

**Comparing
non-ipsative measures
of party support**

Cees van der Eijk (University of Nottingham, UK)
and
Michael Marsh (Trinity College, Dublin, Ireland)

Paper presented at First European Conference on Comparative Electoral Research,
University of National and World Economy, Sofia, 1-3 December 2011

Comparing non-ipsative measures of party support

Cees van der Eijk and Michael Marsh

Abstract

Non-ipsative measures of party preference (preference ratings for each of the parties of a political system) have become common in election studies. They exist in different forms, such as thermometer ratings or feeling scores, likes and dislikes scores, or choice propensities. Usually only one of these is included in a single survey, which makes it difficult to assess the relative merits of each. The Irish National Election Study 2002 (INES2002) contained three different batteries of non-ipsative party preferences, a feature which allows a systematic comparison between them. This paper investigates some properties of these different indicators. We focus mainly on the relationship between non-ipsative preferences and actual choices. This relationship is particularly revealing in a STV electoral system that allows voters to cast multiple ordered votes for candidates from different parties. Additionally, we investigate the latent structure of each of the batteries of party preferences and the relationships between them.

We conclude that the three instruments are not interchangeable, that they measure different kinds of preferences. If the purpose is to study electoral choice and the process leading up to electoral choice, then the propensity to vote for a party is to be preferred over thermometer or feeling scores, and over likes/dislikes scores.

Authors:

Cees van der Eijk, Professor of Social Science Research Methods
and Director of Social Sciences Methods and Data Institute
University of Nottingham, Nottingham NG7 2RD, UK
Email: cees.vandereijk@nottingham.ac.uk

Michael Marsh, Professor of Comparative Political Behaviour
Dean of Arts, Humanities and Social Sciences, Trinity College Dublin
http://www.tcd.ie/Political_Science/Staff/Michael.Marsh
Email: mmarsh@tcd.ie

Introduction and theoretical framework

Election studies often include some kind of question about respondents' preferences for each of the various parties in a country. Such questions exist in different forms, of which we review here the three most commonly used. When considering how they are used in the extant literature we find two main interpretations of what the responses to these questions represent conceptually: affect and electoral utility.¹ This raises the question whether each of the different forms in which these questions are asked is equally valid for each of these interpretations, and what the consequences are of any lack of (construct) validity. This is the main focus of this paper.

The kind of survey questions we investigate are non-ipsative preferences for political parties.² They ask respondents to indicate for each of the political parties separately how strongly they like or dislike them, or how warm or cold their feelings for those parties are, or how likely it is that they will ever vote for them. One way of interpreting the responses is that they measure *affect*. Affect is a concept that refers to the emotive associations generated by exposure to the stimulus in question. These associations may be positive, neutral or negative, and measures of affect purport to provide information about their strength and direction. In a more specific form, the concept of affect is ubiquitous in electoral studies in the form of party identification (and particularly its so-called 'tribal' aspects). Affect ratings from non-ipsative preference questions differ from party identification because the latter is generally seen as exclusive (and hence ipsative), but are similar in terms of evoking emotive associations which themselves may have consequences in terms of inducing perceptual screens³ or behavioural predispositions.

The second way in which non-ipsative preferences are interpreted is as indicators of *electoral utility*. This concept was introduced by Downs (1957) in a theoretical form, referring to the expected benefits a voter would derive from a political party winning the election. Formulated differently it refers to the attractiveness of *voting* for a party in a particular election. Utility is, much more than affect, contextually anchored, dependent on the characteristics of the election at hand and of each of the parties at that time. It is thus not necessarily stable over time.

Clearly, affect for and electoral utility of parties are likely to be correlated. Affect is likely to be one of the factors that contributes to utility, but it is not the only factor that matters in this respect. Other possible determinants of utility include pocket-book

¹ Much of the relevant literature is not explicit in this regard, and frequently data from non-ipsative preference measures are used without conceptual reflection. The actual use of such data, and the interpretation given to analytical findings does, however, reflect these two different streams of interpretation.

