Computer Based Content Analysis on the Party Manifestos of the 1998 Dutch Elections

Miranda W.M. de Vries
Department of Political Science
School of Public Affairs
University of Nijmegen
PO Box 9108
6500 HK Nijmegen
The Netherlands
M.deVries@bw.kun.nl

Paper for the ECPR Joint Sessions, University of Mannheim, 26-31 March 1999
Workshop on Estimating the Policy Positions of Political Actors
1. Introduction

The main goal of this paper is to provide positions of political parties on a number of ideological dimensions for the Dutch political parties in 1998. Even though only 1998 will be studied in this research, the idea is that in order to study party competition or cabinet formation, valid policy positions of political parties are necessary. Therefore also other databases and methods for obtaining party positions than the data and method applied for 1998, will be reviewed shortly in this paper. The party positions of 1998 will be obtained by means of computer-based content analysis of the party manifestos.

In the second section, different data sets and different means of collecting party positions will be introduced. After that, a short introduction on content analysis will be given. The actual content analysis of the party manifestos of the Dutch political parties, in 1998, will be studied in the fourth section. A short summary and discussion will be given in the last section.

In the context of party positions, an ideological dimension is ideally a coherent group of issues that form an idea about how society should be organised. These ideas about how society should be organised need only hold for one country at one period. And this ideology should be the same for both the political leaders and the citizens. For instance, on the traditional left-right dimension, parties on the left side claim that more state interference in the economy for social equality is better than less state interference, whereas the parties on the right side of the spectrum claim the opposite. This dimension gives us information about how parties and citizens that support these parties feel about tax policy, welfare state expansion, public vs. private ownership etc. Another dimension is often an immaterial values dimension that reflects liberal vs. authoritarian views on issues such as abortion, euthanasia, freedom and religion.

Important for research on party competition and coalition formation is that these so-called 'ideological' dimensions reduce the high dimensional policy space, into a low dimensional space. Even if this low dimensional space is obtained inductively, by using methodological techniques to reduce dimensionality, a low dimensional space is preferable because it makes testing party competition models or cabinet formation easier.

2. Data sets

In order to determine positions of political parties on a number of ideological dimensions, the first step is to decide which kind of data to use. The following types of data sets for positioning parties in the Netherlands are at hand:
- Election surveys
- Voting on legislation
- Data on expenditure
- Expert data on positions of parties on issues
- Manifesto data

The first database, the Dutch Election Studies (NKO) is held around every general election ever since 1967. Since 1977, questions on perceived policy positions of parties on some issues are included. The respondents are asked to place parties on seven or ten point scales. For instance: "Euthanasia: Where would you place the PvdA (i.e. the social democratic party) on this line? 1. Forbid euthanasia, ..., 7. Allow euthanasia” or "Income Differences: Where would you place the VVD (right wing liberals) on this line? 1. Larger differences, ..., 7. Smaller differences". The information gathered by these questions is the kind of material that is necessary for placing parties on more than one dimension. A drawback however, is that these questions on issues are asked since 1977, which means that there are only seven cases, i.e. the elections or coalition formation in 1977, 1981, 1982, 1986, 1989, 1994 and 1998 (forthcoming). Another disadvantage of these election studies is that these surveys only contain information of positions of parties on a small number of policy issues. On average, five or six questions about party positions on issues are asked. This is not very much, since I want to examine the dimensionality of the party-space and then make a configuration of the political parties. In order to obtain ideological dimensions that are predictive for policy positions of parties, more policy issues are needed.

Also technically, the methods like Factor Analysis or Multi-Dimensional Scaling that can be applied in order to reduce the issue space, need more variables in relation to the number of parties. Another problem with these mass-survey data is that they represent what the voters think the parties should do, and not what the parties actually do.

Voting on proposals for legislation and expenditure data seem suitable because these data are 'hard evidence', but they have a disadvantage too. Namely, it is often hard to distinguish between coalition partners in these data. During a governmental period, parties that govern together tend to spend money together and vote together. Moreover, package deals are often made, which makes it even more difficult to determine policy positions of individual parties. Even though expenditure data actually show what parties do when they are elected, - instead of reflecting promises - the fact that parties do not govern alone make these data unfit for our task too. Also, Kraan (1990) demonstrated in his thesis that different governments hardly show any differences in expenditure patterns. The expenditure- and voting data are, to my opinion, useful if the goal is to study cabinet-parties versus opposition-parties, but they are less valid for obtaining 'individual' party positions.
Several expert-survey databases are not very useful for our goal, since they have only gathered positions of political parties on one ‘left-right’ dimension (Castles and Mair, 1984). Since I want to obtain party positions on a number of dimensions, these data are not sufficient. For an overview of expert judgements on one left-right dimension, I refer to the appendix in Multiparty Government (Laver and Schofield, 1990).

An important expert survey on party positions is the Laver and Hunt (1992) database. In this survey, country experts are asked to assign policy positions to political parties on several issues. This database is very broad and interesting questions are dealt with. For one reason, the questions are interesting because they are designed for the 1989 and adjusted for 1994, which means that issues that are particularly important in these periods are included. Whereas in other data sets important issues are missing and issues that are no longer political issues - i.e. issues that divide parties - are still included. In this survey, country experts have positioned parties on eight policy issues, on a scale from 1 to 20. It is also interesting that, in contrast to other data sets, information about the weights of these policy issues for the parties, and data on policy- versus power motivation of political parties are included. Despite the fact that the small number of cases makes this otherwise attractive data-set unfit for testing party competition models or coalition formation in retrospect, this is an important development.

