
Cutpoint	No	No

Theme: Cultural Protection

Measure	Early SNP	Late SNP
Eccentricity	2	2
Degree	6	13
Rank Degree	23	19
Salience	0.3	2.0
Rank Salience	27	15
Centre	Yes	No
Cutpoint	No	No

TABLE 5.1

Measures of Centrality - themes reflecting Cultural factors

Rank Salience			Rank Degree	
	Early	Late	Early	Late
1	Criticism	Org./Structure	Devolution	Criticism
2	Devolution	Criticism	Criticism	Org./Structure
3	Foreign Trade	No Content	Referencing	Devolution
4	Economy	Devolution	Symbolic Injustice	Quality of Life
5	No Content	Oil/Energy	Industry, Economy	Referencing
6	Referencing	Agriculture		Agriculture
7	Org./Structure	Referencing	Foreign Trade	Economy
8	Industry	Economy	Historical Just.	Oil/Energy
9	Local Govt.	Industry	Housing	Symbolic Inj.
10	Agriculture	Education, Employment, Foreign Relations	Org./Structure	Employment, Industry

TABLE 5.2

Central Themes

	Early SNP	Late SNP
No. of Vertices	13	13
No. of Edges	25	29
M_p (Planar Mobility)	-2	-6
M_s (Spatial Mobility)	11	7

TABLE 5.3

Comparative Mobility Reduced Graphs

Theme: Historical Justification

Measure	SNP 1	SNP 2	SNP 3	SNP 4
Eccentricity	4	2	2	3
Degree	4	16	17	1
Rank Degree	4.5	7.5	14.5	27
Salience	6.4	2.5	1.1	0.1
Rank Salience	5	16	2.1	28
Centre	No	Yes	No	No
Cutpoint	No	No	No	No

Theme: Symbolic Injustice

Measure	SNP 1	SNP 2	SNP 3	SNP 4
Eccentricity	4	2	2	2
Degree	4	20	19	11
Rank Degree	4.5	5	9.5	15
Salience	3.2	3.5	2.2	1.1
Rank Salience	11	12	14	21.5
Centre	No	Yes	No	Yes
Cutpoint	No	No	No	No

Theme: Cultural Protection

Measure	SNP 1	SNP 2	SNP 3	SNP 4
Eccentricity	5	2	2	3
Degree	1	6	13	2
Rank Degree	14.5	24	20	26
Salience	1.6	0.4	1.9	0.5
Rank Salience	18	26.5	15	26
Centre	No	Yes	No	No
Cutpoint	No	No	No	No

TABLE 5.4

Measures of Centrality - themes reflecting Cultural factors, four consecutive data points

	Rank Salience				*Rank Degree*			
	SNP 1	*SNP 2*	*SNP 3*	*SNP 4*	*SNP 1*	*SNP 2*	*SNP 3*	*SNP 4*
1	*Devolution*	*Org.Struct.*	*Org.Struct.*	*N/C*	Devol.	Devol.	Org.Struct.	Devol.
2	N/C	Devol.	Crit.	Org.Struct.	Refer.	Crit.	Crit.	Crit.
3	Criticism	Criticism	N/C	Devol.	Crit.	Econ.	Devol.	Oil
4	Economy	Economy	Devol.	Industry	Hist. Just.	Ref.	Q.O.L.	Q.O.L.
5	Hist.Just.	For. Trade	Oil	Education	Symb. Inj.	Symb.Inj.	Ref.	Industry
					Economy			Economy
								Employment
6	Refer.	N/C	Agric.	Oil	Education	Industry	Econ.	
7	Q.O.L.	Ref.	Industry	Agric	Industry	For. Trade	Agric.	
8	Agric.	Industry	Ref.	Econ.	Q.O.L.	Hist. Just.	Oil	Agric.
9	Educ.	Agric.	Economy	Q.O.L		Agric.	Industry	Org.Struct.
10	Employ	Loc. Gov.	Employment	For. Rel.	For. Rel.	Housing	Symb.Inj.	Ref.
11	For^{n}Rel					Org.Struct.		
12	Symb.Inj							

TABLE 5.5

Degree and Salience

Period	No. of vertices	No. of components	Average Degree
SNP 1	22	4	2.4*
SNP 2	28	2	12.8
SNP 3	28	3	15.7
SNP 4	28	2	11.1

*Major component only 2.7

TABLE 5.6

Full Graph Complexity

Period	No. of vertices	No. of components	Average Degree
SNP 1	9	2	2
SNP 2	11	1	2.9
SNP 3	8	1	2
SNP 4	18	4	1.9

TABLE 5.7

Reduced Graph Complexity

Rank Degree	Rank Salience
1 Economy	1 No Content
Education	2 Economy
Employment	3 Oil/Energy
4 Industry	4 Devolution
Oil/Energy	5 Social Services
Quality of Life	

TABLE 5.8

Degree and Salience, Conservative

	CON	SNP 1	SNP 2	SNP 3	SNP 4
%	14	11.2	7.1	11.5	15.9
rank	1	2	6	3	1

TABLE 5.9

Salience of the Theme "No Content"

Degree

	SNP 1	SNP 2	SNP 3	SNP 4	CON
SNP 1	*	.65	.72	.45	.20
SNP 2	.65	*	.73	.58	0
SNP 3	.72	.73	*	.75	.41
SNP 4	.45	.58	.75	*	.51
CON	.20	0	.41	.51	*

TABLE 5.10

Rank Correlation Co-efficients (Degree)

Salience

	SNP 1	SNP 2	SNP 3	SNP 4	CON
SNP 1	*	.55	.72	.47	.35
SNP 2	.55	*	.72	.52	.46
SNP 3	.72	.72	*	.67	.58
SNP 4	.47	.52	.67	*	.64
CON	.35	.46	.58	.64	*

TABLE 5.11

Rank Correlation Co-efficients (Salience)

Doc. No.	Text Ref.	Eccentricity of Oil/Energy	Degree of Oil/Energy	Eccentricity of Historical Just.	Eccentricity of Foreign Trade	Eccentricity of Foreign Relation	Eccentricity of Quality of Life	Rank Degree of Quality of Life	Rank Degree of Oil/Energy	Rank Degree of Devolution	Average Degree	Assigned to E = early L - late
001	part 1 is replica of 1948 constitution	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	1.5	½	E
002	"hostilities in Europe" "Tom Johnston"	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1	1.6	E
003	in Hanham (1969) dated 1946 current till much later	3	8	4	4	5	4	5.5	2.5	5.5	3.7	?
009	"London Labour Government" last date 1946	in minor component	1	N/A	2	in minor component	N/A	N/A	N/A	10.5	2	E
018	housing figures 1918-1951	N/A	N/A	N/A	N/A	N/A	4	7	N/A	3	2	E
053	Scottish Congress 1950-1964	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	3	1.25	E (?)
067	"Commonwealth" "Imperial Conference"	4	1	6	6	iso-lated	iso-lated	17.5	11	1	1.6	E (?)
049	"British Steel Corporation"	N/A	N/A	N/A	N/A	2	N/A	N/A	N/A	N/A	3.1	L
056	"League of Nations Representation" "British & Imperial purposes"	N/A	0	3	N/A	3	N/A	N/A	N/A	1	1.2	E
033	"TAYSIDE" "W. Ross the Secretary of State"	N/A	N/A	3	4	N/A	3	2	28	6	5.3	L

TABLE 5.12

Graph measures, undated documents

Doc. No. \ Group	Rank Degree Early	Rank Degree Late	Rank Salience Early	Rank Salience Late	Rank* Degree Early	Rank* Degree Late	Rank Degree Cons.	Rank Salience Cons.
003	0.66	0.64	0.28	0.24	=	=	0.31	0.27
033	0.65	0.51	/	/	=	=	0.11	/
033*	=	=	=	=	0.66	0.62	=	=
019	.51	.37	.39	.32	=	=	.07	.31

*not including the vertex "oil/energy".

TABLE 5.13

Rank Correlation Co-Efficients

Docs. 033, 003, 019

6. Assessment

6.1 Introduction

An assessment of the method which has been described must cover both the methodological issues and the potential significance of the method. The methodological issues can be conveniently grouped into those which concern the content analysis and those which are concerned with the graph model. In this Chapter the first section is about the content analysis, that is the reliability and validity of the procedure which has been used and the overall validity of using Content Analysis in the study of political science.

The second section contains a critique of the graph - theoretic model along a number of dimensions, namely

i) the form of the model
ii) the robustness
iii) the model as a measurement tool
and iv) the practicality of using the model.

The third Section sets out the potential uses and, in some cases, the implications of, a structural model in the analysis of text.

6.2 The Content Analysis

The critique of the Content Analysis can be divided into three main areas

i) the general validity of using Content Analysis as a technique
ii) the reliability of the results obtained
and iii) the validity of the particular results obtained

6.2.1 The general validity

The fundamental validity of the use of Content Analysis in the analysis of communication hinges on two matters, what place public communication is to be assigned in the analysis of the group being studied and how far a systematic, empirical, sometimes mechanical, method distorts the content of communication that was written to be read subjectively. Both these points are argued extensively elsewhere (e.g. Kripp endorf (1980), Osgood (1957)).

6.2.2 Reliability

The reliability of a Content Analysis depends upon the reproducibility and accuracy and the stability of the results obtained. Each of these is discussed in turn below.

6.2.2.1 Reproducibility

Reproducibility is the degree to which different coders are able to produce the same coding, possibly at different times or places. In the coding of the document set 7 different coders were involved, 5 in the initial stages and 2 others in the last stage of coding. In the initial, trial coding, an agreement of over 90% at theme level was achieved between the 5 coders. Once this level of agreement had been reached, the bulk of the coding devolved on one coder. Late in the project two new coders were brought in to code the SNP 4 data. After a short learning period the new coders were able to reproduce the coding of the original coder, again to an accuracy of over 90% at theme level. At the same time the original coder recorded some of his earlier work as a check and these agreed completely.

There was, however, a very considerable difference in the sub-theme coding, each coder producing different supporting sub-themes for the same theme coding. This could not be eradicated by a short period of training (it should be remembered that all the coders, with the exception of the author, were in any case from a political science background and might have been expected to agree on the very narrow part of the conceptual domain involved) and all analysis of the subtheme level was stopped.

In particular the coders were able, at theme level, to agree with the creator of the categories. They were able to categorize events in the way he did, an important factor in considering the objective adequacy of the categorization. (See 6.1. 3.3)

6.2.2.2 Stability

Stability is the degree to which a coding process remains reproducible (usually by the same coder) over time. As mentioned in Section 6.2.2.1, the coder responsible for the bulk of the work was able to reproduce the same coding of a given document after an interval of 18 months.

6.2.3 Validity of the particular results obtained

The validity of the results obtained from the Content Analysis rests on the sampling validity, the semantic validity, the predictive and correlational validities and the validity of the categorization. (Krippendorf 1980, p.158). These are discussed below.

6.2.3.1 Sampling Validity

The sample of Scottish National Party documents was large,

but not systematic (for example not stratified, or random...). Given the circumstances of the project it could not have been. Consequently there must be some doubt as to the accuracy of the results, particularly for the very early SNP 1 data. However, the document set was large (107 documents) and contained both very general documents (such as manifestoes, which do cover much party policy) and a variety of types of document, intended for a variety of audiences. Certainly it was acceptable to the political scientists as sufficient. All the themes were coded and, in that sense, all the possible interactions could have occurred.

The Conservative sample was very small and only included as a rough test of the model - if the model could not have discriminated between two different parties there would have been very little chance of discriminating between documents from the same party at different times. As it happened the picture of the Scottish Conservative party even with this small sample was surprisingly plausible. It would seem that the heart of the structure of such documents appears quite quickly. (Compare, for example, how few bars of a piece of music are required to place the composer - again it is a matter of structure, the composition of the orchestra, the balance of the instruments, the harmonic structure - rather than the tune. Nevertheless, being able to recognize Mozart is not the same as analysing the complete works!)

The sample, then, was adequate for the purpose, but not systematic.

6.2.3.2 Semantic Validity

The semantic validity, that is the assignment of recording

units to the 'correct' category, was achieved by the development of a list of supporting sub-themes. In the early stages of the work, before the final coding was made, there had been some confusion as to the assignment of paragraphs. These were resolved by conference until the sub-theme categories for each theme were agreed and until all the trial data could be coded with sub-theme support. In the final coding only 14 paragraphs (0.1% of all themes coded) were coded as "other".

One further linguistic qualification has to be made. When the relation involved is a mechanical one (co-location in a paragraph), the result could depend heavily on the style of the author and the purpose for which the document was intended.

In practice a large sample of paragraphs will overcome this problem.

6.2.3.3 The validity of the categorization

The validity of the particular categories chosen requires that the categories be both subjectively and objectively adequate.

Taking subjective adequacy to mean that the categories "do not do violence to the [political scientists substantive] feel ... for the structure of what is described (Goodenough 1969), the coding is, almost trivially, subjectively adequate. The political scientists involved in the project were both able to code unambiguously from the final codebook (although not from the first draft of the codebook) and were able to recognize the structures as they appeared after coding.

Objective adequacy, defined as providing "an alien with the knowledge he needs to use [the coding] in a way that [the substantive researcher] will accept as corresponding with the way [he himself] uses it" (Goodenough 1969) is, in this case achieved by the outside coders agreement with the substantive researchers coding. (see also 6.1.2.1)

6.2.3.4 Predictive and Correlational Validity

Beyond these two concepts of validity is the question of the general validity of the categories for the subject matter in hand. This general validity rests on the correlation between the categories (and the resulting structure) and outside events or behaviour. In this sense this study is only a beginning.

6.3. The graph model - assessment

The graph model has proved simple, direct in its pictorial impact, provided a number of measures which have illuminated the structure of the idea-systems under consideration and, as a result, has generated a number of structural hypotheses and means of testing those hypotheses which could not have been arrived at using a different kind of model. The section which follows is a critical assessment of the model, in particular the form of the model, the robustness/sensitivity of the model, the use of the model as a measurement tool and the practical aspects of using the model.

6.3.1 The form of the model

A graph shows a single binary relation. The very simplicity of the graph, which has been so useful for our purposes, particularly in the interpretation of the results, is generally

found to be too restrictive. In much of the work describing systems, of ideas or anything else, graphs are drawn only to be instantly abandoned and the edges or arcs labelled with weights, names and any other information which is thought necessary. Many systems do not even start out as graphs, simply as pictures. There is in fact a tension between the amount of information a diagram should carry and its use as an abstract mathematical entity. In the following sections the possibility of extending the graph model to contain more information whilst retaining at least some of its mathematical properties is considered.

6.3.1.1 Layers

One such possibility arises from extending the relational structure to include two or more (clearly defined and uniform) relations. A simple way to visualize this is to think of a number of graphs stacked one behind the other (Figure 6.1), each graph forming a layer of the structure. Each layer would retain the measures as they have been defined with the same interpretations. Some other metric would be necessary to link the layers together.

6.3.1.2 Levels

Intuitively, the categories of the Content Analysis are of different levels of generality and abstraction. "Economy" "Criticism.." and "Devolution" would, for example, in ordinary usage, include "Agriculture", "Symbolic Injustice" and "Post-Independence Constitution" respectively. The possibility of an independent relation (inclusion) distorting the observation of the chosen relation (co-location) has to be eliminated if

the results are to be unbiased. In the codebook that was used the themes are virtually independent, each theme name/allocation having a list of different supporting subthemes. (The theme 'Economy' for example does not cover "economics" or "the Scottish Economy" as such, terms which would include "agriculture" as sub-themes. Rather it is a category for non-specific references to that economy, so that the question is not "Do the SNP associate agriculture with the Scottish Economy?" which would seem inevitable, but "Do the SNP associate agriculture with general references to the economy?")

Nevertheless, the relation between the themes and sub-themes is not very clear and certainly not uniform. A systematic method of producing a hierarchical structure of levels is required.

It is possible that Componential Semantics would be helpful in this context. Componential semantics is defined by Leech (1981) as "reducing a word's meaning to its ultimate contrastive elements". If the domain were such as to admit of this kind of analysis, it would on the one hand provide a set of independent categories and on the other a clear definition of the relation between the levels. However it would seem that the most successful componential analyses have been on domains in which there is an independent dimension of contrast, colour terms (the spectrum) for example, or kinship terms (time). No such dimension springs to mind for political terms.

Other classifications are possible, particularly where

one relation dominates, constitutent classifications (A is a part of B) or a "genus, species, variety" type of classification. A particularly interesting classification is that given by Agar (1981) which introduces a "stage-process" contrast, i.e. the classification is based on verbs and doing, rather than nouns and naming. It may be that a conceptual leap of this kind is required to achieve a satisfactory multi-level structure in a domain as amorphous as political text.

The problem of level has provoked a deal of discussion in Computer Science, Psychology and in Artificial Intelligence (Hofstadter 1979, Newell 1981). What seems clear is that multi-level models are required. What is not clear is how they might be adequately constructed. A picture of a multi-layer, multi-level structure is shown in Figure 7.1. It is taken from Farbey, Mitchell and Webb (1980) and includes the possibility that not all the relations included in the model are universal (i.e. shared by the whole linguistic community) but that some are idiosyncratic to particular individuals or factions. This point is discussed in greater detail in that paper.

6.3.2 Robustness

The sensitivity of the model to small changes in the data, the addition and deletion of a small number of documents needs to be considered for each type of measure, included viz, those based on distance, those based on degree and the measures of complexity.

6.3.2.1 Distance

The measures based on distance are very sensitive to the

addition of data. A single sentence, by linking two themes not previously linked can cause a major disturbance in the distance values. This does reflect a facet of the substantive domain - a single sentence in a political document can have a dramatic effect on party policy. In general, though, the distance measures are distinctly nervous.

A numerical test of the sensitivity of the distance measures might have been provided by the comparison between SNP 2 and "early" SNP, which contained a substantial number of documents in common. Unfortunately the diameter of the graph in each case was 3 and the radius 2, i.e. nothing was more than 3 steps away from anything else to start with and no nearer than 2 steps, so that the addition of an extra link could only have altered the centrality / distance by 1. In the event 3 out of the 27 themes had different centralities in the two graphs.

6.3.2.2 Degree

Throughout the project degree seemed to be the least sensitive (most robust) measure. In Figure 5.10 for example, degree changes more slowly than does salience and, in the comparison of the SNP documents over time, and with Conservative documents, it was rank degree which changed more slowly than rank salience and which was the better discriminator between the parties.

Further, the correlations on rank degree between SNP 2 and "early", and between SNP 3 and "late", each pair having a large number of documents in common, were 0.99 and 0.96 respectively. (rank salience was also stable at 0.99 and 0.98 respectively).

Rank degree, then, is a robust measure.

6.3.2.3 Complexity

As measured by average degree, the complexity of SNP 2, SNP 3, early and late are shown in Table 6.1. In the first comparison, SNP 2 and "early", the difference in the means is significant, in the second it is not (paired comparison t-test, 99% significance level).

Group	Av. Degree	Significant	Diff.
SNP 2	12.8	Yes	
"early"	11.5		
SNP 3	15.7		No
"late"	15.3		

TABLE 6.1. - Comparison of Average Degree

Erring on the side of caution, complexity must be treated as a sensitive measure. As with centrality, the occurrence of a single new link could alter the other measures of complexity, planarity for example. These measures are therefore also sensitive to small changes.

6.3.2.4 Reduced graphs

All of the comments above apply only to the full graph. For a reduced graph based only on the most significant connection the addition of one new link is less likely to change the reduced structure.

In sum, the model is sensitive generally to small changes in the data, the exception being the measure of rank degree which changes slowly.

6.3.3 The model as a measurement procedure

A major theme of the project has been the assignment of numbers to properties of the graphs studied in order to describe and compare the graphs concisely and precisely. It is important to the further use and interpretation of the model that the question of whether and how far this assignment constitutes measurement is discussed together with the question of the validity of these measures. In this Section the assessment of these measurement implications of the work are discussed.

6.3.3.1 A summary account of modern theory and philosophy of measurement can be found in Finkelstein (1982). Measurement is there defined as "the process of empirical, objective, assignment of numbers to properties of objects or events of the real world in such a way as to describe them". At face value this definition could be taken for a description of the programme of this work. However, a careful dissection of the implications of the definition reveals a number of deep-rooted problems.