The manifesto data set is the largest data set on party positions⁴. Party manifestos in 20 democracies, since 1945, are studied by means of content analysis. However, content analysis is said to be a dubious method to gather information about party positions because the procedure counts the number of sentences that are concerned with each issue, i.e. it measures the salience of issues. Thus, the scores on the issues do not inform us about the positions of parties on these issues. Most issues in the data base are valence issues which means that all parties agree on the goal - like environmental protection -, but probably differ on how to reach this goal, or which goal or issue should get priority. In less than half of the cases, positive and negative attitudes on issues are distinguished. Moreover, the coding scheme provides ‘only’ 56 issues divided in seven policy domains. This seems a large set of issues, but uncertainty about the importance of the various issues present, and the importance of the issues that are absent in the coding scheme, has risen. One of the strong points of the data set, namely that so many countries can be compared, because the same coding scheme is used all over the world, and throughout the whole post war period, is to my opinion also the downfall of this data set. It has led to issues that do not seem to fit to the Dutch political situation. Issues like ‘political corruption’, ‘foreign special relationships - negative’, ‘economic planning’, ‘controlled economy’ and ‘nationalisation’, are not political issues in the Netherlands, but are nevertheless included in the coding

scheme. This is a small problem since these issues simply get extremely low or zero scores, and low or zero variance.

On the other hand, significant issues that should be included in the analysis but are not, i.e. issues on unemployment, inflation, and women's liberation, also cause problems. Many immaterial issues like abortion, religion, homosexuality and euthanasia are, as a group incorporated in issues in the data set, for instance in issues about 'traditional morality' and 'special social groups'. I believe that these immaterial issues deserve their own place in the coding scheme, and should be coded separately.

For the Netherlands, I have checked the relation between issues within a domain with factor analysis. Harman in his “Modern Factor Analysis” gives a profound overview of factor analysis. In this work both the foundation and mathematical background, as well as different models of factor analysis are presented (Harman, 1976). Factor analysis refers to a set of techniques whose goal it is to represent a large set of variables into a smaller set of hypothetical variables. Starting point for any kind of factor analysis is the correlation matrix of the variables. If any set of variables is highly correlated, it can be treated as an imperfect measure of the same phenomena. If at the same time several sets of variables are mutually independent, this can mean that the data set consists of a series of uncorrelated 'latent' variables (Kim and Mueller, 1978). In two different kinds of situations, factor analysis can be performed. At the first extreme, the researcher has absolutely no idea about what and how many underlying dimensions there are. In that case, exploratory factor analysis is used as a way to find the minimum number of latent factors that can account for the observed covariation. In this type of analysis the data are explored for possible data reduction. On the other extreme, the researcher anticipates for instance two factors, and uses the analysis as a means to test his hypothesis. This is referred to as confirmatory factor analysis (Scott Long, 1983).

I have performed confirmatory factor analysis for the Netherlands (1946-1994) to examine whether the issues within the seven domains are highly correlated, and thus are different issues that actually measure the same phenomena. The expectation is that issues within one domain are highly correlated and can be represented with one latent dimension. For instance, the first ten issues (per101 - per110) measure the extent to which parties approve or object to further increase of the external relations of a state. Factor analysis of these ten issues leads to four factors - underlying dimensions - all with an eigenvalue greater than one. The eigenvalue is the most common criterion by which the number of common factors (latent dimensions) can be determined. The average number of underlying dimensions that are detected when performing factor analysis on these groups of issues is three per domain. This does not comply with my expectations, it is one underlying dimension for every domain. Even two dimensions could be satisfying as long as the explained variance of the first factor would be high. For all policy domains the explained variance for the first factor is low. The highest explained variance for the first factor is in domain two 'Freedom and Democracy' which consists of only four issues, and adds
up to 40%. In summary, the confirmatory factor analysis did not confirm the expectation that within each domain one latent phenomena is examined.

Another method to examine the coherence between the issues within a domain is reliability analysis. With this method, scales can be constructed of a variety of related - correlated - items. In a scale, responses to each of the items can be summed and result in a score for each case, i.e. in our case a score on for example external relations for every party. Reliability analysis starts with constructing a scale consisting of items that the researcher believes ‘belong together’. If the correlation between the items in a scale and the scale itself are high, the scale is said to be reliable. If, for a particular item the correlation between this variable and the scale is low, this item can be deleted and the reliability of the scale improves. The reliability analysis uses a coefficient to test the internal consistency of a scale. The most common coefficient is Cronbach’s alpha, which is based on the average correlation of items within a test. Since the alpha is an average correlation it ranges between zero and one; the closer the coefficient is to one, the better the scale. For the items within a domain, reliability analysis is performed, not to construct scales, but again to examine the coherence of the domains. Again the results are not promising, the best scales i.e. the scales with the highest Cronbach alpha are for domains two "Freedom and Democracy" and six "Fabric of Society". The scales have alpha's .39 and .36 respectively. These alphas can improve if items are deleted, but this is not done since this scaling was performed to get an impression of the coherence of the issues within a domain, and not to construct scales.

In this section, an assessment of content analysis of the manifesto data is presented. In summary, we have studied how the data are obtained, what the data look like, and how the data are divided in seven policy domains. Also the absence of some important issues and the presence of unimportant issues, and the lack of coherence between the various issues within the domains is discussed. I acknowledge these drawbacks, but I have used this data set and applied multi dimensional scaling for obtaining party positions for the Netherlands in the period 1945-1994 (De Vries, 1999). Moreover, note that this data set is used often in research on party positions (see among others Klingemann et. al, 1994). The reason is that this data set includes more issues than other data sets, and it also runs over a longer period of time. This data set is the largest and most comprehensible set available.

The last option (that I discuss) is to perform computer-based content analysis also based on the party manifestos. This is not done yet for the Netherlands but will be performed for the Dutch 1998 party positions in this paper. Unfortunately, I will not be able to compare the results of this analysis with the other data sets discussed in this section, since at the time of writing, other data for 1998 have not been published yet.
As is said, for policy based coalition formation theories and party competition models, we need multi-dimensional data on party policy positions. The only suitable database in retrospect is manifesto data set, but this is by no means a perfect database especially since it was not designed to measure policy positions. So, the lack of systematic data on more than a single dimension of ideology makes obtaining policy positions and thus testing theories a tricky enterprise. It is therefore, to my opinion absolutely necessary to develop good data sets on party positions.

3 Computer-Based Content Analysis of the Party Manifestos: theory

Weber (1990) defines content analysis as "a research method that uses a set of procedures to make valid inferences from text". Content analysis is basically a method for data reduction. In 1969, Holsti defined three requirements for content analysis:
1. Objectivity: which means that each step in the research process must be carried out on the basis of explicitly formulated rules and procedures
2. Systematic: the choice of categories for coding should be done systematically, i.e. consistent to rules that are defined to answer the research question.
3. Generality: the finding should have theoretical relevance (Holsti, 1969).

In 1990, Weber suggests that content analysis should meet reliability and validity requirements. Nowadays, these two requirements are most referred to in texts on content analysis. The first demand - reliability - is very much like objectivity in Holsti’s work. This means that the results of content classification should be invariant over time and over different coders. The last requirement - validity - refers to the content of the research and is more or less the ‘systematic’ and ‘generality’ requirement together. The method performed by the manifesto research group is referred to as quantitative content analysis, carried out by expert coders. In the manifesto content analysis documents are selected, a coding scheme is constructed, and a counting unit is selected with which the programmes can be analysed.

The original goal of the ECPR research group of the manifesto data (MRG), was to operationalise a specific model of party competition, which held that parties compete with each other in terms of salience of particular issues. So, the manifesto data measure the relative emphasis placed on an issue and not the parties' positions on the issues. Unfortunately, the scarcity of policy data has led to using the MRG data as policy positions.

As we saw in the previous section, obtaining policy positions of parties in the presence can be done with any kind of method such as expert surveys, elite surveys, mass surveys, or content analysis. In order to test theories on party competition or coalition formation in retrospect, the data on party
positions should be improved. Since for past policy positions the formal texts, i.e. the party programs, are the best possible information sources, Laver and Garry (1998) propose two important improvements for the Manifesto data. They suggest the following two different methods to refine the content analysis:

- To redo the content analysis with a fine grained coding scheme. This new coding scheme contains at least bipolar and most of the time three-polar issues (i.e. pro, con and neutral), and also the issues are hierarchically structured, starting with defining to which domain an issue belongs and then refining more and more until it fits in a specific category. This method is similar to the original manifesto coding since expert coders perform the coding. The improvement lies in the more refined coding scheme, which allows us to obtain - positional - policy positions instead of measuring emphasis and at the same time is expected to have higher coder reliability.

- The second, more revolutionary, alternative to the qualitative expert coding is quantitative computer coding, which distinguishes itself from the former method by perfect ‘coder’ reliability and ease of use especially when large documents are considered (Laver and Garry, 1998).

I have decided to use the latter method for content analysis of the Dutch manifestos of 1998, for the following reasons. First, the reliability is better than in expert coding. Secondly, although computer coding is said to be less valid, because it mechanically codes without being able to ‘understand’ the context, the other side of the coin is that expert coders are almost inevitably biased. The coder - being an expert - is familiar with the basic ideas of the parties, and can thus be inclined to interpret similar statements differently, if they belong to different parties. So, computer coding has validity advantages too. Thirdly, also the fact that the first results of computer coding of the manifestos by Laver and Garry are promising has led to this decision too. Laver and Garry have performed computer coding as well as the revised - fine grained - expert coding for Britain in 1989 and 1997 and Ireland in 1992 and 1997. They compared the results of these methods with the original MRG results and the Laver and Hunt expert surveys. They consider their results encouraging because the cross validation of the different methods is good, i.e. high Pearson correlations exist between the party positions derived by different methods. The estimates of economic left-right positions showed correlations above .84 (Laver and Garry, 1998). In conclusion Laver and Garry suggest that “computer coding of huge volumes of virgin text may be a viable undertaking” (Laver and Garry, 1998).