"..empirical to objective"

Finkelstein adduces two features to his definition of an empirical process (p.7). The first "..that it must be the result of observation..." is true of the procedure adopted here. The second ".. the concept of the property measured must be based on an empirical relation" is true once the categories have been established (degree, for example, is a matter of counting), but the categories themselves are both idiosyncratic and arbitrary, although they are both adequate

and replicable 6.1.3.3.) There is no theory of themes, or theme creation. For the same reason the process as a whole is not objective, although again, once the categories are fixed, the construction of the graph to the calculation of the measures follows directly (and could be computerized).

"properties and events"

"Measurement presupposes something to be measured..... the concept of 'degree of hotness' as a theoretical construct, interpreting the multitude of phenomenon involving warmth, is necessary before one can construct a thermometer" (Finkelstein 1982, p.10).

The 'something to be measured' in this case is presumably ideology, rather, the properties of an ideology, and there exists very little agreement as to what ideology is, where it is located or what it does. The stipulative definition provided here (ideology is located in the content of the documentary output of a group) has proved useful in delineating a finite, objective body of data and the further restriction to an idea-system (an "ideology", even as defined here, would carry some functional overtones) has further limited the definition of what is to be measured.) The eccentricity of a particular theme, for example, is, in the context of the model, clearly defined. The concept of "centrality" in political ideology is not. In fact, one of the most important uses of the model may prove to be in forcing clear definitions of such concepts as "centrality" and "complexity".

In this study (Chapters 2 and 5) definitions for each of these concepts and interpretation of each measure have been

proposed and, within the limitations of a single date set, hypotheses, for which these measures provide a testing ground, have been worked out. Beyond that, however, the usefulness of the measures will always have to be established by their correlation with theory and with events and behaviour in the relevant substantive domain.

6.3.3.2 Types of scale

Of the measures found most useful in this study, degree eccentricity and salience are ratio scales, their ranks are ordinal scales. The measures of complexity, average degree and planarity, for example, are generally ordinal scales. The concept of complexity is, other than in these measures, so ill-defined, that a statement such as "the SNP ideology is twice as complex as Conservative ideology" is virtually meaningless. As in the previous section, the measurement problem arises from a lack of conceptual clarity.

6.3.4 The practical use of the model

The procedures and programs developed were designed for immediate, one-off use in the context of the research project. For the wider application of the model a considerable benefit would result from developing a computerized system incorporating the experience gained in the present project, which could be used by a subject expert and acted to systematize and assist in the process of categorization and coding. In this section, the requirements of such a system as they have become evident in the course of the work, are presented. An outline system architecture is proposed in Chapter 7, (and in Farbey, Mitchell and Webb 1980).

6.3.4.1 The role of the expert

In political science, as in much other work where the focus of attention is not sharply defined by existing theory, the role of the expert in transforming the available data into a coherent and systematic account, is a crucial one. It is the basis of scholarship and scholarly advance. This is particularly true of the interpretation of documentary evidence: documents are written by humans for other humans to read. No one has ever written a manifesto for computers. It would therefore seem essential, in any system for analysing documents, to retain ordinary (expert?) human understanding within the system.

There is however, an unavoidable tension between the need to retain this human understanding and a programme for measurement and scientific objectivity such as that undertaken in this work. In this work the problem has to some extent been resolved by insisting on subjective and objective adequacy (i.e. agreement on the categories and coding by a number of people including the experts) rather than objectivity in any absolute sense. More could now be done, using the experience gained in this project and incorporating it into an expert system, as the term is used in A.I. using the machine as an "intelligent assistant"(Davis 1978) to ensure reliability (section 6.1.2) but leaving the expert the final judgement.

6.3.4.2 Computerization

Some work (Codebook 1978) has been done on a fine grained coding scheme, in which the categories were "idea-elements" rather than themes. Apart from any theoretical difficulties

which were encountered, this work could not be sensibly carried out without a computerized system for coding. The sheer volume of data precluded any large-scale manual analysis.

Most of the arguments, in fact, for computerized Content Analysis, as put forward, say, in the General Enquirer (Stone et al 1966), apply to this work. The difference between that approach and ours lies in the approach to the development of categories rather than in the need for automation.

In sum, any large scale application of the method requires a substantial development of software in the form of an expert system for the analysis of content.

6.4. Summary

In this Chapter the Content Analysis and the graph model have been assessed in terms of their validity, reliability, formal properties, robustness and as a measurement procedure.

6.4.1 The Content Analysis coding was found to be subjectively and objectively adequate and reliable. The available documents, however, did not constitute a systematic sample and some of the results (particularly SNP 1) must be treated with caution.

6.4.2 The graph model proved to be direct, simple and plausible and particularly useful in this early study in generating hypotheses about the nature of change in systems of ideas.

6.4.2.1 It was found that the model could usefully be extended without losing the benefit of the mathematical model.

6.4.2.2 The sensitivity of the model in relation to small changes in the data was considered. Centrality and complexity were found to be sensitive measures, average degree less so.

6.4.3 The use of the model as a measurement tool has been examined. It was found that the use of the model had forced clarification of a number of concepts, centrality, for example, or complexity, and had provided sensible measures for these more narrowly defined concepts. However the usefulness of the measures will ultimately have to be established by their correlation with events and behaviour in the relevant, substantive domains as well as by the development of theory and clarification of concepts within those domains.

6.4.4 The effectiveness of the model would be greatly enhanced by further programs for the graph analysis and by the computerization of the coding procedure, in such a way so as not to lose the benefit of specialist expertise in the substantive domain.

6.4.4.1 An extension to the model and an outline of the computer system required to realize the model are presented in the next Chapter.

7. Extensions of the model

7.1 Introduction

In this chapter possible extensions to the model are set out. These extensions are partly based on the experience gained on working with the existing model and partly on relevant ideas taken from other disciplines, in particular Cognitive Science and Cognitive Anthropology.

In section 7.2 a generalised form of relational structure model proposed and in section 7.3 a top-level system architecture which could form the basis for a practical system is proposed. In 7.4. basic research which would have to be done to clarify the model are set out. The presentation is a summary of Farbey, Mitchell and Webb (1980).

7.2 Extension to a general relational structure model

7.2.1 The kernel of the proposed model is, a generalization to a more complex relational structure

$$S = \langle A,R \rangle$$

where A is a set of elements (non-empty)
and R is a set of binary relations on A (non-empty)
The restriction to binary relations is a practical one - there is no restriction in principle (see 6.3.1.1 and Figure 6.1).

7.2.2. Beyond this the model is, in principle, multi-level, that is, allowing for different levels of conceptual generality, each with its own relational structure. The levels which have already been examined, theme, sub-theme and idea-element, provide a starting point for the derivation of levels.

7.2.2.1 Other multi-level structures would also provide clues

in the derivation of levels Abelson (1973), Cutler (1980), Simon (1969) as well as the work on componential analysis (Nida, Louw and Smith, 1977, and Leech, 1981). Abelson's work (1979) suggests the possibility of incorporating two kinds of graph, those which represent a consensus or common area over the particular collection of documents under study and those which are individual to selected subsets of the documents. In the context of ideology, for example, the 'generally shared graphs' might be the relations shared by all political parties, whereas the individual relations would be those peculiar to the SNP.

A diagram of the model is shown in Figure 7.1.

7.2.3 Finally, as a representation of the environment in which the idea-system exists, Norman's (1980) model of a cognitive system provides a view on the mechanisms by which an idea-system might respond to changes in the environment. (Figure 7.2)

In this representation individuals (or groups, depending on where the idea-system is considered to be located) each have their idea-systems, part of which is shared and part of which is individual. They communicate with each other and have access to manifest idea-systems, for example, a collection of documents. In the process of interaction all the idea-systems change, for example, two people reading the same text-book might end up with an increase in their shared ideas, and new documents are produced which alter the shape of the manifest idea-system.

An important distinction needs to be made in the mechanism by which the individual systems and the documentary system

change. Norman's model for an individual system includes a control mechanism through which all ideas are filtered. No such mechanism exists for the documentary system which can only grow by accumulation.

What has been pointed up in the preceding discussion in Chapter 5 is that this change by accumulation may be contingent, but not random in a narrow domain.

7.3 A computerized model

An outline system architecture is given in Figure 7.3. Four distinct subsystems would be required

a 'code-book development' program

a 'coding' program

a 'structuring' (model-building) program

a 'descriptive' program

The first program would be, in effect, an expert system for developing the code-book, taking over the work described in section 4.4. The second would be an automation of the work of the coders and the third would produce the graphs from the data. The final program would produce the graph theoretic measures and statistics of the measures.

There would need to be a computerised data-base on which these programs would operate and operating system to control the programs and provide a user-interface.

7.4. Before such a model could be implemented, there is much work to be done on theory of the model. There would also have to be work done in defining and connecting the levels of Figure 7.1. The most promising area in which to look for such theory would seem to be in Cognitive Science, in particular

in expert and knowledge based systems (e.g. Waterman and Hayes-Roth, 1978, Schank 1979, Carbonell 1978, Charniak 1976 among others).

7.5 An extended model has been proposed in this chapter consisting of a multi-layered, multi-level structure. An outline architecture for realising this model has been described.

8. The relation of the work to cognate disciplines

8.1 Introduction

There is no direct precursor of the work presented here, other than McNeill (1979), which appeared after the work on the project had begun, and graphs such as those of Gellner (1979) and Hofstadter (1979 p.30), which were not intended as graphs but which happen to be graphs and therefore amenable to the analysis presented here.

Nevertheless there is work in other fields which has directly influenced and shaped the present study. That work is summarised here, in order to acknowledge this debt and also to locate the results in a framework of existing theory wider than that provided by Scottish Nationalism. This latter is necessary if the work is to be appreciated as a general tool for modelling idea-systems.

Where it is appropriate, the contribution which the use of the model might make to these other disciplines is noted.

8.1.1 The work and ideas which have formed this extra field may be useful grouped in the following way:

i) the documents which constituted the raw material for this study were a collection of artifacts. This immediately places the study as a part of the science of the artificial (Simon 1969) which in turn had a bearing on the kind of analysis that was possible.

In particular these artifacts were linguistic artifacts ('text') and therefore other systematic analyses of text were relevant. These included Content Analysis, the particular form

used has been described in Chapter 4, Computational Linguistics, Artificial Intelligence, Cognitive Mapping, Semantic Memory and aspects of Semantics.

ii) the purpose of the project has been to discover a system or structure of ideas in the documents and therefore Systems and Structuralist ideas have been important

iii) a further objective has been the comparison of different idea-systems. Such comparisons are a part of the purview of Cultural and Cognitive Anthropology and these too, have provided a useful source of concepts.

8.1.2 The discussion which follows has been organised into three sections corresponding to these three major groupings. Section 8.2 deals with the analysis of text, Section 8.3 with System and Structure and Section 8.4 with Anthropology.

8.2. The Analysis of Text

In this Section the relevant ideas taken from other fields concerned with the study of artifacts, in particular those concerned with the systematic analysis of texts are described. These include

i) Simon's work on the Sciences of the Artificial

ii) Computational Linguistics and Artificial Intelligence

iii) Cognitive Mapping

iv) aspects of Semantics

8.2.1 A Science of the Artificial

Simon (1969) defines an artifact as "a meeting point - an 'interface' in today's terms - between an 'inner' environment, the substance and organisation of the artifact itself, and an 'outer' environment, the surroundings in which it operates."

In the case of a document, the interesting inner and outer environment are not the physical ones of a piece of paper and, say, the weather. The interesting inner environment, the one which limits the design of the artifact, is language. The outer environment is the audience for the document (c/f 7.2).

8.2.1.1 The implications of the inner environment as language are straightforward. First, although the study is concerned with content and not with form, the very analysis of that content depends on the use of language by the writer and by the analyst and is limited by that dependence. For example, an analysis of paragraphs such as that used here will depend on the paragraphing ability of the writer and some method for overcoming this problem has had to be found.

Second, other analyses of language which are concerned with semantics or content are related to that presented here.

8.2.1.2 The outer environment, the people who read a document, determines the conditions in which the document succeeds or fails in its goals, which are, roughly, to inform and persuade. If the document is properly designed it will be adapted to this outer environment (Simon p.11), this potential audience. This adaptation which will affect both form and content, must be taken into account when comparing different documents.

There is a distinction to be made here between a single document and a collection or series of documents. A single document may be adapted to a situation, it does not adapt. It is a structure, not a behavioural system. A series of documents may act as a contingent system, that is the structure may change in response to the outer environment. As shown in

Chapter 5, in the case of Scottish Nationalism at least, the collections of documents taken over time do show slow, structural change of this sort.

8.2.1.3 There is one other aspect of Simon's work which is of direct concern and that is his description of complexity. Simon writes

".....complexity frequently takes the form of hierarchy and that....hierarchy....is one of the central structural schemes that the architect of complexity uses" (Simon p.87).

The analysis in Chapter 2 has already demonstrated that complexity is not a simple notion in itself. Moreover, the problems of establishing hierarchical levels in concept networks and idea systems are acute, heterarchy and slippage from one level to another being the order of the day (see also Hofstadter p.651).

That said, Simon's conclusions that the complexity of a structure depends critically on the description of it and his discussion of hierarchy have influenced this work, in particular in the proposals for an improved model (Chapter 7) and in the view taken of structural change.

8.2.2 Artificial Intelligence and Computational Linguistics

In recent years there has been a very considerable development of relational models of idea-systems, fuelled by the need in Artificial Intelligence, Computational Linguistics and Cognitive Psychology generally, to provide computers with a 'deep kind of knowledge' (Winograd 1973) and to represent that knowledge 'in the form of programs in which we can explicitly express the connections between different parts of the system's knowledge thus enriching the possibilities for interaction'

(Winograd p.155). The motivation is thus different from ours and the models developed are different, but the suggestions for improving the model (Chapter 8) and much of the concept of an idea-system have derived from work in these specialities. The models closest to the graph-theoretic model are the semantic network, scripts and conceptual structures, which are summarised below.

8.2.2.1 Semantic networks

'A semantic network purports to represent concepts expressed by natural-language words and phrases as nodes connected to other such concepts by a particular set of arcs called semantic relations' (Simmons 1973). Semantic nets have been used extensively in A.I. as a means of representing knowledge, for example in Waterman et al (1978); Hofstadter (1979); Schank and Colby (1973) and others.

Semantic nets in general differ from the graphs presented in this study in two ways. First, the theme is a high level concept, more general than that usually put forward in these networks. Second, and more importantly, semantic nets are not constrained to only one relation, but will normally have a labelling on the edges. This is perfectly proper, and efficient when the network is used for storing information. It is very difficult, though, to interpret notions such as 'distance' on a net with more than one relation. Additional assumptions are required which require theoretical justification of some kind. Alternatively, the relation has to be a very general one - idea A is somehow connected to idea B - to encompass all the relations of the net.'

As it happens, the relation we have chosen, co-location, is not a semantic one at all. There is no reason, in principle, why the relation should not be a semantic one (inclusion or implication for example) provided it is uniform.

Some semantic nets are graphs. Hofstadter's example (Hofstadter 1979 p.370) is a graph and his 'intuitively sensible notion of nearness' is the notion we wish to measure precisely.

8.2.2.2 Conceptual dependency structures

Another network model which has been used extensively is the Conceptual Dependency Network (Schank 1973). These networks were developed to 'represent the meaning of natural language....in order to provide a representation that will account for the meaning of an..utterance..in an unambiguous way and one that can be transformed back into that utterance or back into any other utterances that have the same meaning' (Schank 1973).

Like semantic networks, conceptual dependency networks are not graphs and therefore do not lend themselves to structural comparison.

8.2.2.3 Belief systems (scripts)

Abelson (1973) has presented a network model or belief systems with six levels of structures: elements, atoms, molecules, plans, themes and scripts. The 'units at each level are structurally bonded according to specified rules in order to produce units at the next highest level'. Again, these are not graphs. On the other hand, many of the concepts of Abelson's work (7.2.2.1) have been useful and certainly

it tackles the higher levels of themes and scripts. Perhaps the difference in approach is rooted in the difference of purpose. Abelson was concerned to provide a representation for belief-systems in general, for the purpose of simulation. This would lead naturally to a classification system starting with a generalised system of conceptual dependency (Schank's) and moving to general linguistic categories at theme level, 'submission', 'betrayal' and 'co-operation' for example.

We have been concerned to provide a structural basis for comparison and measurement, which has enforced a uniform relation and a very narrow range of categories. There is, nevertheless, a degree of overlap and the presentation of an improved model (Chapter 8) has drawn on Abelson's and other multi-level structures in A.I. generally.

note: a 'belief-system' and an 'idea-system' are not necessarily the same thing. A 'belief-system' as expanded by Abelson (1979) contrasts with a 'knowledge system' and characteristically includes beliefs for which there is no consensus, existence beliefs, alternative worlds, evaluative components, episodic material, and beliefs which are unbounded or with variable credence. In general, knowledge systems do not have these features. An idea-system is an amalgam of the two.

8.2.3 Cognitive Mapping

Cutler's (1980) work has formed a part of the general background to the study. He too, moves immediately from a graph to a network which is appropriate to his problem.

8.2.4 Aspects of semantics

Certain aspects of Semantics have been important in

the theoretical background to this work. These are

i) the concept of a lexical system

ii) semantic relations

iii) componential semantics

which are discussed in turn below.

8.2.4.1 A lexical system

The idea of a lexical system, or field, has been widely studied in modern linguistics (for example Berlin and Kay 1969). A lexical system is defined by Lyons (1969) in the following terms

"....we will assume that at least some vocabulary items fall into lexical systems, and that the semantic structure of these systems is to be described in terms of the sense relations holding between the lexical items.....these sense-relations can be paradigmatic....or syntagmatic....the sense of a lexical item is the set of relations which hold between the item in question and other items in the same system."

In other words a lexical item acquires meaning connotatively as well as denotatively.

Analogously, the 'sense' of a theme in an idea-system depends on the relations between the theme and others in the same system i.e. on its connotative aspects as well as its denotative ones.

8.2.4.2 Semantic relations

The only relation explored in this work was co-location. As has already been observed, semantic relations for example, hyponymy, synonymy and antonymy, might usefully provide another source of structure.

8.2.4.3 Componential analysis

The systematic structuring of a text and in particular the development of categories which will serve as vertices in the model, can be achieved in a number of ways. One such method, the method of componential analysis (c/f 7.2.2.1 and 6.3.1.2) although not feasible in this study, might provide a basis for categorisation in similar work.

Componential analysis (Palmer 1981) involves 'the total meaning of a word being seen in terms of a number of distinct elements or components of meaning'. These elements are generally taken in pairs of binary opposites (male-female, adult-child) or along a clearly defined dimension (generation differences or degrees of relationship in a kinship system) and the most successful applications of the technique have been in fields where these kinds of distinction can be made, the analysis of kinship terms, for example, or colour terms. (Leech 1981, Berlin and Kay 1969). Componential Analysis does not handle all semantic relations equally well however, hyponomy being a case in point. (Palmer 1981 p.108). In general, the themes in this study had no clearly defined dimension and were, as was shown by the disagreement of expert coders at sub theme level, too abstract to admit easily agreed components.

In the content analysis described in Chapter 4, the categories were chosen not according to linguistic criteria but to substantive interest. Such categories are not necessarily independent and there is a small amount of overlap in the theme categorisation. To the extent that there is overlap, the corresponding relations which emerge could

be linguistic i.e. inherent in the use of language, rather than characteristic of the idea-system itself. Had the themes been amenable to componential analysis (and had it been practicable) such analysis would have provided a systematic solution to the problem. In the event the themes of interest were not strongly opposed or dimensioned and no componential analysis was attempted.

A possible method of obtaining independent categories might be to perform a principle component analysis on the results of a coding and then to recode on the now orthogonal components. This would retain substantive categories of interest and maintain the independence required by componential analysis.

8.2.4.4 Semantic memory

Semantic networks have been studied in psychology in the context of Semantic memory. Semantic memory contains words and the features which define them. Because of the speed with which words can be stored and retrieved, it is believed that some form of network model is appropriate for this kind of memory. Collins and Quillian proposed a hierarchical model of storage in which words are stored at different levels of the hiearchy with the location of each word signifying how it is categorised (Gregg 1975). Although the assumption of an hierarchical net does not adequately explain the available experimental results, it, too, forms a part of the background to the importance of structure in systems of ideas.