In the next section the computer content analysis of the Dutch manifestos in 1998 will be performed. The following steps, which are the basic steps for any kind of content analysis, will also be taken in order to perform computer content analysis for the manifestos of 1998 for the Netherlands:
- Selection of the documents to be coded
- Defining the coding units (sentence, quasi-sentence or word)
- Defining the coding categories (dictionaries)
- Performing the analysis.

4. **Computer-Based Content Analysis of the Party Manifestos: practice**

High reliability of computer coding and promising empirical results of Laver and Garry’s content analysis, are parts of the reason to perform this type of analysis for the Netherlands in 1998. The criticism on content analysis, namely that it provides salience scores and not positional scores, will be met in this analysis. For the main ideological dimensions positional scores will be derived. Only for the valence dimension 'environmental protection' saliency is interpreted as positional scores.

Also an important reason for this analysis is the urge to explore possible methods that might solve the lack of reliable, and valid estimates of party positions in retrospect. Better methods will result in better estimates of policy positions, which will improve empirical testing of studies on party competition in general and coalition formation in particular. This section should therefore not only be seen as a means to collect data for 1998, but most of all as a first effort towards improving the data on policy positions in general. However, this is a tryout and many improvements should be made before applying this method in retrospect. Especially the dictionaries should be refined, before applying the computer based content analysis in retrospect.

Computer based content analysis is performed to obtain policy positions of the main political parties, in 1998. The first step, was to choose the documents to be analysed. The five largest parties - CDA, D66, GrLi, PvdA, and VVD - are included in the analysis, and the documents are the party manifestos. In this method, the next step is to design dictionaries of words that are relevant for a particular policy domain. These dictionaries are the coding schemes. In this procedure, the coding unit is a word. If these steps are accomplished, the computer analyses the texts word by word, and counts the number of words associated with each coding category.

After choosing the coding unit and the documents, the dictionaries have to be developed. Designing the dictionaries is a vital part of content analysis. Since there are no good dictionaries at hand, I decided to design my own dictionary. I did not have to start from scratch, because I decided to use the policy domains distinguished by Laver and Garry (1998). They applied the following categories for both the refined expert coding and the computer coding:

- Economic policy
- Social values
- Political reform
- Law and order
- Environmental policy
The next step is to allocate words to these coding categories. The dictionaries, made by Laver and Garry, are used as reference dictionaries. These reference dictionaries, the manifestos, and common sense was used, to design the dictionaries. It is preferable to use other important documents of political parties in the same period, to develop the dictionaries. In this way we would avoid using the same texts for designing the wordlists and performing the analysis and this would improve the validity of the results. Now, the same texts are used for both. This is of course not very elegant, but being limited by time and lack of available reference documents has led to this procedure.

The main requirements for a word to be added to a wordlist are:
- it should have substantive meaning for a category, i.e. a validity requirement,
- it should be interpreted the same in all texts, i.e. there should be little room for ambiguity about the meaning of the word, and
- it should be discriminating between the parties, i.e. the frequency of a word turning up in a document should be significantly different for the different parties.

For the categories distinguished above, we designed bipolar wordlists. For economic policy and social values bipolar dictionaries are created, content analysis is performed. The same procedure is done for the categories ‘law and order’ and ‘political reform’. Dictionaries for these two domains were constructed, but these categories were later removed from the analysis, because the parties did not substantially distinguish themselves on these domains.

Only for environmental policy the wordlist is uni-polar, since references to environment are made in ‘protection sense’ only. No sane party would actually state that it wants to waste the environment. Even if a party is not willing to spend a lot of money on environmental protection, it will only make positive references. The difference in the extent to which parties choose for environmental protection is measured by comparing the percentage of words dedicated to environmental protection with the total amount of words in the manifesto.

The wordlist for left-wing words for the economic left-right dimension consists of typical left-wing words like ‘care’, ‘state’, ‘insurance’, ‘health’ and ‘social security’. Words on the others side of the left right dimension are ‘growth’, ‘budget-deficit’, ‘individual’, and ‘stimulate’. These words have substantive left or right wing meaning, and the frequency of these words in the different documents differs.

The social value wordlist provides a list of ‘liberal words’ on the one and ‘conservative words’ on the other hand. Parties distinguish themselves on this scale and the words are relevant for each category. On this dimension, liberal words are ‘emancipation’, ‘equality’, ‘self-determination’, ‘freedom’ and ‘ethnicity’, whereas conservative words are ‘value’, ‘traditional’, ‘religion’, and ‘family’.
The dictionary that examines ecological commitment is a very large dictionary. This is the case because environmental words are usually not ambiguous. All words with ‘energy’, ‘soil’ or ‘resources’ or even the term ‘environment’ itself, refer to environmental protection and are only used in a pro-ecological sense. The wordlists for the domains left right, social values and environmental protection can be found in appendix A. These dictionaries will be set down in Dutch. Anglo-Saxon readers will understand the meaning of the scales from the short overview above, and keeping the dictionaries in Dutch gives the opportunity to verify the results and to extend the same analysis with new documents.

The frequencies of words from the wordlists - obtained with the computer program KWALITAN (1998) - occurring in the documents, are portrayed in the following table.