8.3. System and Structure

The terms 'system' and 'structure' are both widely and

variously defined. In this section a working definition for each term is provided together with a brief outline of the relation between this work of other system and structural analysis.

8.3.1 System

The graph model is a system by virtue of being a 'structured set of objects....together with the relationships between them' (Open University Systems group 1981, Klir 1969).

More complex definitions (The City University 1982, Boulding 1956, Checkland 1981, Jordan 1968, Woschni 1982) generally involve notions of purpose, behaviour and coherence, whoch are not appropriate, a priori, to an idea-system which may consist of a single document. These notions are discussed, in turn, below.

8.3.1.1 A single document, or even a group of documents can only acquire purpose by ascription. Purpose resides in the author or reader.

8.3.1.2 Similarly, a document has no behaviour. However a series of documents may show structural change which, although contingent, can be reasonably construed as adaptive behaviour. Such systemic behaviour is an empirical property to be observed (using the method described in this work or otherwise) and related to an external environment for justification and validation. It has been shown in Chapter 5, that the SNP data does show slow structural change and that therefore it possesses structure over and above that imposed on the data by the coding and model. This point is taken up again in section 8.3.2.2. in the discussion of Structuralism.

8.3.1.3 It will be noticed that the model allows for more than one component in the graph i.e. an idea system does not have to be connected at the particular level being investigated, or along the particular relation being investigated: all elements will be connected at the very general level 'are members of the group of documents being studied'.

8.3.1.4 In sum, a graph is a system by virtue of being a relational structure. Other systemic properties, coherence, behaviour and purpose are matters of context and empirical observation.

8.3.2 Structure

'Structure' in this study refers in the first instance to a relational structure or graph, i.e. a system involving a binary relation (Harary 1972, Bridge 1977). As such it carries no structuralist overtones, in the sense that the term is used in Philosophy, Linguistics (and Literary Criticism) or Anthropology. There are, nevertheless points of contact, in particular the distinction between the imposed structure and the underlying structure mentioned in the previous section (8.3.1.2) comes from structuralist thinking (Piaget, 1973), and also the later hypotheses which were examined in 5.9 have a structuralist ring to them. The connection between this work and structuralism works both ways in that graph models provide a quantitative tool for structural comparison.

In the next sections the major themes of structuralism that impinge upon this work are set out.

8.3.2.1 A convenient starting point for the description of

structure as used in structuralist analysis is Piaget (1973). Piaget defines structure as a 'system closed under transformation' and adduces 'wholeness', 'transformation' and 'self-regulation' as necessary properties of structure. He also identifies certain features of 'structuralism' which are relevant here, namely the concept of a (non observable) structure beneath an apparent one, the concept of 'universal' structures and the distinction between systems or structures which have structure by definition, or by fiat, and those which are formed (better 'transformed').

A similar view is taken by Robey (1973) in introducing the notion of Structuralism.

Robey defines Structuralism in two senses, the first 'as that approach to any subject which has as its object..... the laws of solidarity, the reciprocal relations of the different facts under observation, rather than considering these facts in isolation'. (This could be interpreted as a 'systems approach' see Open Systems group (1981) for example).

In the second, narrower, sense, Robey states that '(structuralism)..has grown out of the supposition that the theories and methods of structural linguistics are directly or indirectly applicable to the analysis of all human culture, in so far as all of these, like language, may be interpreted as a series of signs'.

In a later essay in the same book Todorov (1973) remarks 'Structuralism is a scientific method implying an interest in impersonal laws and forms, of which existing objects are only the realisations'.

In all cases wholeness, universality and 'hidden kernels' are seen as key features of a structuralist approach.

8.3.2.2 Taking these points in order: an idea-system (and its graph-theoretic model) clearly is a whole and the properties which are measured are properties of that whole. However, to begin with, the system is a whole by fiat - the idea-system is identified with the documents. It is a whole of the kind Piaget (1973 p.97) refers to as 'global' - it arises simply as the 'union of (its) components'. No deep structure is suggested from which this whole is generated. By contrast the slow structural change which has been observed in this work is an empirical finding and this, together with the results given in Chapter 5 which show the Conservative party documents to have a radically different structure from that of the SNP, suggests strongly that the impact of a particular theme cannot be understood except in relation to the other themes in the system, and that there is an analytic whole, or deep structure, which is reflected in the group of documents.

8.3.2.3 Moreover, the kind of change which has been observed in the SNP data over the periods available seems to be sufficiently slow to match the concept of 'transformation' as put forward by Piaget, as 'an intelligible change', which does not transform things beyond recognition at a stroke, and which always preserves invariance in certain respects. There appears to be not an '<u>invariant</u> (my emphasis) form with different contents' (Levi-Strauss 1973) but a slowly changing one.

8.3.2.4 Obviously, the units of analysis, the themes, are not

self-regulating. Nor for that matter is the system. There is no way of interpreting or incorporating the notion of feedback without going beyond the system under study to include events and people - the environment. What the method can do is to map the results of changes.

8.3.2.5 With regard to the problem of 'surface' and 'deep' structure, in so far as the graph theoretic model gives a snapshot of the whole structure which obtains and can be observed at a particular time, it cannot be used to suggest a deep structure. This is also true of the graph-theoretic features which describe the manifest content. However, some of the features are not of content but of form and these undoubtedly suggest a meta-structure, especially as they change more slowly than do the content features. Moreover, the manner in which these features have emerged as important through the comparative study of the idea-system over time, ties up with the structuralist view that 'the search for an invariant factor can only be carried out... by superimposing .. different works one on the other.... regarding them as though they were a series of transparencies' (Todorov 1973), or again that the business of structuralism is 'to seek behind tangible appearances a more solid organisation of reality at a deeper level' (Levi-Strauss 1973).

8.3.2.6 The concept of 'universality' cannot be established in a single study, depending as it does on the comparison of more than one type of system. It may be that the method proposed here will facilitate such an endeavour, being both systematic and quantitative.

8.3.2.7 In sum, although the approach implied by modelling thematic structure and also many of the problems and concepts, are very much like those to be found in the structuralist approach, it differs from structuralism in its concentration on observable (and, in all probability, consciously devised) thematic structure which is peculiar to an individual idea-system.

8.4. Cultural anthropology

The central problem of this study has been the quantitative description of structure in systems of ideas. The immediate use of this description has been to compare idea-systems: in the practical application these have been systems from a political party over time, or to a lesser extent, two parties. However, the use of the model is not confined to political systems. If the categorisation and content can be made substantively acceptable, the graph and its associated measures will be useful in comparing any two idea-systems. In particular there ought to be a use for them in those areas of anthropology which are directly concerned with comparing idea-systems. A part of such work is a part of structuralist anthropology and has been described above. Two other areas of the subject which are related to this work are described below.

8.4.1 Folk models

Holy and Stuchlik (1981) have developed the concept of a folk model to describe the 'notions, concepts, ideas, etc. which are somehow relevant...to the conduct of...life' and which are 'to a large extent shared'. These notions, because, and in so far as, they are shared, have to be more or less

coherently structured'.

The idea-system of a political party, in so far as it reflects the shared ideas, notions etc. of its members concerning the ordering of political life, is a folk model in this sense.

Moreover, the model developed here would seem to capture the structure implied in Holy and Stuchlik's definition.

8.4.2 Cognitive anthropology

The aim of cognitive anthropology is to 'understand and describe the world of people in other societies in its own terms, as it is conceived and experienced by the people themselves'. (Casson 1978).

In pursuit of this aim, cognitive anthropologists have developed methods which cross-cut those presented here. In particular its attempt to empirical analysis of classification and language, the concept of an 'ideational system' qua system (Keesing 1981), the notions of subjective and objective adequacy, and the placing of this system in a wider political and social setting have all been a part of the background assumptions to this work.

An important concept which has been adapted directly from an anthropological study was the concept of the level of presentation, which led directly to the hypothesis presented in 5.9.4. Much the same phenomenon has been described by Hofstadter (1979 p.323) in his attack on the isomorphism between two semantic networks.

8.5 In this chapter the theoretical background to the study has been set out. This background has had to be developed

across a number of existing disciplines corresponding to the salient features of the data and the purposes of the analysis. These were first, that the data consisted of a collection of artifacts, second that the purpose was partly to discover structure or system in the content of the documents, third, that a further objective was the comparison of the resulting idea-systems and finally that the documents were to be compared over time.

The major contributing disciplines have been Anthropology, Linguistics and Cognitive Science. The relevant concepts from each discipline have been summarised.

The contribution from Anthropology centered on Structuralism and Cultural Anthropology. It was found that approach implied by a thematic structural model was like that of Structuralism but differed from it in the concentration on manifest observable structure. In this view systemic properties, such as persistence are matters for empirical observation rather than theory.

Several notions were incorporated directly into the work from cognitive and cultural anthropology, namely the concept of a folk model and the concept of level of presentation. The concepts of subjective and objective adequacy, although more generally used, were in practice taken from these writings.

The study of linguistics provided a source of possible relations for the graph and pointed to a possible resolution of problems of level and generality in certain cases through the use of Componential Analysis.

Cognitive Science, in particular the work of Simon and

the work on semantic networks provided some theory, in particular the setting of the problem as one in the sciences of the artificial and a focus on level as found in semantic networks. A particularly important notion for further research was the possible use of an expert system in coding and in developing the code-book.

8.5.1 From the survey some general concepts in the study of idea-systems have emerged. These include concepts about change in idea-systems, the type and mechanism of change, the control of change (accumulation) and the speed of change. It was found that the nature of the change observed would be consistent with an underlying, slowly changing but contingent system.

The significance of the concept of level was remarked, both the level of presentation and its effect on the apparent level of complexity and the problems associated with differing levels of generality in the themes. Contained within these latter are the problems of heterarchy or hierarchy which hamper the development of an orthogonal system of themes.

The similarity between the 'folk-model' and the definition of an ideology as being part of the shared idea-systems of its members implies that the model could be applied to the study of such systems. More generally it seems that the model will provide an empirical tool for the empirical study of manifest idea-systems, such as those which form the subject matter of the contributing disciplines.

9. Summary and Conclusions

9.1. Summary

In this thesis a quantitative model of thematic structure in the content of written documents has been developed and tested on a sample of 126 documents, mostly of the Scottish National Party.

The model is a graph, or simple relational structure; the invariants of the graph have been used to provide measures to describe and compare different groups of documents within the data set.

The graphs have been derived from the documents using a form of Content Analysis (Chapter 4) in which each paragraph in each document was coded onto a frame of policy themes. Each theme corresponded to a vertex of the graph and the relation 'co-mention in a paragraph' corresponded to an edge.

Three groups of documents were compared - two groups of SNP documents and each of these with a group of Conservative documents. The results of this analysis (Chapter 5) showed the SNP ideology to be a slowly changing system of ideas and one which contrasted markedly with that of the Conservative documents. From this analysis it was possible to suggest further hypotheses about the nature of change in the SNP ideology and in idea-systems generally.

A full assessment of the results and the underlying methodology has been presented (Chapter 6) and suggestions for an extended, computerised model based on this assessment has been described (Chapter 7).

Finally the model has been placed in a wider context (Chapter 8) and links developed between this work and other work on manifest idea-systems in Anthropology, Artificial Intelligence, Linguistics and Psychology.

9.2. Conclusions

9.2.1 Methodological

9.2.1.1 The model in practice has proved simple, direct and plausible.

9.2.1.2 It has been particularly useful for generating a new, structural type of hypothesis.

9.2.1.3 The use of the model has forced clarification of several apparently straightforward concepts, such as centrality, complexity and distance.

9.2.1.4 The use of the model as a measurement tool was hampered by the lack of substantive theory.

9.2.1.5 The effectiveness of the model would be greatly enhanced by further computerisation.

9.2.1.6 Certain of the graph measures were very sensitive, for example centrality and complexity; others, such as rank degree were more robust.

9.2.2 Idea systems

9.2.2.1 The concept of an idea-system qua system is not unreasonable. The idea-system under study showed slow, structural change, consistent with the existence of a persistent underlying system.

9.2.2.2 Such systems may be contingent but are not random.

9.2.2.3 Some systemic properties, coherence, (apparent) behaviour and purpose can be treated as matters of empirical

observation using the model.

9.2.2.4 Certain problems emerged as common to several disciplines concerned with the study of idea-systems. These include problems of level, and the type, mechanism and control of change in such systems.

9.2.2.5 The model could be useful in any of these disciplines - in particular it can provide a structural snapshot of a specific conceptual domain.

9.2.3 Substantive (Scottish National Party)

9.2.3.1 Cultural factors declined with respect to material factors, but not uniformly.

9.2.3.2 The complexity of the ideology increased through the first three periods but decreased at the last. This was due to the level of presentation, rather than the whole ideology, becoming simpler.

9.2.3.3 All the early key themes, except Devolution showed signs of dilution over the period studied.

9.2.3.4 The structure of the SNP ideology altered as the party became a mass public party. The later structure was more like that of the Conservative party than the earlier structure had been.

9.3 This report presents the results of an empirical investigation into the use of graphs as models of structure in systems of ideas, using documents of the Scottish National Party as data. The results of the work have shown the model to provide useful, quantitative measures of system structure in a very soft system. The measures are useful in that they force clarification of structural concepts, allow generation of structural hypotheses and enable soft systems to be characterized and compared. The work could be extended to other manifest idea systems in so far as, and only so far as, these can be represented in documents.

"What's <u>done</u> we partly may compute
But know not what's resisted".

(Robert Burns - Address to the Unco Guid)

Appendix 1 - Document 119

Scottish National Party

14A Manor Place
Edinburgh EH 3 7ES
Tel 031-226 3661

para. no.

1

Your ref
Our ref

George Reid, MP for Clackmannan and East Stirlingshire and SNP Spokesman on Welfare, will take a press conference at 11.00 hrs. in the Caledonian Hotel, Edinburgh. 2

FOR USE AFTER 11.00 HOURS ON MONDAY SEPTEMBER 30. 3

A Scottish National Party MP today charged that the Labour Party had kept a "guilty silence" on the issue of poverty in Scotland 4

George Reid, MP for Clackmannan and East Stirlingshire, told a press conference in Edinburgh that until the SNP revealed the facts, the Labour Party had succeeded in hiding from the people of Scotland the extent of Labour's failure as a party of social reform. 5

He said: "Months ago - with the unanimous backing of the Annual Conference - SNP launched the "War on Poverty" programme with the charge that one million Scots were living on, or close to, the poverty line. Labour spokesmen have made no attempt to deny the charge. Their silence is the silence of guilty men, for they know that the Labour Party, politically dominant for decades in Scotland, must accept responsibility for the fact that Scotland has the worst social and economic record among developed countries in Europe. 6

The Labour Party has tried to keep silence on the issue throughout this election campaign as well. The flood of 7

election promises omits any mention of allocating even a part of the oil revenues for use in Scotland to eradicate poverty. Labour's Development Agency has been given no powers to act on poverty or any other social issue. Incredibly, the Labour Party has refused to respond to the urgent pleas of the Child Poverty Action Group to take the one single action that would do most to help low income families, that is to extend an increased family allowance to the first child. Their pensions' plan - for adequate pensions only in twenty years time - is a cruel Public Relations' hoax at the expense of the country's pensioners.

SNP's policy includes a £2 a week allowance for the first child and old age pensions for a single person equivalent to 3/5ths of a national minimum wage of £25 (at May 1974 prices), as well as a pension equivalent to the national minimum wage as the basic provision for widows and the disabled. 8

Like all SNP's proposals, these policies have been costed. We know that a self-governing Scotland could afford them. I challenge the Labour Party to deny it. 9

The discovery of Scotland's oil gives us an opportunity to eradicate the poverty and neglect which is the legacy of those decades of Labour dominance. But the opportunity will only be realised if the people of Scotland act urgently to ensure Scottish control. Under London control, the economic benefits of the oil, benefits SNP would use to create Scandinavian levels of welfare, will be mortgaged away to the international money-lenders. 10

References

ABELSON, R.P. (1973). The Structure of Belief Systems. In: Computer Models of Thought and Language ed. R. Schank and K. Colby, (W.H. Freeman: San Francisco).

ABELSON, RP. (1979). Differences between Belief and Knowledge Systems. Cognitive Science 3 (1979) 355 - 366.

AYER, .J. (1982). Philosophy in the Twentieth Century. (Weidenfeld nd Nicholson: London).

BERLIN and KAY. (1969). Basic Colour Terms: Their Universality and Evolution . from Casson (ed.) 1981 p132

BOULDING,K. (1956). quoted in Checkland (1981).

BRIDGE, J. (1977). Beginning Model Theory - The Completeness Theorem and Some Consequences. Oxford Logic Guides, (Oxford: The Clarendon Press).

BRAND, J. (1978). The National Movement in Scotland. (London: Routledge and Kegan Paul).

BOWERS, D.A., MITCHELL, C.R. and WEBB, K. (1979). A Dynamic Model of Bicommunal Political Systems - I. Model Description. The City University DSS/DAB/CRM/KW/160.

CARBONELL, J.G. (1978). Automated Ideological Reasoning. Cognitive Science 2 (1978) 27 - 51.

CARNEY, T.F. (1972). Content Analysis. (London: Batsford).

CASSON, R.W. (ed.) (1981). Language, Culture and Cognition. (London: Macmillan).

CHARNIAK, E. and WILKS, Y. (eds.) (1976). Computational Semantics. (Amsterdam: North Holland).

CHECKLAND, P. (1981). Systems Thinking, Systems Practice. (Chichester: John Wiley and Sons).

CITY UNIVERSITY (1982). Prospectus, Department of Systems Science.

CODEBOOK Dimensions of Ideology and Evaluation, Coding Manual, Part I - Document and Paragraph Level Data. The City University (unpublished - 1978).

CONVERSE, P.E. (1964). The Nature of Belief Systems in Mass Publics. In: Ideology and Discontent ed. D. Apter (London: Collier Macmillan).

CUTLER, R.M. 91980). On the trail of the elusive policy. Centre for Empirical Research in International Relations, Geneva.

DAVIS, R. (1978). Knowledge Acquisition in Rule-Based Systems. In: Pattern-Directed Inference Systems eds. Waterman and Hayes-Roth (New York: Academic Press).

DEUTSCH, K. (1966). Nationalism and Social Communication. Cambridge Mass.: The M.I.T. Press.

FARBEY, B. (1977). Quantification in International Relations - Some Research Notes. The City University DSS/BAF/143.

FARBEY, B.A., MITCHELL, C.R. and WEBB, K. (1979). A Meaning and Measurement of Political Ideology: The case of Scottish Nationalism. The City University DSS/BAF/CRM/KW/190.

FARBEY, B.A., MITCHELL, C.R. and WEBB, K. (1980). Modelling Political Belief Systems (unpublished paper).

FINKELSTEIN, L. (1975). Fundamental Concepts of Measurement, Definitions and Scales. Measurement and Control 8.

FINKELSTEIN, L. (1982). Theory and Philosophy of Measurement. In: Handbook of Measurement Science vol. 1 (Chichester: John Wiley and Sons) ed P. Sydenham.

GELLNER, E. (1979). Words and Things (London: Routledge and Kegan Paul).

GELLNER, E. (1983). Nations and Nationalism (Oxford: Basil Blackwell).

GERBNER, G., HOLSTI, R. KRIPPENDORF, K., PAISLEY, W.J. and STONE, P.J. (1969). (eds). The Analysis of Communication Content: Developments in Scientific Theories and Computer Techniques (New York: John Wiley).

GOODENOUGH, W.H. (1969). Yankee Kinship Terminology - A problem in Componential Analysis. In: Cognitive Anthropology ed. Tyler (New York: Holt, Rhinehart and Winston).

GREGG, V. (1975). Human Memory (London: Methuen).

HANHAM, H. (1969). Scottish Nationalism (London: Faber).

HARARY, F. (1972). Graph Theory. (London: Addison-Wesley).

HARVIE, C. (1977). Scotland and Nationalism - Scottish Society and Politics 1707-1977. (London: George Allen and Unwin).

HECHTER, M. (1975). Internal Colonialism (London: Routledge and Kegan Paul).

HOFSTADTER, D.R. (1979). Godel, Escher, Bach - An External, Golden Braid. (Hassocks, Sussex; Harvester).

HOFSTADTER, D. and DENNETT, D. (1981). The Mind's I. (Brighton: The Harvester Press).

HOLY, L. and STUCHLIK, M. (eds). (1981). The Structure of Folk Models (London: Academic Press).

JORDAN, N. quoted in Checkland (1981).

KAMENKA (ed.) (1976). Nationalism - The Nature and Evolution of an Idea. (London: E. Arnold).

KEDOURIE, E. (1960). Nationalism (Essex: The Anchor Press).

KEESING, G. (1981). Theories of Culture. In: Language, Culture and Cognition ed. Casson (London: Macmillan).

KLIR, G.J. (1969). An Approach to General Systems Theory (New York: van Nostrand, Rheinhold).

KRIPPENDORF, K. (1980. Content Analysis. (London: Sage).

LEECH, G. (1981). Semantics - The Study of Meaning (Harmondsworth, Middlesex: Penguin).

LEVI-STRAUSS, C. (1973). Structural Anthropology (Harmondsworth, Midlesex: Penguin).

LYONS, J. (1969). Introduction to Theoretical Linguistics (Cambridge: The University Press).

LYONS, J. (1970). New Horizons in Linguistics (Harmondsworth, Middlesex: Penguin).

MACKIE, J.D. (1972). A History of Scotland (Harmondsworth, Middlesex, Pelican).

MACNEILL, D. (1979). The Conceptual Basis of Language. (Hillsdale, New Jersey: Lawrence Erlbaum Associates).

NAIRN, T. (1981). The Break-up of Britain (London: Verso).

NEWELL, A. (1981). The Knowledge Level. Artificial Intelligence Summer

NIDA, E.A. LOUW, J.P. and SMITH, P.B. (1977). Semantic Domains and the Componential Analysis of Meaning in Cole (ed.) Current Issues in Linguistic Theory (Indiana University Press).

NORMAN, D.A. (1980). Twelve Issues for Cognitive Science. Cognitive Science 4 no. 1, 1-31.

NORTH, R.C., HOLSTI, O.R., ZANINOVICH, M.G. and ZINNES, D.A. (1963). Content Analysis (Evanston, Illinois: Northwestern University).

OPEN SYSTEMS GROUP (1981). Systems Behaviour (London: Harper and Row).

OPEN UNIVERSITY (1981). Graphs, Networks and Design TM361 Unit 10.

OSGOOD, C.E., SUCI, G.J. and TANNENBAUM, P.W. (1957). The Measurement of Meaning (Urbana: University of Illinois Press).

PALMER, F.R. (1981). Semantics. (Cambridge: Cambridge University Press).

PARSLER, R. (ed.) (1980). Capitalism, Class and Politics in Scotland (Farnborough: Gower).

PIAGET, J. (1973). Structuralism (London: Routledge and Kegan Paul).

POOL I de SOLA (1959). (ed.). Trends in Content Analysis (Urbana: University of Illinois Press).

ROBEY, D. (ed.) (1973). Structuralism: An Introduction - Wolfson College Lectures 1972 (Oxford: Clarendon Press).

SCHANK, R.C. (1973). Identification of Conceptualisations Underlying Natural Language in Computer Models of Thought and Language eds. Schank R.C. and Colby K. (San Francisco: (W.H. Freeman).

SIMMONS, R.F. (1973). Semantic Networks - Their Computation and Use for Understanding English Sentences. In: Computer Models of Thought and Language eds Schank .R.C and Colby K. (San Francisco: W.H. Freeman).

SIMON, H.A. (1969). The Sciences of the Artificial. (Cambridge, Mass.: The M.I.T. Press).

SMITH, A.D. (1971). Theories of Nationalism. (London: Duckworth).

SMITH, A.D. (1981). The Ethnic Revival in The Modern World (Cambridge: Cambridge University Press).

STONE, P.J. DUNPHY, D.C. and SMITH, M.S. (1966). The General Enquirer - A Computer Approach to Content Analysis (Cambridge, Mass.: The M.I.T. Press).

TYLER, S. (ed.) (1969). Cognitive Anthropology (New York: Holt Rhinehart and Winston).

URWIN, D.W. (1982). Territorial Structures and Political Developments in the United Kingdom in The Politics of Territorial Identity eds Rokkan S and Urwin D. (London: Sage).

WATERMAN, D. and HAYES, ROTH (eds.) (1978). Pattern Directed Inference Systems (New York: Academic Press).

WEBB, K. (1978). The Growth of Nationalism in Scotland. (Harmondsworth, Middlesex: Pelican).

WINOGRAD, T. (1973). Procedural Models of Language Understanding in Computer models of Thought and Understanding eds. Schank R.C. and Colby K. (San Franciusco: W.H.Freeman).

WOSCHNI (1982). Signals and Systems in the Time and Frequency Domains in Handbook of Measurement Science vol 1, ed. P. Sydenham (Chichester: John Wiley and Sons).

Documents of the Scottish National

Party mentioned in the Text

001 S.N.P. Constitution

002 What about Scottish Housing

003 Policy of the S.N.P.

009 Scotland's Answer to the Crisis

018 Your Scottish Candidate - Wilson

019 Policy of the S.N.P.

033 The Future of Tayside

053 A Scottish Parliament - Now

056 Heckling Points

049 Save Scottish Steel

067 Policy of the S.N.P.

119 Press Release (Appendix 1).

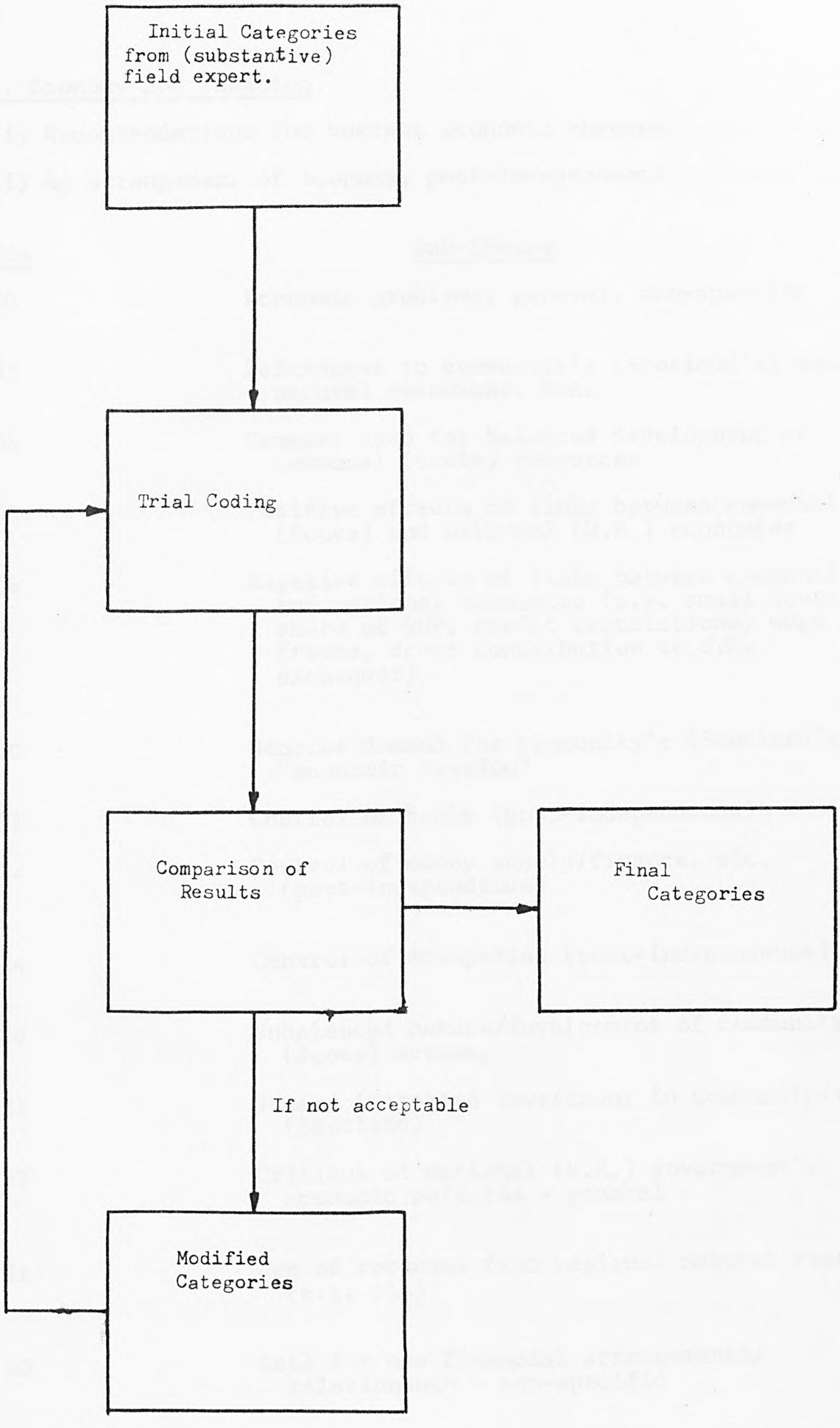


Figure 4.1 - Codebook Development Procedure

06. Economy and Taxation

(i) Recommendations for current economic changes

(ii) Re-arrangement of economy, post-independence

Code	Sub-themes
00	Economic problems; general, non-specific
03	References to community's (Scotland's) wealth, natural resources, etc.
04	General need for balanced development of communal (Scots) resources
05	Positive effects of links between communal (Scots) and national (U.K.) economies
06	Negative effects of links between communal and national economies (e.g. small Scots share of GNP, credit restrictions, wage freeze, Scots contribution to U.K. exchequer)
10	General demand for community's (Scotland's) "economic freedom"
11	Control of Banks (post-independence)
12	Control of money supply/finance, etc. (post-independence)
14	Control of monopolies (post-independence)
20	Unbalanced nature/development of community (Scots) economy
21	Demand increased investment in community/region (Scotland)
22	Critique of national (U.K.) government's economic policies - general
31	Use of revenues from regional natural resource (e.g. oil)
40	Call for new financial arrangements/ relationship - non-specific

Code	Sub-themes
41	Financial arrangements with England (post-independence)
42	Financial arrangements with England (post-devolution)
51	Control of multi-nationals' influence (post-independence)
52	Reduction level of foreign investment in community/region (Scotland)
53	Critique of regional economic organisation (E.E.C), e.g. subsidies
57	Inflation - result of membership of regional economic organisation (E.E.C)
58	Inflation - devolution, self-government or independence as a means of halting
59	Inflation - effects of (e.g. on lower paid, agriculture, tax levels)
60	Inflation - general problem of
61	Inflation - national government efforts to reduce effects of
62	Inflation - connection with 1970's energy crisis, oil price rise, etc.
63	Inflation - connection with increased prices of (imported) food/raw materials
64	Inflation - connection with increased wage levels, etc. (wage demands, rises, wages outstripping productivity, government expenditure/profligacy)
65	Efforts to develop communal (Scots) and national (U.K.) economy - successful
66	Efforts to develop communal (Scots) and national (U.K.) economy - unsuccessful
67	Call for special tax benefits for community (Scotland), e.g. special depreciation allowance, abolition of SET, differential tax rates (e.g. VAT), raising tax thresholds
68	(General) commitment to cut taxes/tax levels
69	Specific actions to cut taxes by national government (e.g. tax cuts for lower paid)

Code	Sub-themes
70	Prices and incomes policy - on national basis (for whole of U.K.)
71	Prices and incomes policy - effects on community (Scotland)
72	Prices and incomes policy - criticism of efforts to undermine by key groups
73	Prices and incomes policy - need for co-operation from trade unions (T.U.C.) and industry (C.B.I.)
74	Prices and incomes - need for statutory restraints
76	Prices and incomes - rejection of statutory restraints: critique of effects
80	Taxation - general critique of tax levels as too high/stultifying, etc.
81	Taxation - effects of cuts as stimulus to economy/ export drive, etc.
82	Taxation - need for an increase to pay for reforms/ social services
83	Commitment/action to cut public expenditure
84	Wealth tax - for introduction
85	Wealth tax - against
86	Criticism of current tax system as unfair, inefficient
90	Other (list)

FIGURE 4.2 - A theme and its sub-themes

He said: Months ago with the unanimous backing of the Annual Conference - SNP launched the "War on Poverty" programme with the charge that one million Scots were living on, or close to, the poverty line. Labour spokesmen have made no attempt to deny the charge. Their silence is the silence of guilty men, for they know that the Labour Party, politically dominant for decades in Scotland, must accept responsibility for the fact that Scotland has the worst social and economic record among developed countries in Europe.

DOC 119

Coded as:	02	Criticism of National (London) Govt.
subthemes	31	Specific critique in particular themes e.g. Education, Housing
	51	Critique of Left-Wing (Labour Party)
	16	Income, Wages and Salaries
subtheme	41	Critique of National Govt. especially with regard to levels of community (Scots) wages compared to elsewhere
	24	Referencing/Comparing to other "nations'" conditions
subtheme	61	Commonwealth/Europe

FIGURE 4.3 - An example of Coding

Code	Theme
01	Agriculture and Primary Products generally; land
02	Critique of national (London) Government
03	Cultural protection
04	Defence
05	Devolution
06	Economy and Taxation
07	Education
08	Employment and Unemployment
09	Reactions of other Community (e.g. English)
10	Fishing
11	Foreign Relations
12	Foreign Trade
13	Health Care
14	Historical Justifications
15	Housing
16	Income, Wages and Salaries
17	Industry and Industrial Relations
18	Legal/Judicial
19	Local and/or Regional Government
20	Natural Resources; Energy, Oil
21	Organisational Structure of Communally-based Political Movement (e.g. SNP)
22	Post-Independence Constitution
23	Quality of Life; Environment
24	Referencing/Comparison to others
25	Social Security

Code	Theme
26	Symbolic Injustice
27	Transport and Communications
90	Other (list code sheet)
91	No Content
92	Uncodable

FIGURE 4.4 - The Theme List

The London Government	have only now admitted
	proclaims
	now impose great material sacrifices
	distorts the Scottish Economy's balance
	invariably acts against Scottish interests
	have not avoided the situation
	obstructs Scottish Reconstruction
	dictatorship is economic
	pillages Scottish assets
	forces Scottish exports
	vainly attempts to bolster up a bankrupt England
	attempts in vain

codes to

The London Government	admits/proclaims
	imposes/distorts
	acts against/avoids/obstructs/ pillages/forces/attempts in vain