Table 1  Counts of coded words in the party manifestos of 1998

<table>
<thead>
<tr>
<th>Party</th>
<th>economic policy</th>
<th>social values</th>
<th>Environment</th>
<th>Total # words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left</td>
<td>Right</td>
<td>Liberal</td>
<td>Conservative</td>
</tr>
<tr>
<td>GrLi</td>
<td>590</td>
<td>93</td>
<td>109</td>
<td>114</td>
</tr>
<tr>
<td>PvdA</td>
<td>738</td>
<td>186</td>
<td>110</td>
<td>106</td>
</tr>
<tr>
<td>D66</td>
<td>656</td>
<td>279</td>
<td>176</td>
<td>177</td>
</tr>
<tr>
<td>CDA</td>
<td>636</td>
<td>176</td>
<td>153</td>
<td>267</td>
</tr>
<tr>
<td>VVD</td>
<td>212</td>
<td>212</td>
<td>64</td>
<td>69</td>
</tr>
</tbody>
</table>

Before computing the positions of the parties on these scales, a short introduction on the political parties will be given, because this improves the understanding of the parties on the dimensions.

- In 1989 three small left-wing parties, the progressive party, the pacifist socialist party and the communist party, merged into GrLi. GrLi, called Green Left is a left-wing alternative for voters who find the social democrats (PvdA) too moderate. The average support for this party is about 3 or 4% of the electorate, but this party is growing and its electoral support in 1998 was approx. 7%.

- The second party, the PvdA, which has its roots in the labour movement, is the largest left-wing party in the Netherlands. This social democratic party usually gains almost one third of the votes and has been a partner in 8 cabinets since 1946.

- The next party in the table is D66, Democrats ‘66, named after the year of their founding. This is a moderate left-wing liberal party that originated as a party proposing constitutional reform. Its newest denomination is being a ‘social liberal’ party. Its support ranges from 4% to 16% of the votes.

- A Christian Democratic Party, the CDA, was officially founded in 1980 but already entered the elections of 1972 with a list. Three small religious parties merged into the CDA. These parties are the KVP (catholic party, the largest of the three), the ARP (protestant party) and the CHU (also a
The CDA and before its founding the KVP is a center party and was a member in every cabinet until 1994. Its support ranges from approximately 20% to 35% of the vote.

- The VVD is a right-wing secular liberal party and is the main opponent of the PvdA on social economic issues. In general, the PvdA and the VVD were alternating coalition parties for the main Christian Democratic Party. The average turnout of the VVD is 15%.

The next step, is to calculate the party positions on the ideological scales. The position of a manifesto on the left right scale depends on the proportion of left wing statements compared to the right wing hits. For example, the position of Green Left on the policy scale left right is the number of words coded right, minus the number of words coded left, divided by the total number of hits on left and right; 93-590 / 683 = -.73. The score for the D66 is 279-656 / 935 = -.40. The same formula works for the value scale. The position on the ecological scale is computed differently. On this salience issue parties always score positive. The position is therefore the proportion of words dedicated to environmental protection from the total amount of words in a document.

The ranges of these positions differ quite a lot, because for social values and left right the number of words in the dictionaries vary quite a lot. For instance, the list and hits for left wing words is larger than for right wing words which means that all scores are below zero. Also the raw scores for environment are very small since they are related to the total number of words in the document. Therefore, I have decided to transform the scores. In each scale the largest score receives the value one and the lowest score receives zero. The other scores on the scale receive their normalised score, which is the raw score divided by the range of the scores in the scale. This is done in order to make comparing positions on different scales easier. If differences in the size of the dictionaries induce a smaller range of scores on a particular scale than on another, and if at the same time these scales are perceived equally important, this would (mis)lead to the conclusion that on one scale parties are more alike than on another. This would influence the perceived party policy positions and would thus cause problems when testing party competition or coalition formation theories in 1998. The normalised scores of the Dutch manifestos in 1998 for left right, values and environmental protection, can be found in Appendix A. For illustrative purposes, we present a graphical representation of party positions on the scales in the following figures.
5 Summary and Discussion

In this paper a short overview on different databases, and methods to obtain positions on dimensions of political parties is given. Then computer-based content analysis is applied for Dutch political parties in 1998. If, in the future, this method turns out to be worthwhile, this might solve the lack of reliable data on party positions, since it is possible to extend this method in retrospect.
Researchers from the Free University of Amsterdam have also been working on the 1998 party positions. Kleinnijenhuis and De Ridder have performed expert content analysis in media, newspapers and television programmes, in the months before the elections until Election Day in 1994 (EJPR Kleinnijenhuis and De Ridder, 1998) and in 1998 (Kleinnijenhuis, et al. 1998). The emphasis of these studies is on issue ownership of parties, and less on issue positions like in our analysis. Still, it is interesting to examine these results and see whether or not they comply with our results, and v.v.

Some issues in their 1998 research are closely related to our scales are environment, social values (in this study denoted as Christian ethics) and left and right. The placement of the parties in relation to each other on the first two issues - in our research denoted as environmental protection and social values - are fairly similar, but not identical to our results.

On the value dimension positions of parties have even changed place. In both studies the Christian Democratic Party is positioned far away from the other parties on the value scale, and the remaining parties lie closely to each other. However, on Kleinnijenhuis' scale, the most liberal party is the PvdA, whereas on our scale the most liberal party is D66. Of course, these analyses were done with different word-lists, different documents and for different purposes, but since the same parties are analysed the same positions should be found.

It is, however, encouraging that on the economic policy domain in both analyses we find the position of the Christian Democratic Party further to the left on the left-right scale - even at the left-hand side of D66 - than its traditional position. In most studies before 1998, the CDA is positioned to the right of D66, and the fact that both studies show the same deviance from the CDA’s traditional position, increases the confidence in the scaling method. Also, based on merely face validity, the CDA manifesto of 1998 does seem to be more ‘social’ i.e. left wing than the program of D66. Finding this unanticipated shift of the CDA in both analyses, complies with the idea of high reliability of computer based content analysis.

In order to construct decent scales, more research on content analysis and especially on designing word-lists is necessary. Also, a criterion should be designed to validate the results. Laver and Garry made an attempt by studying the correlation between the scales constructed by different methods. Such a thing should be applied and extended, such that a good criterion can be found, to evaluate the data reduction method. If this method proves to be as promising as it seems, the next task is to perform computer-coded content analysis in retrospect.
References


Appendix A

Wordlists used for Content Analysis of the party manifestos in 1998

A.1.1 Economic Policy Domain
Wordlist left

overheid  gekoppeld
sociale  gekoppelde
sociaal  koppelingswet
zorg  arbeid
zorgen  instellingen
werk  betaald
investeren  betaalde
zekerheid  uitkering
basis  uitkeringen
staat  minimumloon
draagvlak  armoede
voorzieningen  armoedevl
investeringen  armoedebestrijding
zekerheid  anti-armoedebestrijding
gezondheidszorg  solidariteit
vervoer  bijstand
koppelen  minimum
koppeling  herverdeling
ontkoppeling  basisvoorzieningen

A.1.2 Economic Policy Domain
Wordlist right

Infrastructuur  staatsschuld
lastenverlichting  tekort
individuele  concurrentiepositie
beperking  concurrentie
groei  stimuleren
particuliere  ontplooien
versterking  ontplooiing
deregulering  flexi…
besparingen  veiligheid
industrie  afdrachten
financieringstekort  orde
vrijheid  marktsector
zelfstandig

15
A.2.1 Social Values Domain
Wordlist conservative

alleenstaanden
allochtoon
allochtonen
alleenstaande
asiel
asiel-
asielaanvragen
asielfunctie
asielverzoek
asielzoekers
asielzoekers
asielzoekerscentra
beginsel
bordeelverbod
burgerlijk
burgerlijke
burgermaatschappij
defensie
defensie-
defensie-identiteit
defensie-industrie
defensie-uitgaven
defensiec Capaciteit
defensiepersoneel
defensieuitgaven
door migratie
echt genoot
evangelie
familie
familieband
familiebedrijven
familiebeleid
familieleden
familieverband
familievriendelijke
geruwd
gezinde
genof
ge loof
ge loofwaardig
ge loofwaardige
ge loofwaardigheid
gezin
gezinnen
gezins-
gezinsbeleid
gezinslid
gezinsplanning
gezinsverantwoordel
gezinsvoogden
gezinsvoogdij-inste
gezinsvorming
gezinsvriendelijk
gezinsvriendelijke
gezinsvriendelijkhe
gezin_
godsdienst-
godsdiensten
huwelijk
huwelijks
immigranten
immigrantenkoordenen
immigratie
immigratie-
inburger-
inburgering
inburgeringsbeleid
inburgeringscontract
inburgeringscursus
inburgeringsprogram
inburgeringstraject
kerk
kerkdiensten
kerkelijk
kerkelijke
kerken
koningin
koninklijk
koninklijke
koninkrijk
koninkrijksverband
koninkrijksverhoudi
levenbeschouwelijke
levenbeschouwelijk
levenbeschouwing
migranten
migrantenbeleid
migrantenkringen
migrantenvraagstuk
migratie
migratie-
migratie-effecten
migratieoverschot
migratiestromen
migratievraagstuk
monarchie
naturalisatie-
naturaliseren
norm
normatieve
normen
normvervanging
opvoeden
opvoeding
opvoedings-
opvoedingsaspect
opvoedingsklimaat
opvoedingsondersteu
ouder
ouderbetrokkenheid
ouders
religie
religieuze
remigratie
respect
respecteert
respecteren
re_nTEGRATIE
re_nTEGRATIEbeleid
toelatingsbeleid
toelatingsregeling
traditie
tradities
traditioneel
traditionele
verantwoorde
verantwoordeli
verantwoordelijk
verantwoordelijkheid
verantwoordelijkhe
verantwoording
verantwoordingen
verboden
waarden
waardengemeenschap
zingeving
zingevingsvragen
medisch-ethische
kloren
genetherapie
genetische
euthanasie
euthanasie/hulp
zelfdoding
A.2.2 Social Values