Figure 4.5 - Idea-element Coding

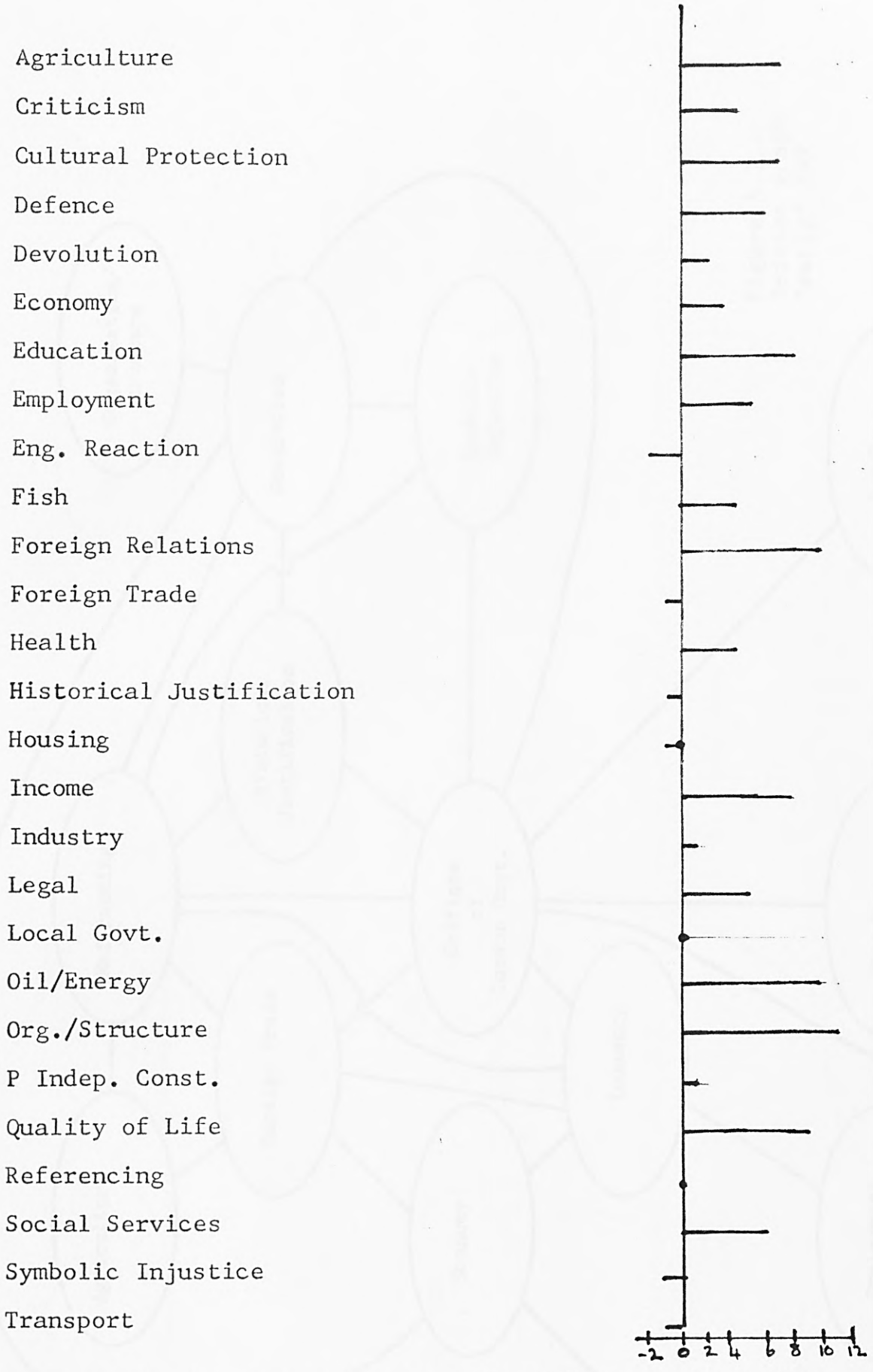


Figure 5.1 The change in degree "early" SNP v. "late" SNP

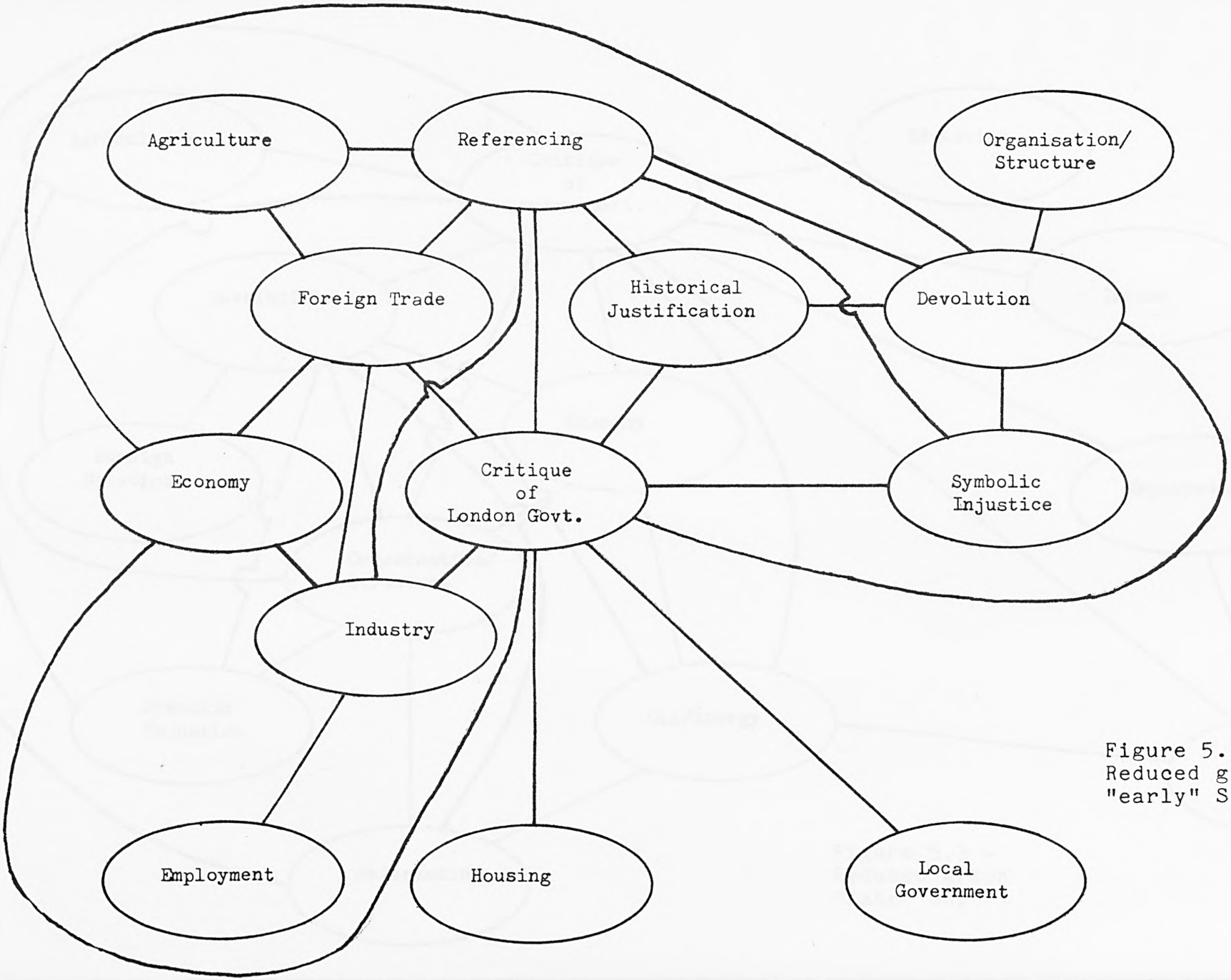


Figure 5.2 -
Reduced graph
"early" SNP

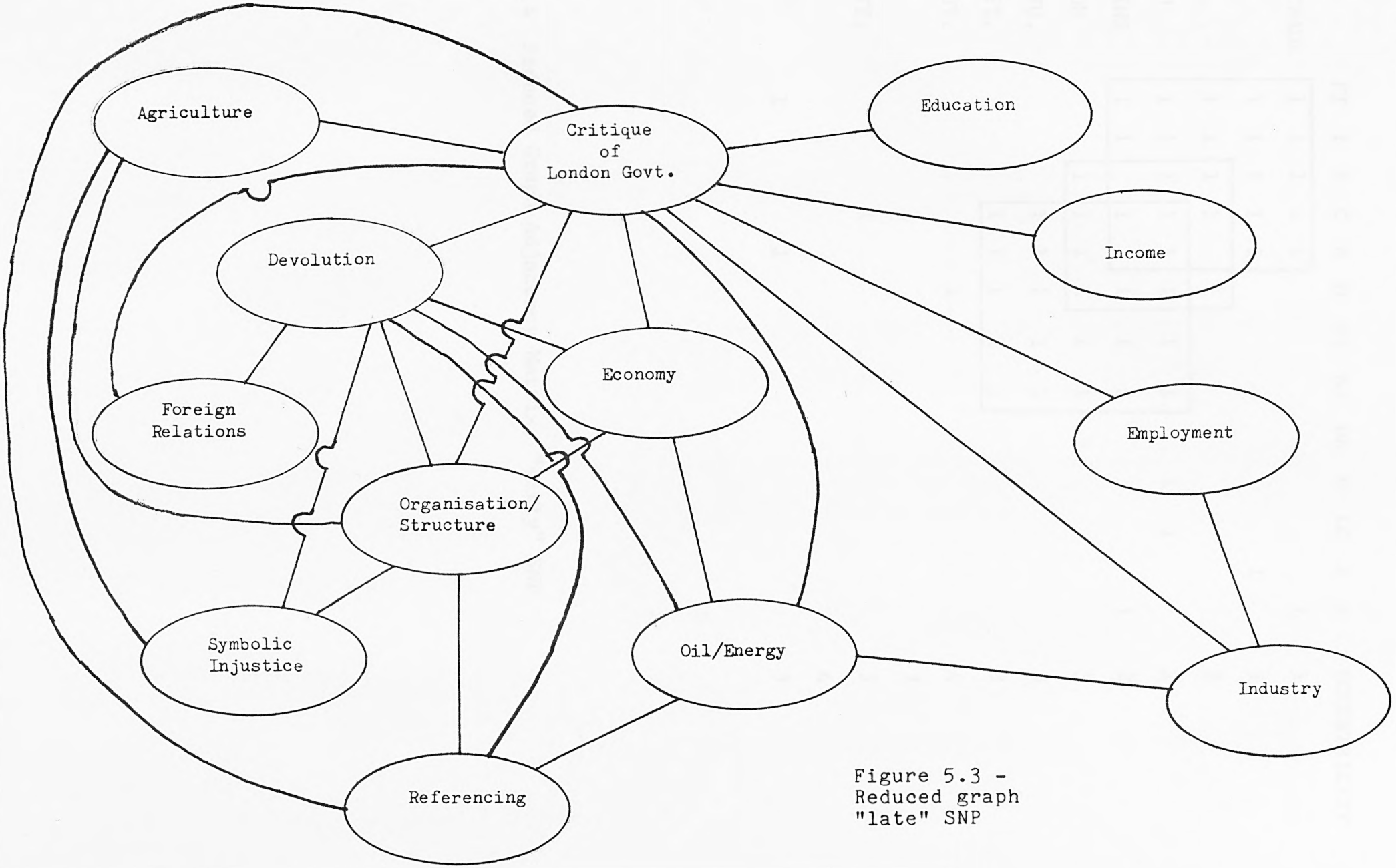


Figure 5.3 - Reduced graph "late" SNP

THEME	FT	I	E	C	R	D	SI	HJ	OS	H	LG	E	A	ECCENTRICITY
FOREIGN TRADE	1	1	1	1	1								1	3
INDUSTRY	1	1	1	1	1							1		3
ECONOMY	1	1	1	1		1								2
CRITICISM	1	1	1	1	1	1	1	1		1	1			2
REFERENCING	1	1		1	1	1	1	1					1	2
DEVOLUTION			1	1	1	1	1	1	1					3
SYMB. INJU.				1	1	1	1	1						3
HIST. JUST.				1	1	1	1	1						3
ORG. STRUT.						1								4
HOUSING				1										3
LOCAL GOVT.				1										3
EMPLOY.		1												4
AGRIC.	1				1									3

Figure 5.4 Reduced Graph Adjacency Matrix, "early" SNP

THEME	SI	FR	A	R	D	OS	CL	E	OE	I	E	I	E	ECCENTRICITY
SYMB. INJU.	1				1	1								3
FOREIGN REL^{S}.		1			1	1	1							2
AGRICULTURE			1	1	1	1	1							2
REFERENCING			1	1	1	1	1		1					2
DEVOLUTION	1	1	1	1	1	1	1	1	1					2
ORG. STRUC.	1	1	1	1	1	1	1	1	1					2
CRITIQUE OF LONDON		1	1	1	1	1	1	1	1	1	1	1	1	2
ECONOMY					1	1	1	1	1					2
OIL/ENERGY				1	1	1	1	1	1	1				2
INDUSTRY							1		1	1	1			3
EMPLOYMENT							1			1				3
INCOME							1					1		3
EDUCATION							1						1	3