Domain

Wordlist liberal

alleengaande
alleengaanden
anti-discriminatieb
anti-discriminatiew
anti-racisme
antidiscriminatiebe
antidiscriminatieco
antifoltercommissie
bewustwordingsproces
bewustzijn
civil
discriminatie
discrimineren
discriminerend
draagmoederschap
eiceldonatie
emancipatie
emancipatiebeleid
emanciperen
etniciteit
etnisch
etnische
gelijk
gelijkberechtiging
gelijke
gelijkgestemde
gelijkgestemden
gelijkheid
gelijkheidsbeginsel
gelijkwaardig
gelijkwaardige
gelijkwaardigheid
gerechtigd
gerechtigden
gerechtigheid
gerechtvaardigd
gerechtvaardigde
gescheiden
gezinshereniging
grondrecht
grondrechten
homoseksuele
humanitaire
keuzevrijheid
leefvormen
lesbische
levensbe_indiging
levensvormen
liberale
liberalen
liberalisatie
liberalisering
medeburger
medeburgers
menningsuiting
menningsuitingen
mens
menselijke
mensen
mensenhandel
mensenrechte
mensenrechten-
mensenrechtenaspect
mensenrechtenbeleid
mensenrechtencommis
mensenrechtersgroep
mensenrechtenmorum
mensenrechtenorgani
mensenrechtenpolitii
mensenrechtschend
mensenrechtsituat
mensenrechtenstaanda
mensenrechtenverdra
mens_
minderheden
minderhedenbeleid
minderheid
minderheidsgrp
minderheidsgroep
minderheidsgroperi
minderheidsrechten
multiculturaliteit
multicultureel
multiculturele
non-conformisme
onderdrukking
partner
partnerregistratie
partners
partners/samenwonen
partnersvoogdijschap
prostitutie
racisme
racistisch
ralpatronen
samenlevingen
samenlevingsvorm
samenlevingsvormen
samenwonen
samenwonenden
seksueel
seksuele
sexueel
subcultuur
tolerantie
tolereert
verdraagzaamheid
vluchtelings
vluchtelingen-
vredvredesproces
vreie
vrije
vrijheid
vrijheidszin
zelfbeschikking
<table>
<thead>
<tr>
<th>Dutch Term</th>
<th>English Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>aardbol</td>
<td>ecotax</td>
</tr>
<tr>
<td>aarde</td>
<td>ecotoerisme</td>
</tr>
<tr>
<td>aardgas</td>
<td>emissie</td>
</tr>
<tr>
<td>aardgasbaten</td>
<td>emissiearme</td>
</tr>
<tr>
<td>aardgasvoorraden</td>
<td>emissieplafonds</td>
</tr>
<tr>
<td>aardolie</td>
<td>emissierechten</td>
</tr>
<tr>
<td>aardwarmte</td>
<td>emissiereducties</td>
</tr>
<tr>
<td>afval</td>
<td>emissies</td>
</tr>
<tr>
<td>afval...</td>
<td>energie</td>
</tr>
<tr>
<td>ammoniak</td>
<td>energie-</td>
</tr>
<tr>
<td>anti-verdrogingsbel</td>
<td>energie-...</td>
</tr>
<tr>
<td>atoomenergie</td>
<td>erodeert</td>
</tr>
<tr>
<td>bedrijfsmilieuplan</td>
<td>fiets</td>
</tr>
<tr>
<td>bestrijdingsmiddele</td>
<td>fiets-...</td>
</tr>
<tr>
<td>bio-industrie</td>
<td>gasprijs</td>
</tr>
<tr>
<td>bio...</td>
<td>geluidhinder</td>
</tr>
<tr>
<td>bodem-</td>
<td>geluidshinder</td>
</tr>
<tr>
<td>bodembescherming</td>
<td>geluidsnormen</td>
</tr>
<tr>
<td>bodems</td>
<td>groen</td>
</tr>
<tr>
<td>bodemsanering</td>
<td>groene</td>
</tr>
<tr>
<td>bodemvervuiling</td>
<td>groenfondsen</td>
</tr>
<tr>
<td>brandstof</td>
<td>groenstructuren</td>
</tr>
<tr>
<td>brandstofdrager</td>
<td>grondbeleid</td>
</tr>
<tr>
<td>brandstoffen</td>
<td>grond...</td>
</tr>
<tr>
<td>broeikaseffect</td>
<td>herbebossing</td>
</tr>
<tr>
<td>broeikasgassen</td>
<td>herbruikbaar</td>
</tr>
<tr>
<td>broeikassen</td>
<td>herbruikbare</td>
</tr>
<tr>
<td>chemicali_n</td>
<td>hergebruik</td>
</tr>
<tr>
<td>chemische</td>
<td>hergebruiken</td>
</tr>
<tr>
<td>co2</td>
<td>hergebruiksector</td>
</tr>
<tr>
<td>congestie</td>
<td>kernafral</td>
</tr>
<tr>
<td>congestieproblemen</td>
<td>kerncentrale</td>
</tr>
<tr>
<td>congestiegebieden</td>
<td>kerncentrales</td>
</tr>
<tr>
<td>delfstoffen</td>
<td>kernenergie</td>
</tr>
<tr>
<td>dienstfietsen</td>