Figure 5.5 Adjacency Matrix: Reduced Graph, "late" SNP

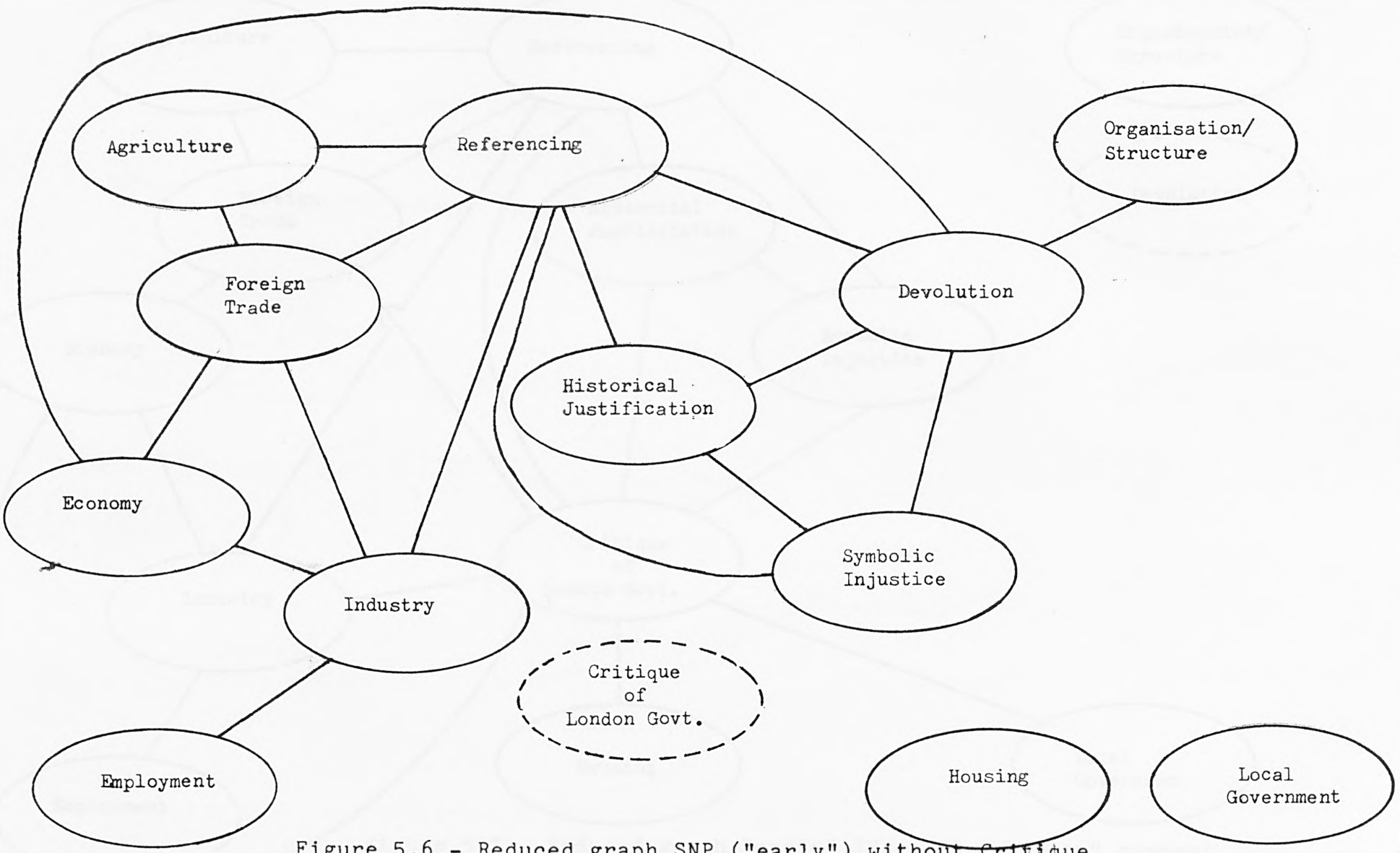


Figure 5.6 - Reduced graph SNP ("early") without Critique

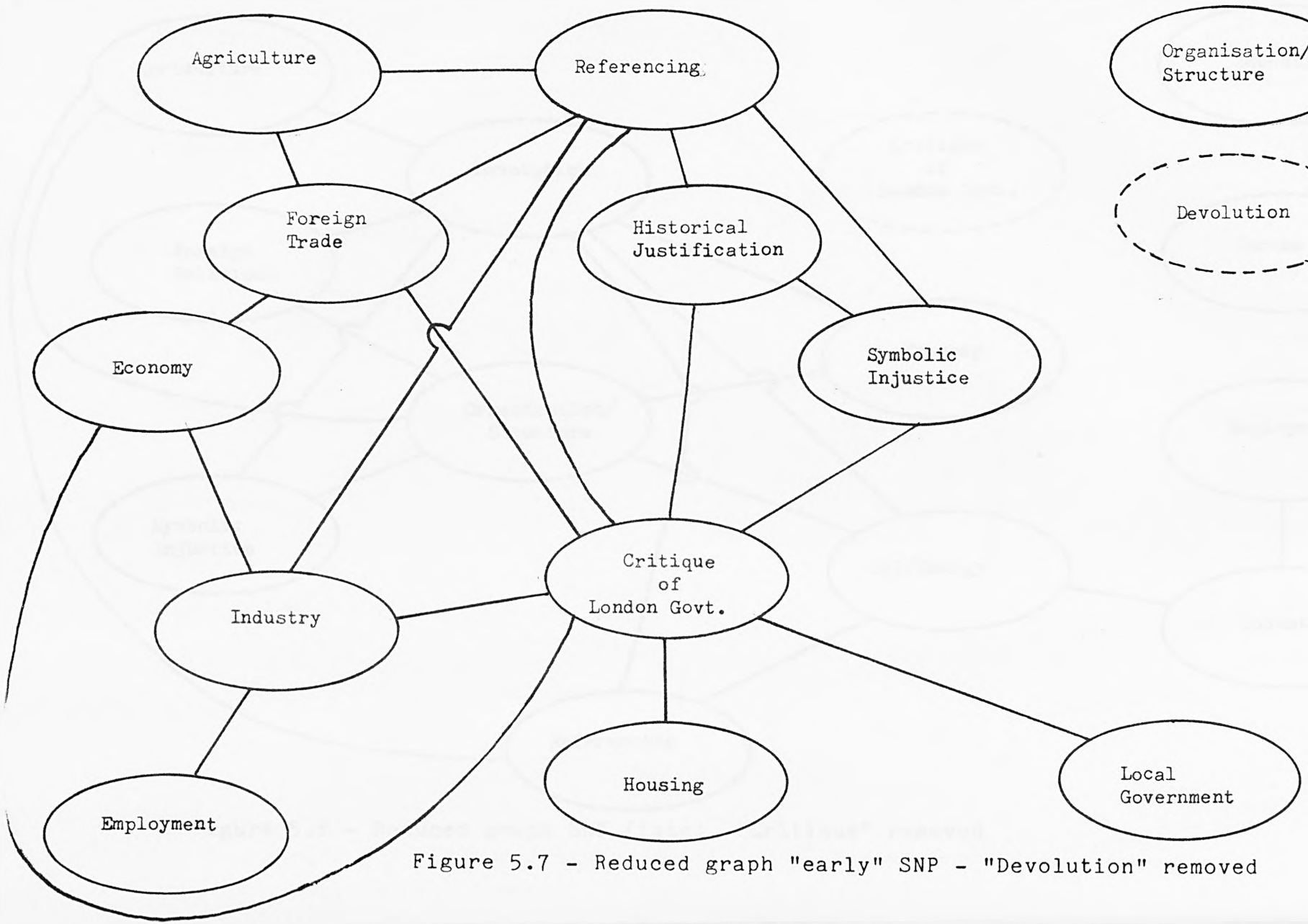


Figure 5.7 - Reduced graph "early" SNP - "Devolution" removed

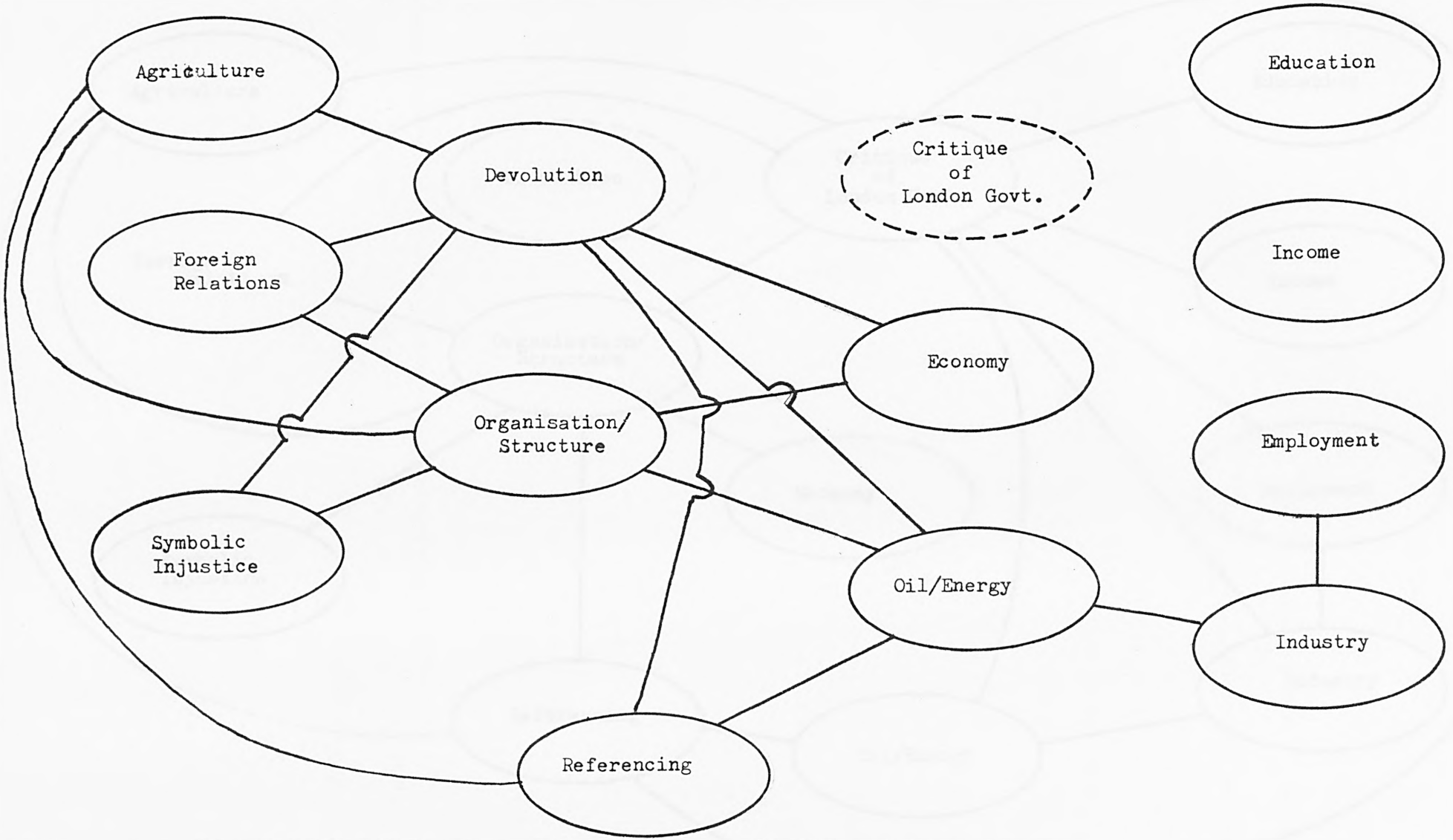


Figure 5.8 - Reduced graph SNP (late) -"Critique" removed

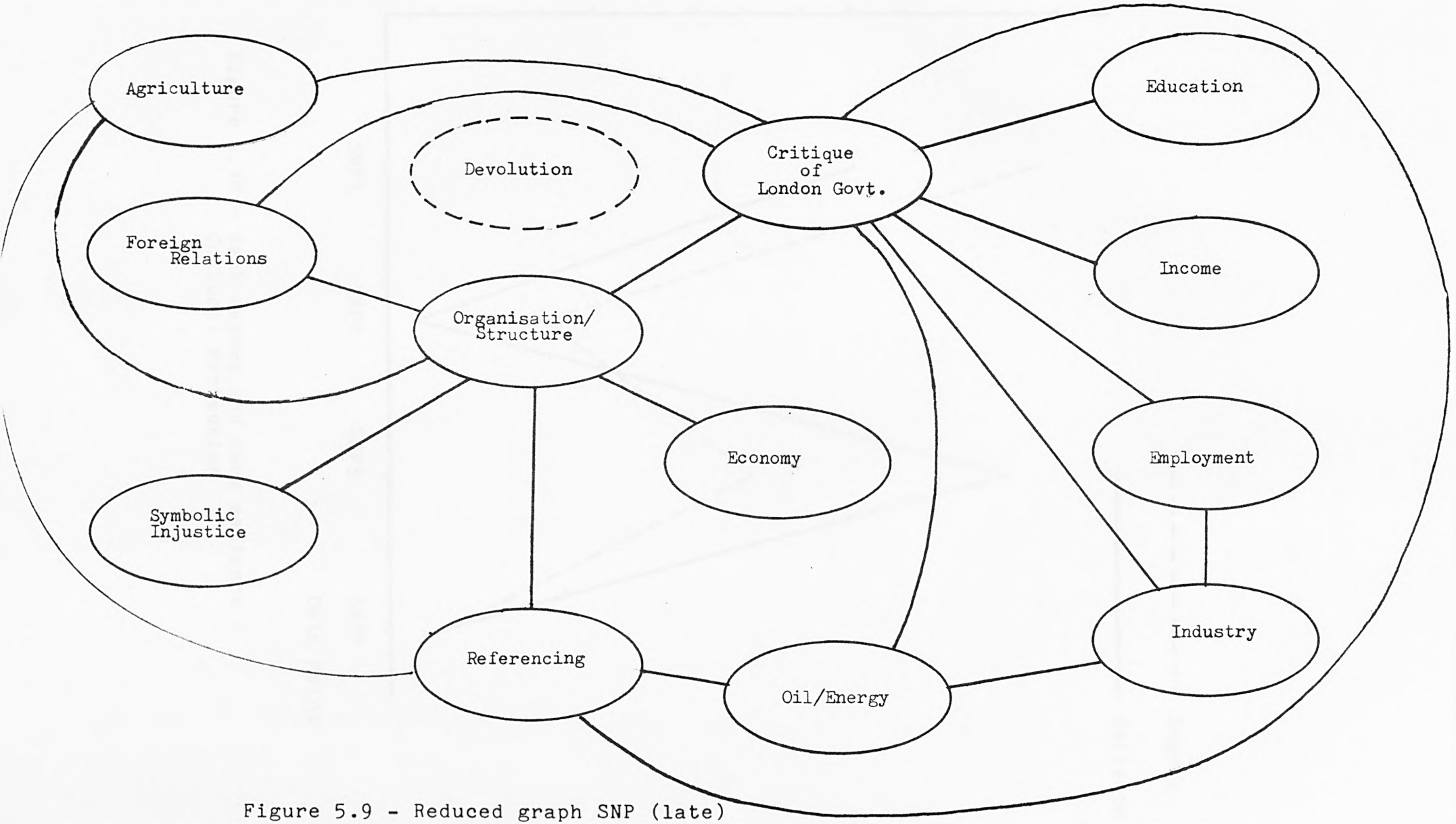


Figure 5.9 - Reduced graph SNP (late)
- "Devolution" removed

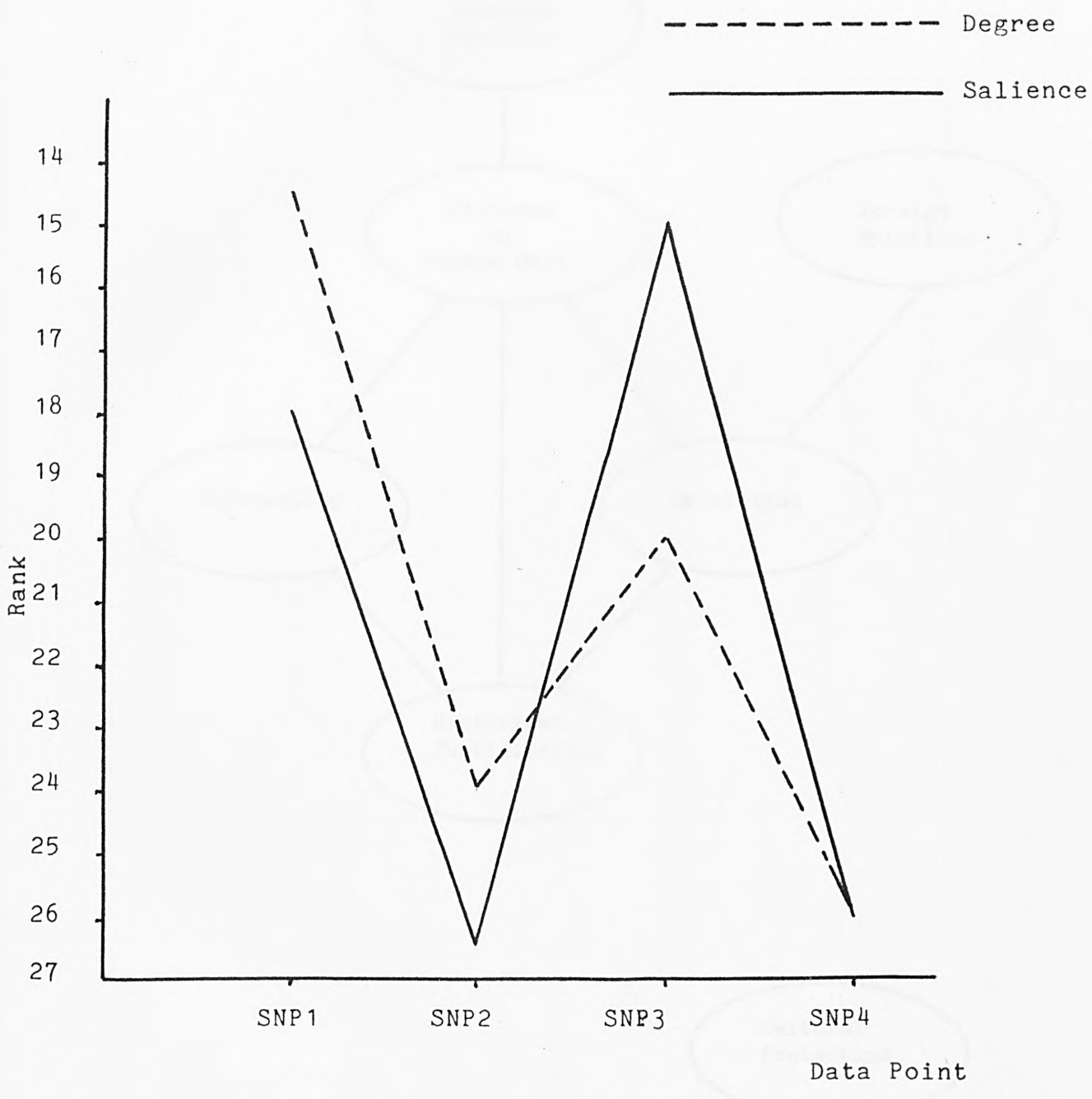


Figure 5.10 - Rank degree and rank salience - Cultural Protection

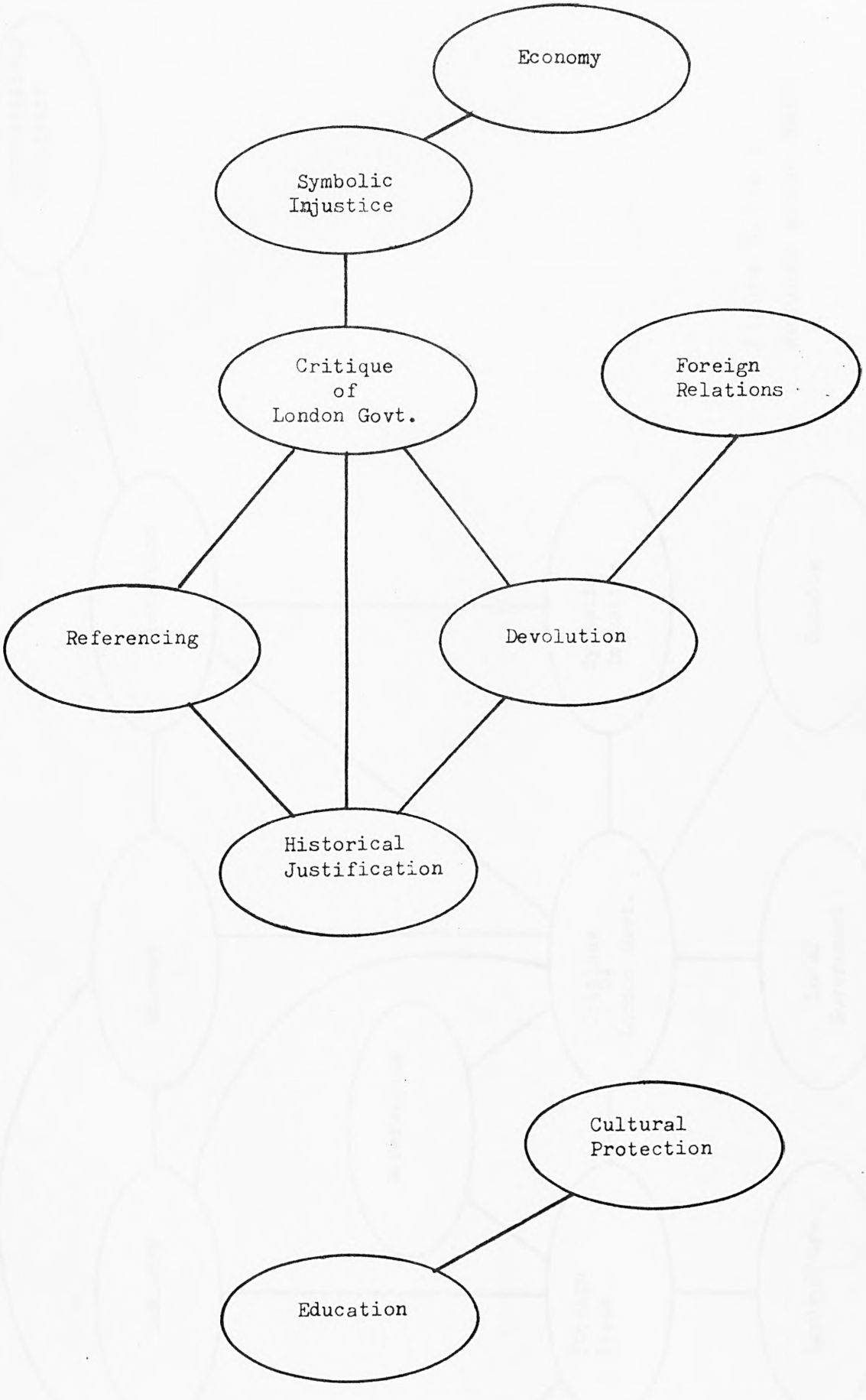


Figure 5.11a - Reduced graph SNP1

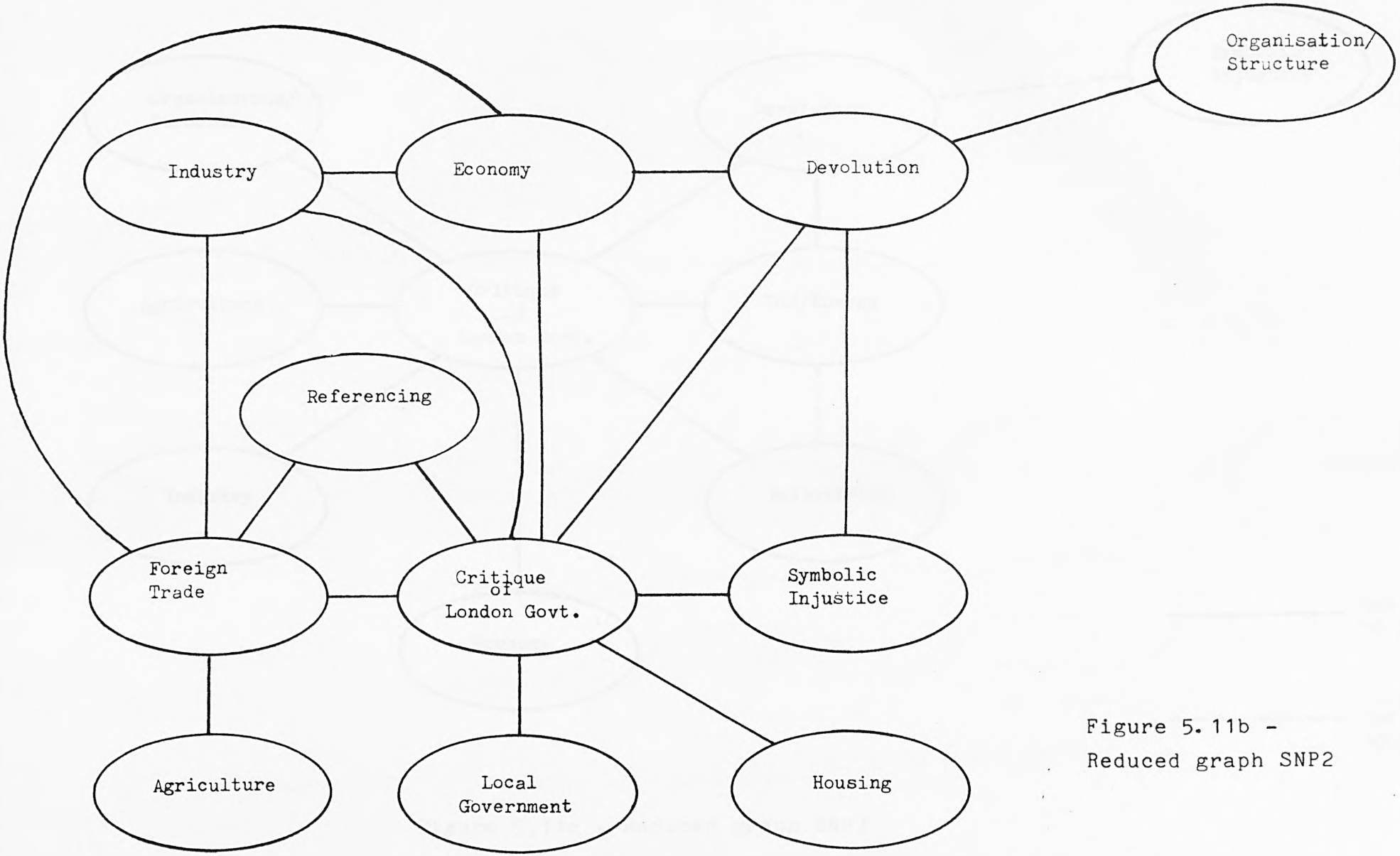


Figure 5.11b -
Reduced graph SNP2

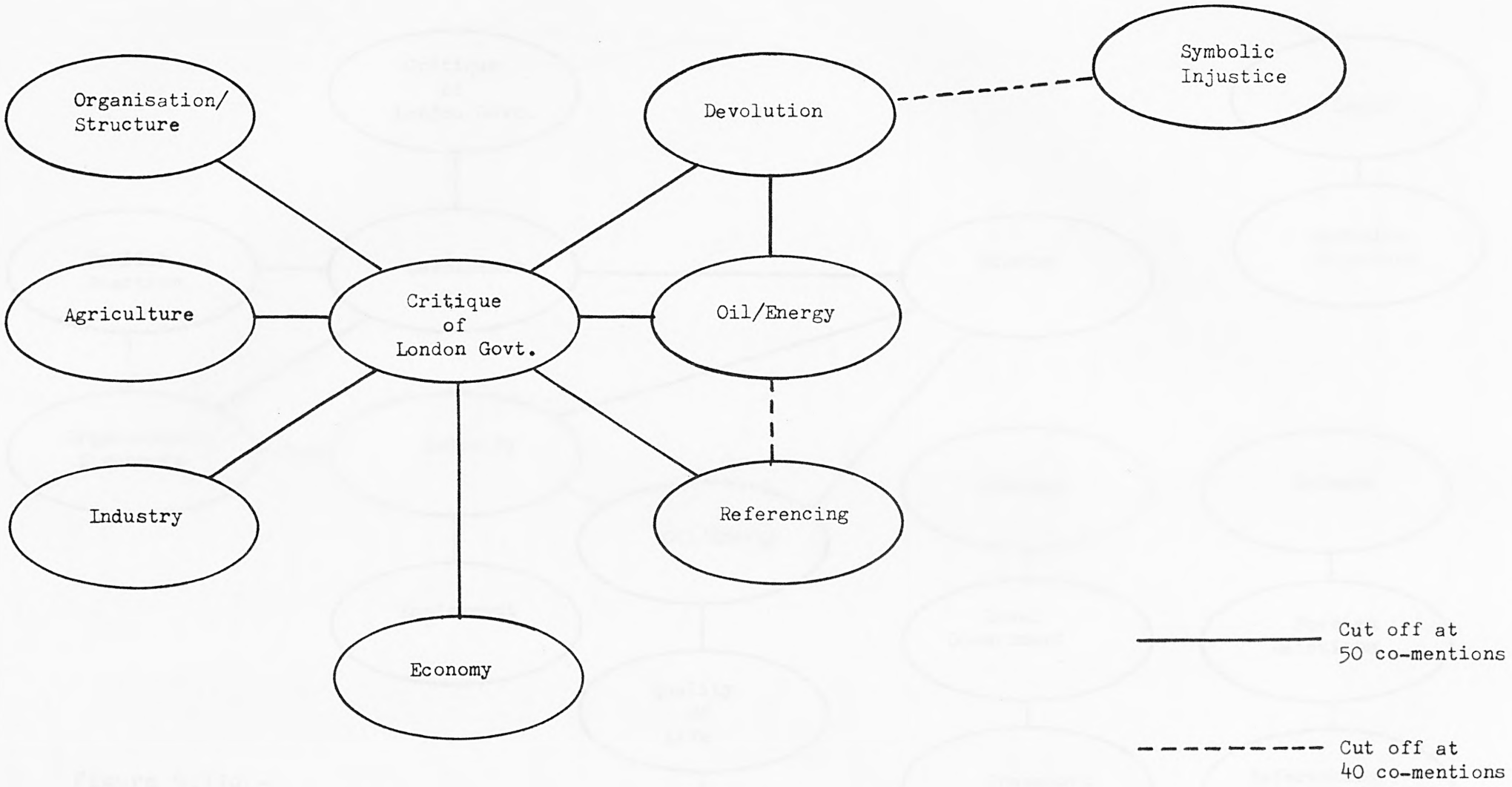


Figure 5.11c - Reduced graph SNP3

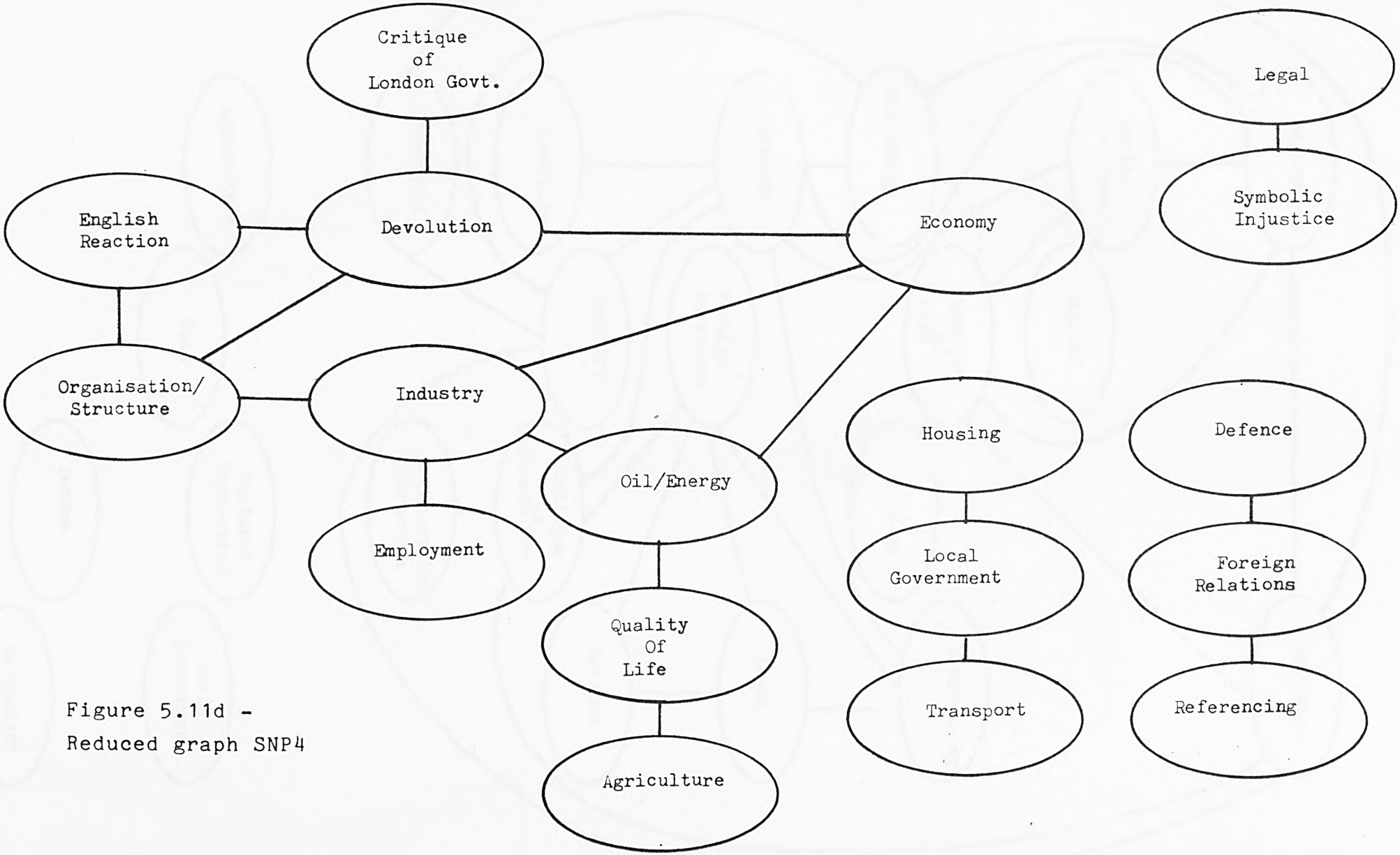


Figure 5.11d -
Reduced graph SNP4

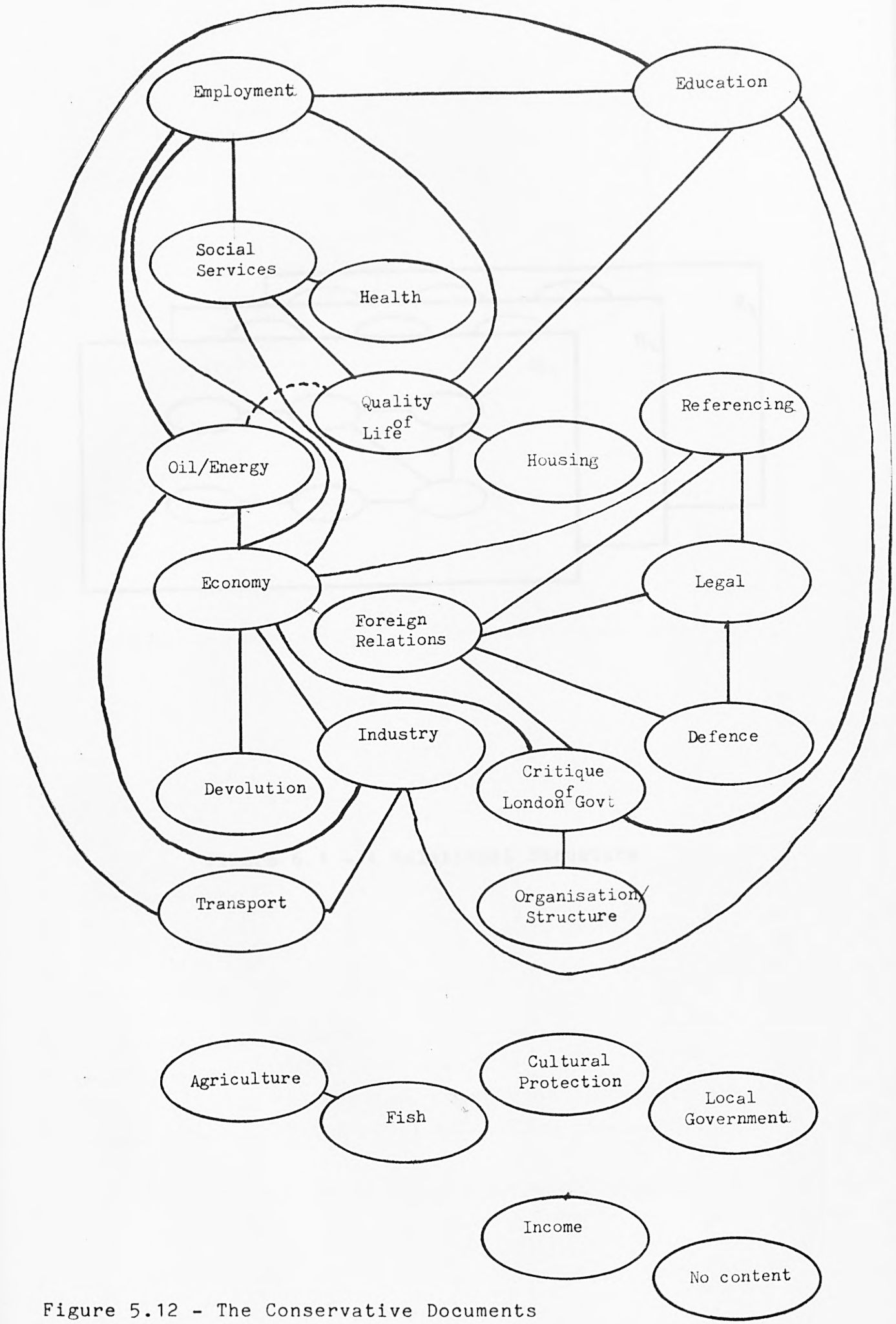


Figure 5.12 - The Conservative Documents

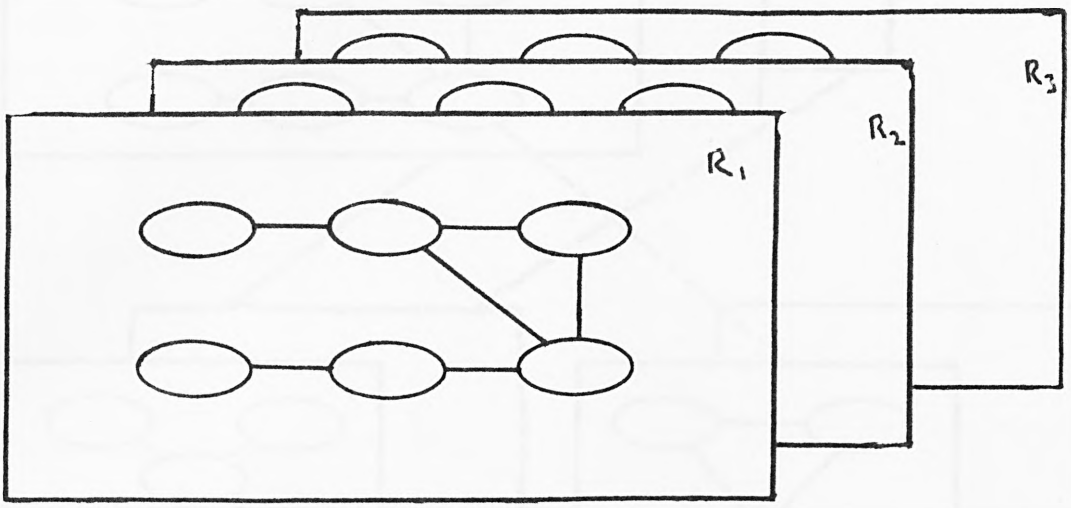


Figure 6.1 - A Relational Structure

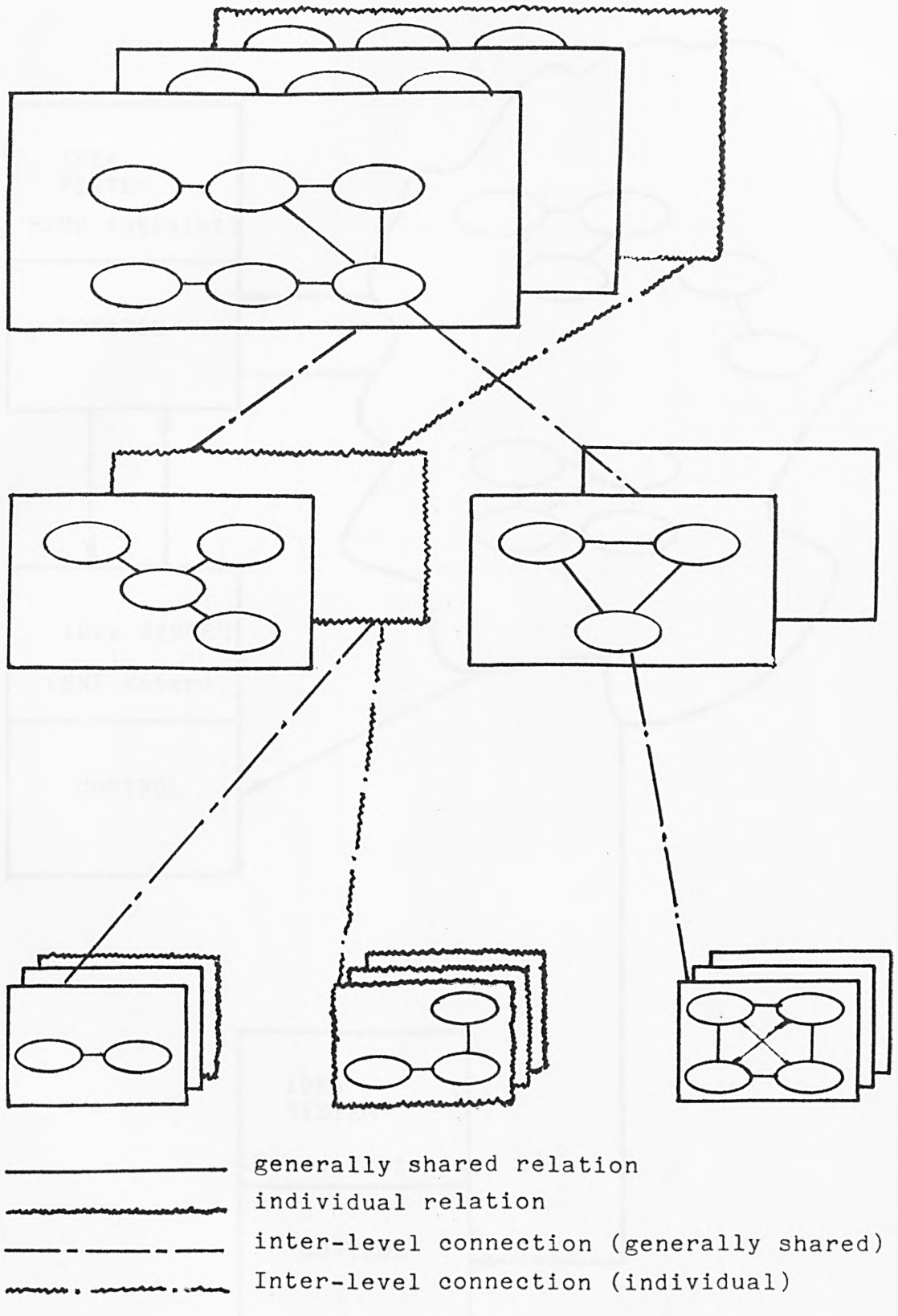


Figure 7.1 - A Multi-level Relational Structure

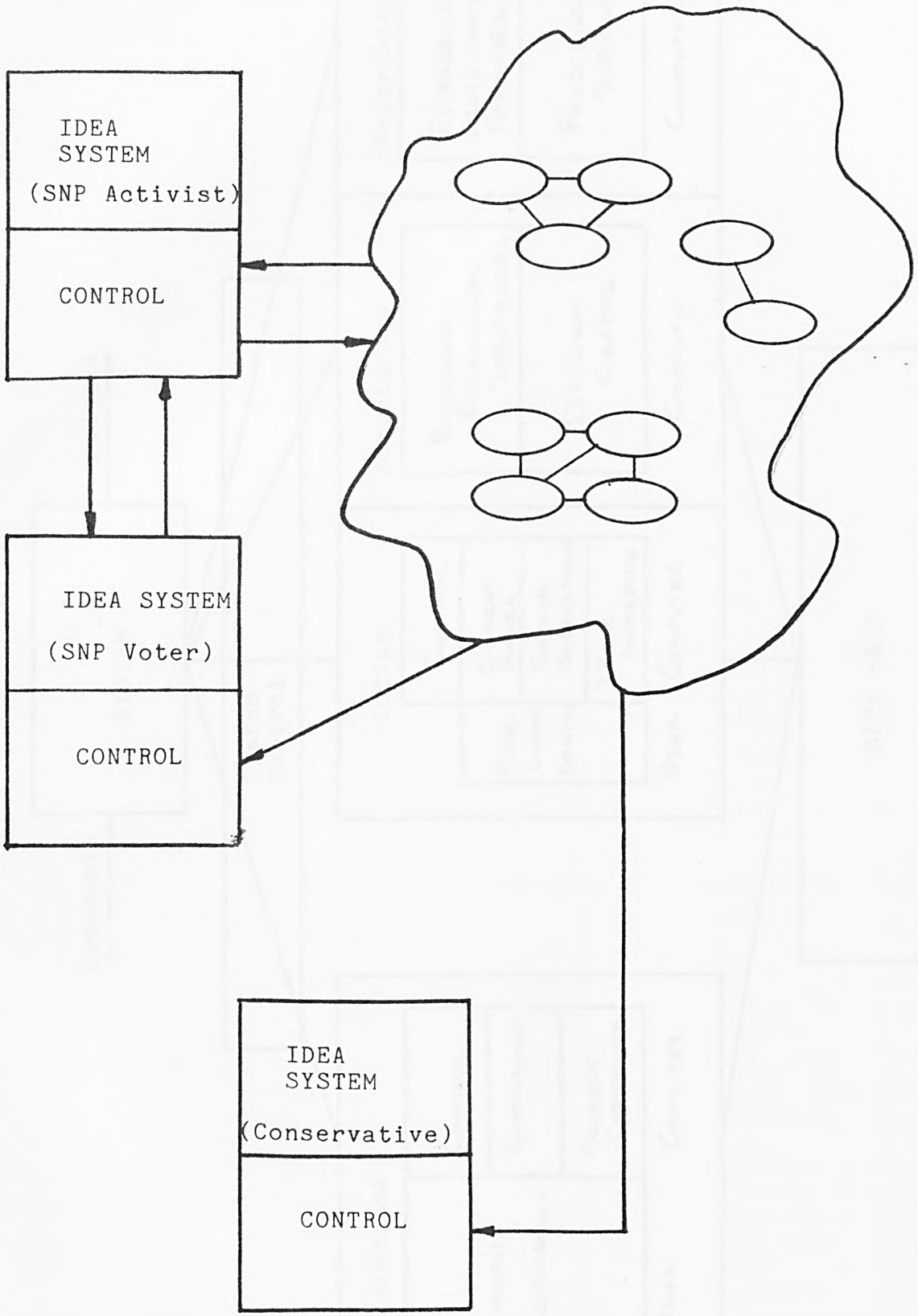


Figure 7.2 - The Environment of an Idea-System

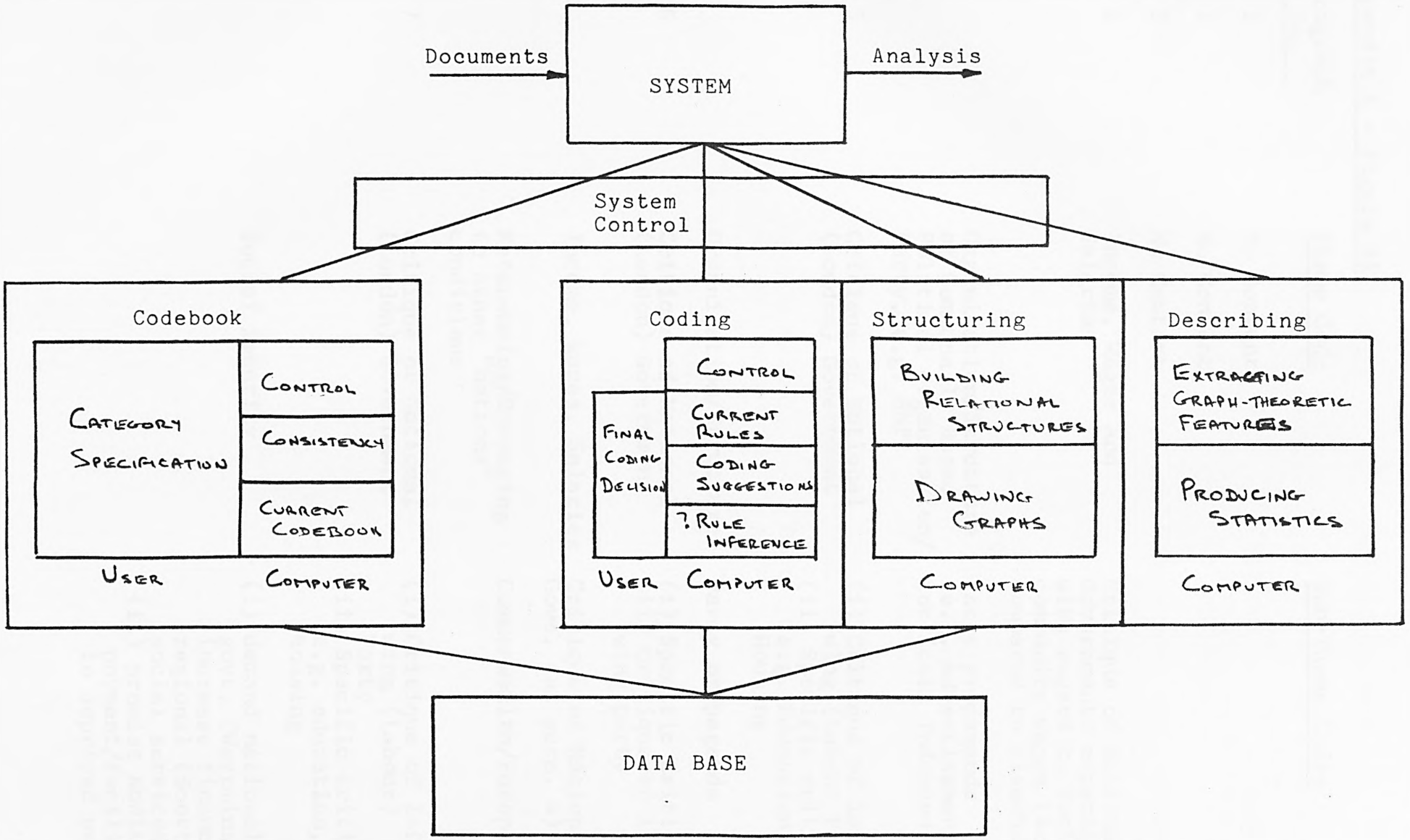


Figure 7.3 - A top level system architecture

Appendix 1 - Figure A1

Paragraph No.	Theme Code	Sub-Theme Codes
1	No Content	
2	No Content	
3	No Content	
4	Income, Wages and Salaries	Critique of National Government: especially with regard to lack of community wages (Scots) compared to elsewhere
	Organisation/Structure of Communally Based Political Organisation/ Party, e.g. SNP	Party propaganda (e.g. advertisement for Scots Independent)
5	Critique of National (London) Government	(i) Critique of left-wing (Labour Party) (ii) Specific critique e.g. Education, Housing
	Organisation/Structure	Party propaganda
6	Critique of National (London) Government	(i) Specific critique (ii) Critique of left-wing party
	Income, Wages, Salaries	Critique of National Govt. (as para. 4)
	Referencing/Comparing to other "nations" conditions	Commonwealth/Europe
7	Critique of National (London) Government	(i) Critique of left-wing (Labour) Party (ii) Specific critique e.g. education, housing
	Social Security	(i) demand national govt. (Westminster) increase finance to regional (Scottish) social services (ii) promise additional payment/facilities to deprived person:

Paragraph No.	Theme Code	Sub-Theme Codes
7 (cont.)	Natural Resources: Energy/Oil	(i) Use of oil revenues - Social Security (ii) Use of oil revenues - Quality of life/ environment
8	Social Security	(i) Pension obligations honoured in an indep. Scotland (ii) Promise additional payment/facilities to deprived persons
	Organisation/ Structure (as in para. 4)	(i) SNP post independence (ii) Party propaganda (as in para. 4)
9	Devolution	Unspecified demand
	Critique of National (London) Government	Critique of left-wing (Labour) party
	Organisation/ Structure (as in para. 4)	Party propaganda (as in para. 4)
10	Natural Resources: Energy/Oil	(i) Statement of community ownership of natural resources (especially oil) (ii) Use of oil revenues - Social Security
	Critique of National (London) Government	(i) specific critique (ii) critique of left-wing (Labour) party
	Referencing (as in para. 6)	Scandinavia