<td>kernkwaliteit</td>
</tr>
<tr>
<td>diesel</td>
<td>kerosine</td>
</tr>
<tr>
<td>dodewaard</td>
<td>klimaat</td>
</tr>
<tr>
<td>drinkwater</td>
<td>klimaatsysteem</td>
</tr>
<tr>
<td>drinkwater-voorzien</td>
<td>klimaatverandering</td>
</tr>
<tr>
<td>drinkwaterbedrijven</td>
<td>kooldioxide</td>
</tr>
<tr>
<td>drinkwatervoorzieni</td>
<td>krachtscentrale</td>
</tr>
<tr>
<td>dumpen</td>
<td>kringloopcentra</td>
</tr>
<tr>
<td>duurzaam</td>
<td>kringloopfuncties</td>
</tr>
<tr>
<td>duurzaam...</td>
<td>kringlopen</td>
</tr>
<tr>
<td>eco-tax</td>
<td>kunstmest</td>
</tr>
<tr>
<td>ecologie</td>
<td>kustgebied</td>
</tr>
<tr>
<td>ecologisch</td>
<td>kuststroken</td>
</tr>
<tr>
<td>ecologische</td>
<td>kustuitbreiding</td>
</tr>
<tr>
<td>ecologiseren</td>
<td>kustvaart</td>
</tr>
<tr>
<td>ecopolis</td>
<td>leidingwater-afvalw</td>
</tr>
<tr>
<td>ecotax</td>
<td>luchtvervuiling</td>
</tr>
<tr>
<td>ecotoerisme</td>
<td>mest</td>
</tr>
<tr>
<td>emissie</td>
<td>milieu</td>
</tr>
<tr>
<td>emissiearme</td>
<td>milieu</td>
</tr>
<tr>
<td>emissieplafonds</td>
<td>natuur</td>
</tr>
<tr>
<td>emissierechten</td>
<td>natuur ...</td>
</tr>
<tr>
<td>emissiereducties</td>
<td>nitraat-richtlijn</td>
</tr>
<tr>
<td>emissies</td>
<td>nucleaire</td>
</tr>
<tr>
<td>energie</td>
<td>nul-emissie-situati</td>
</tr>
<tr>
<td>energie-</td>
<td>oppervlaktewater</td>
</tr>
<tr>
<td>energie-...</td>
<td>oppervlaktewateren</td>
</tr>
<tr>
<td>erodeert</td>
<td>radioactief</td>
</tr>
<tr>
<td>fiets</td>
<td>recycling</td>
</tr>
<tr>
<td>fiets-...</td>
<td>regenwouden</td>
</tr>
<tr>
<td>gasprijs</td>
<td>rijwielpaden</td>
</tr>
<tr>
<td>geluidhinder</td>
<td>ruimtebeheer</td>
</tr>
<tr>
<td>geluidshinder</td>
<td>ruimteclaims</td>
</tr>
<tr>
<td>geluidsnormen</td>
<td>ruimtetoeveling</td>
</tr>
<tr>
<td>groen</td>
<td>schone</td>
</tr>
<tr>
<td>groene</td>
<td>schonere</td>
</tr>
<tr>
<td>groenfondsen</td>
<td>schoon</td>
</tr>
<tr>
<td>groenstructuren</td>
<td>sputitbussen</td>
</tr>
<tr>
<td>grondbeleid</td>
<td>statiegeldsysteem</td>
</tr>
<tr>
<td>grond...</td>
<td>statiegeldsystemen</td>
</tr>
<tr>
<td>herbebossing</td>
<td>temperatuurstijging</td>
</tr>
<tr>
<td>herbruikbaar</td>
<td>uitlaatgassen</td>
</tr>
<tr>
<td>herbruikbare</td>
<td>uitstoot</td>
</tr>
<tr>
<td>hergebruik</td>
<td>vervuilde</td>
</tr>
<tr>
<td>hergebruiken</td>
<td>vervuildeende</td>
</tr>
<tr>
<td>hergebruiksector</td>
<td>vervuiler</td>
</tr>
<tr>
<td>hergebruikt</td>
<td>vervui...</td>
</tr>
<tr>
<td>kernafral</td>
<td>verzurende</td>
</tr>
<tr>
<td>kerncentrales</td>
<td>verzuring</td>
</tr>
<tr>
<td>kernenergie</td>
<td>warmte-krachtkoppel</td>
</tr>
<tr>
<td>kernkwaliteit</td>
<td>warmtemkracht</td>
</tr>
<tr>
<td>kerosine</td>
<td>warmtekrachtkoppeli</td>
</tr>
<tr>
<td>klimaat</td>
<td>water</td>
</tr>
<tr>
<td>klimaatsysteem</td>
<td>worldklimaatconfer</td>
</tr>
<tr>
<td>klimaatverandering</td>
<td>wereldmilieufonds</td>
</tr>
<tr>
<td>kooldioxide</td>
<td>windenergie</td>
</tr>
<tr>
<td>krachtscentrale</td>
<td>windenergieplan</td>
</tr>
<tr>
<td>kringloopcentra</td>
<td>windmolens</td>
</tr>
<tr>
<td>kringloopfuncties</td>
<td>windpark</td>
</tr>
<tr>
<td>kringlopen</td>
<td>zee</td>
</tr>
<tr>
<td>kunstmest</td>
<td>zeehaven</td>
</tr>
<tr>
<td>kustgebied</td>
<td>zeemilieu</td>
</tr>
<tr>
<td>kuststroken</td>
<td>zoetwatervoorzienin</td>
</tr>
<tr>
<td>kustuitbreiding</td>
<td>zonne</td>
</tr>
<tr>
<td>kustvaart</td>
<td>zuiderzee-hsl</td>
</tr>
<tr>
<td>leidingwater-afvalw</td>
<td>zwavelverbindingen</td>
</tr>
</tbody>
</table>