FIGURE A1 - A specimen coding

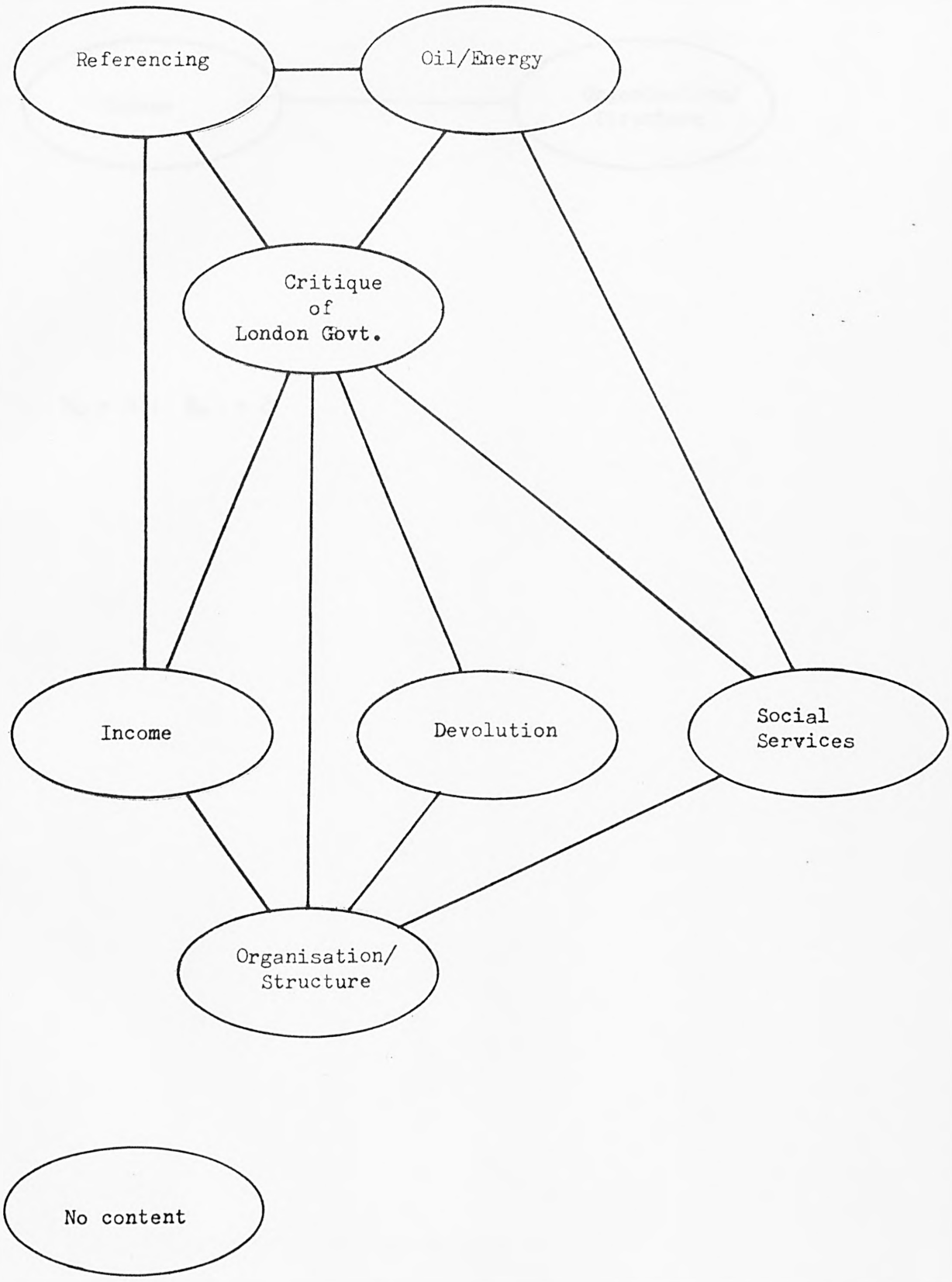
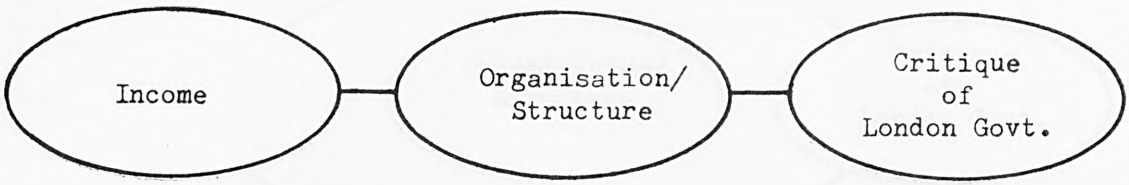


Figure A2 - The graph of document 119

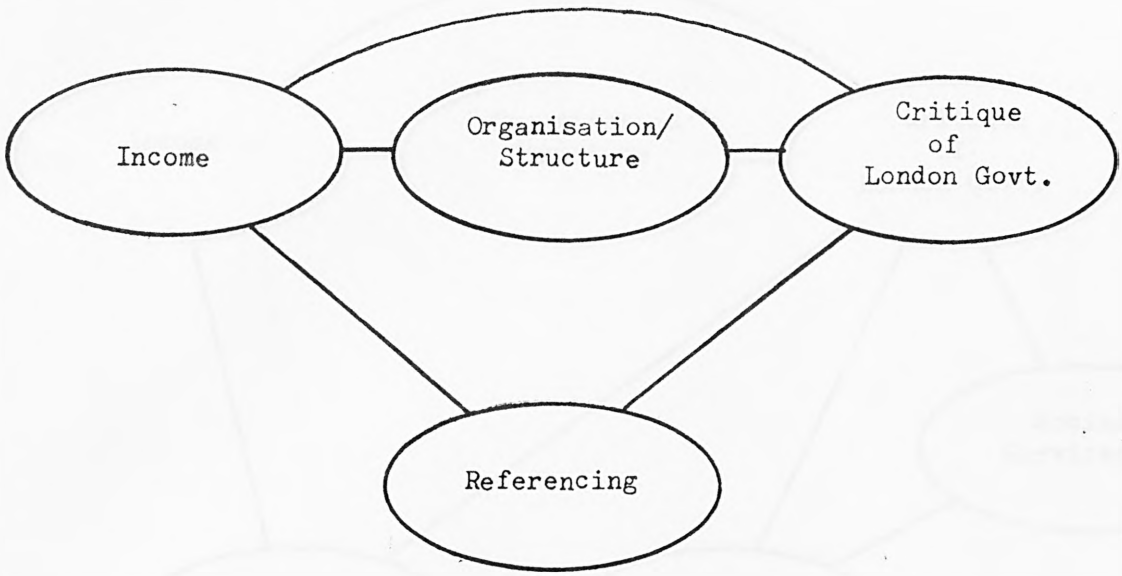
$M_p = 0$; $M_s = 2$

Figure A3 - - The emerging pattern of mobility
(at paragraph 4)



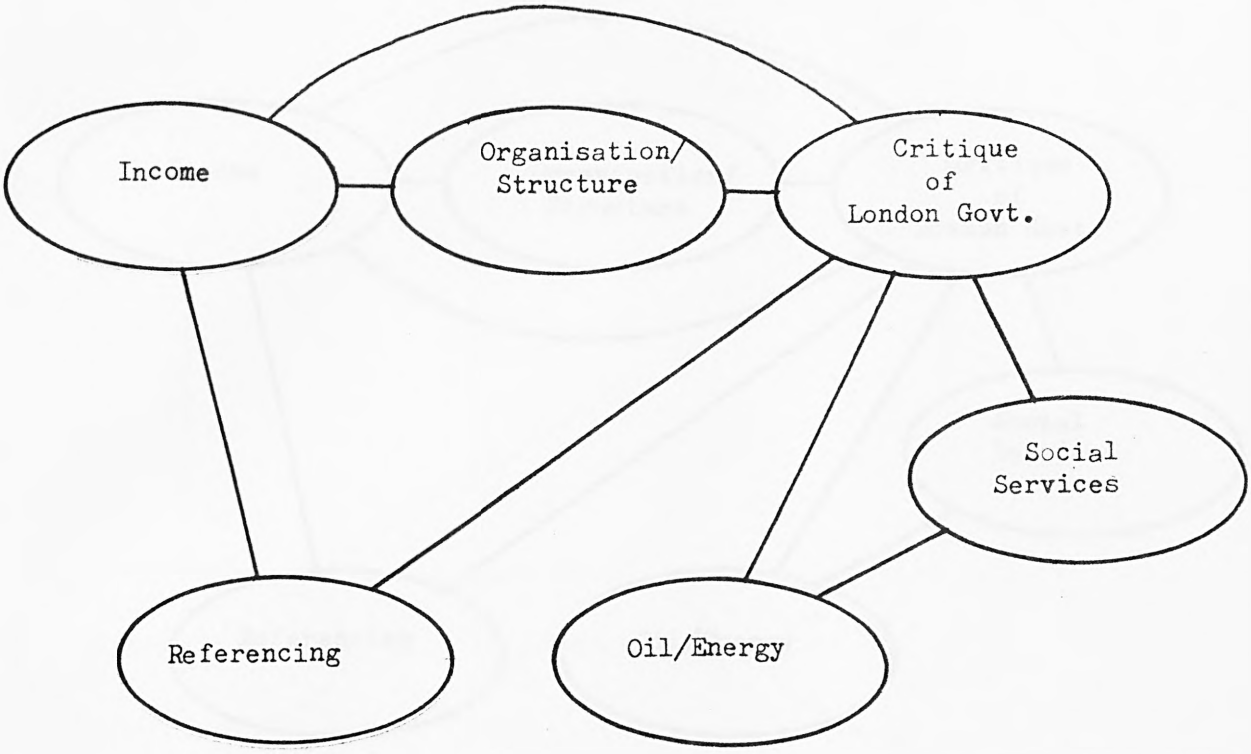
$M_p = 1$; $M_s = 4$

Figure A3 - at paragraph 5



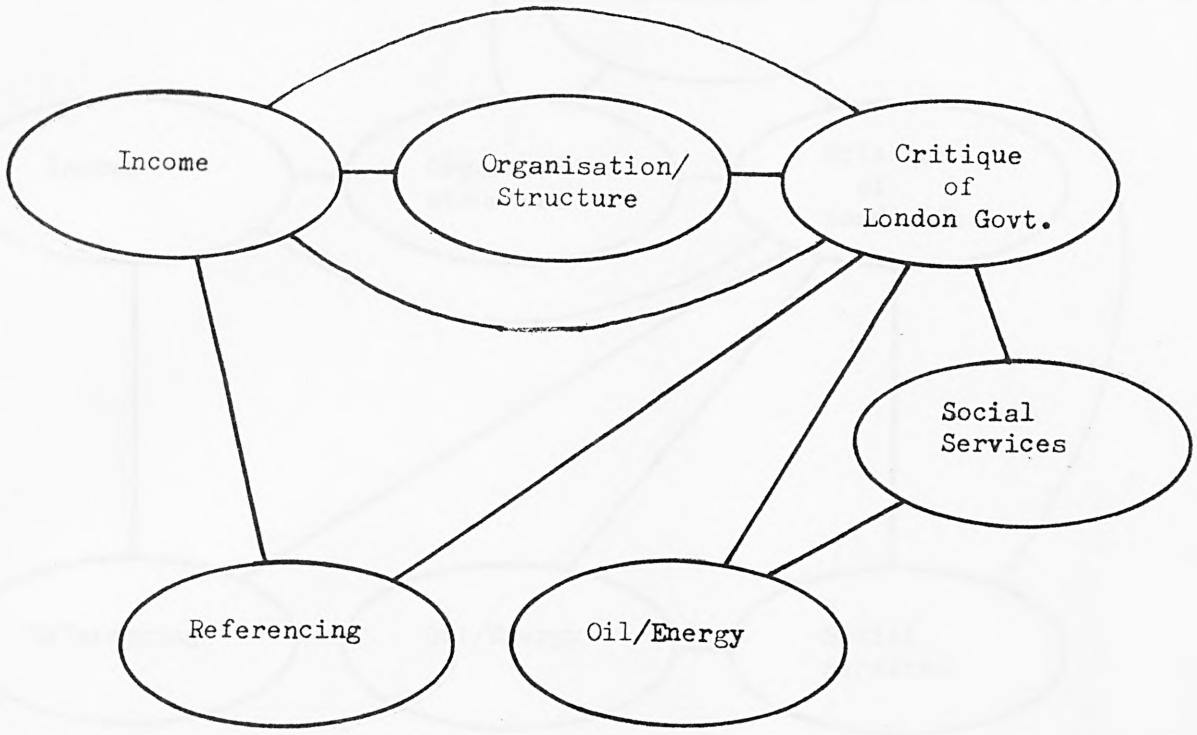
$M_p = 0$; $M_s = 4$

Figure A3 - at paragraph 6



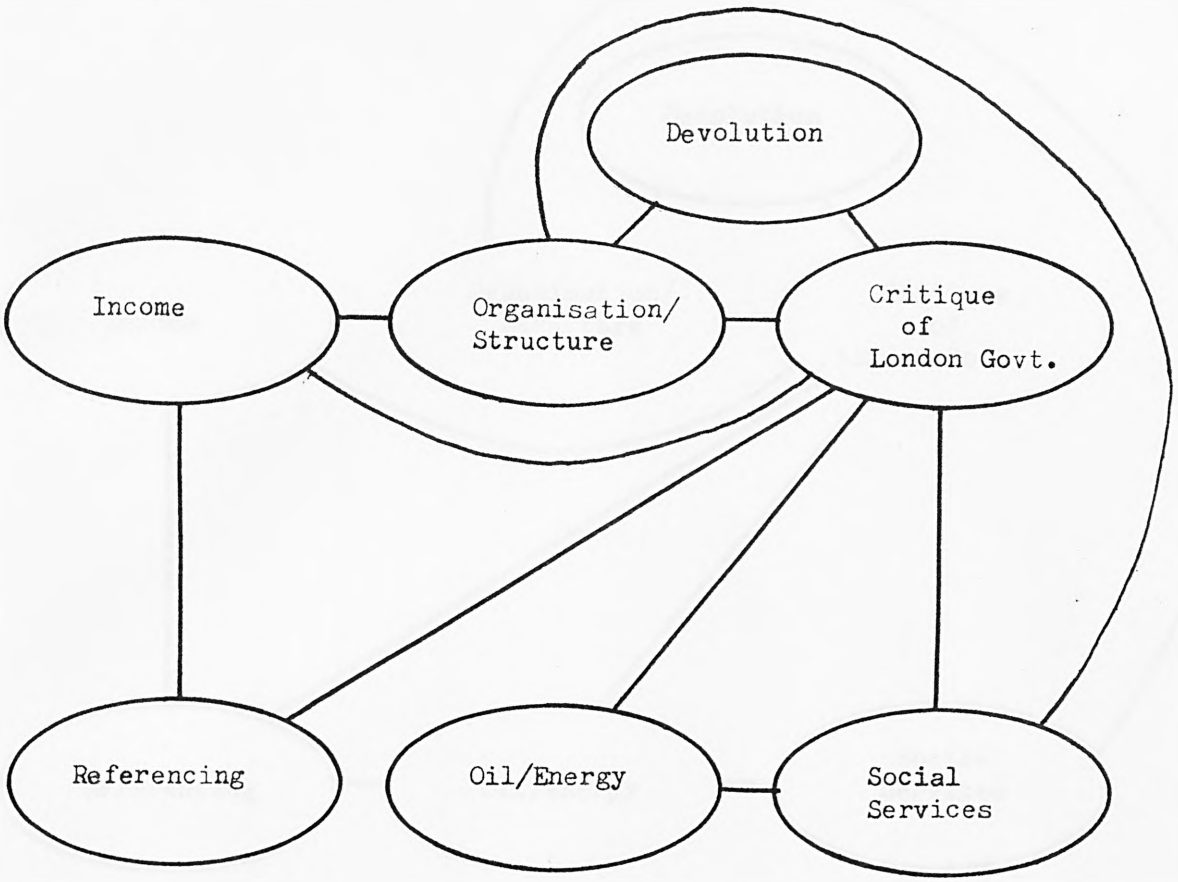
$M_p = 1$; $M_s = 8$

Figure A3 - at paragraph 7



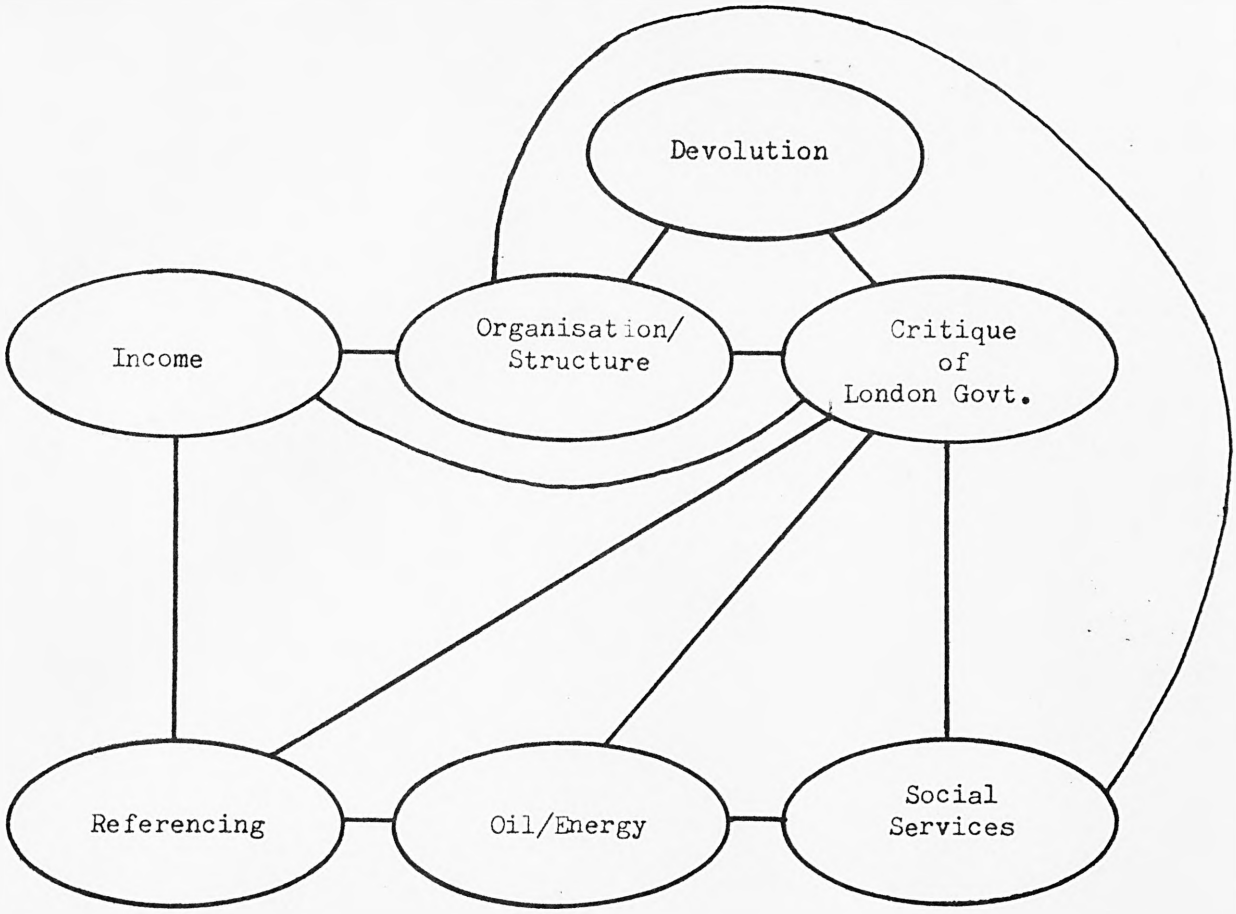
$M_p = 0$; $M_s = 6$

Figure A3 - at paragraph 8



$M_p = 0$; $M_S = 7$

Figure A3 - at paragraph 9



$M_p = 0$; $M_s = 6$

Figure A3 - at paragraph 10