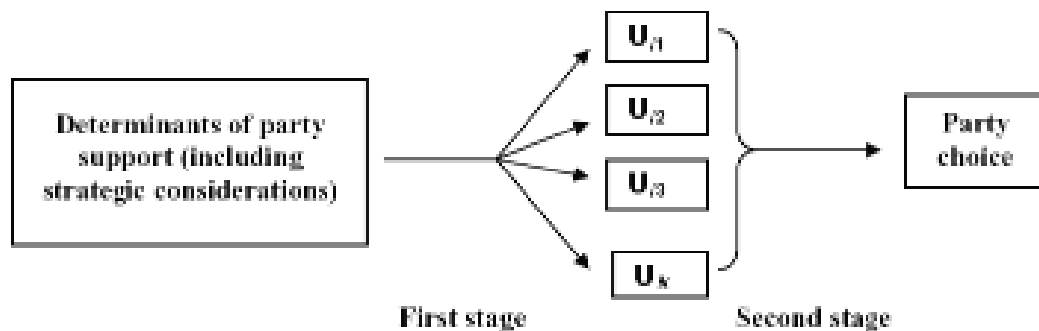
² Preferences are called ipsative if they are constrained by a fixed-sum rule. In most countries, electoral choice is ipsative, as the voter is only allowed to express a single preference on the ballot. Voting for one party thus implies not-voting for each of the other ones. Ranked preferences are ipsative in a different way: having indicated which of a series of stimuli is preferred most implies that others are ranked lower, etc. In all these cases, the preferences for the various choice options are thus not mutually independent. In the case of non-ipsative preferences, a respondent expresses a degree of preference for each of the options, where the response given for any option does not in any way constrain the responses for other options. Most often, non-ipsative preferences are elicited in the form of preference ratings for the various options.

³ Generally these refer to processes of selectivity when confronted with communications and new information: selective exposure, interpretation, acceptance, and retention.

considerations; similarity between voter and party qua ideology, policy positions, or cleavage location; perceptions of competency of parties and their leaders; and the perceived capacity of parties to affect the course public policy. On the other hand, utility considerations may well spill over into affect, particularly if they happen to be stable between successive elections, resulting in repeated reinforcement. Yet, utility may be generated by non-affective factors, and affect does not have to originate in utility considerations. Affect and utility are therefore conceptually distinct, yet to what extent they can be distinguished in practice can only be assessed empirically.

For assessing construct validity of each of the non-ipsative preference questions as measures of affect or utility we need, ideally, not only ‘face validity’ considerations, but also empirical criteria that can be derived from the observable implications of the respective concepts. Because actual voting is generally ipsative, such a criterion can be formulated, at least for the utility concept. The distinction between utility (which is non-ipsative in character) and voting (which is ipsative) implies a two-stage model of electoral choice, as illustrated in Figure 1. The first stage involves the development of utilities; the second stage consists of the determination of choice by some procedure, often conceived of as utility-maximization.

Figure 1: The two-stage model of electoral choice



The distinction between utilities and choice, which implies a two-stage model in the analysis of electoral politics has a large number of analytical advantages that were perceived by Philip Converse in 1968 (see Converse 1974:742-3). Bingham Powell (2000:160) likewise argues that information about choice (which party someone voted for) does provide us with sufficient information to understand (or explain) voters’ electoral behaviour. Indeed, neither proximity models nor directional models of voter choice can be tested explicitly when only the outcome of the process (the choice made) is known (cf. Rabinowitz and McDonald 1989; MacDonald, Listhaug and Rabinowitz 1991). Van der Eijk et al. (2006) have elaborated additional analytical advantages that can be obtained from the availability of empirical information about utilities in combination with choice.⁴ For this paper it suffices that this conceptual foundation provides an empirical criterion for construct validity: validity of non-ipsative preferences as measures of utility requires that highest preference and actual choice coincide. We will

⁴ See also van der Eijk and Franklin (1996), and van der Brug et al. (2007a).

assess the performance of different non-ipsative measures in this respect, but before doing so, we first introduce the data.

Data, measures and context

The source of our data is the Irish National Election Study (INES) of 2002 (see Marsh et al. 2008).⁵ This dataset contains three different non-ipsative preference measures, thus allowing rare possibilities for a comparative assessment of their characteristics and validity. The INES is unusual in having included all three different forms of these questions in the same study. Most election studies select only one of them for inclusion, which makes it virtually impossible to make a comparative assessment of their validity as measures of affect or utility.⁶ The Irish case is of particular interest, as its STV electoral system offers more, and more demanding elaborations of our validation criterion than most other electoral systems.

Three measures of non-ipsative party preferences

The questionnaire of the INES contains three non-ipsative preference questions. One of these is the ‘classic’ feeling thermometer that derives from the American National Election Studies. The question is used to rate not only parties, but also national politicians and candidates in the voter’s constituency:

I'd like to ask you how you feel about some Irish politicians, using what we call the 'feeling thermometer'. The feeling thermometer works like this: If you have a *favourable feeling* (a warm feeling) towards a POLITICIAN you should place him/her somewhere between 50 and 100 degrees; If you have a *unfavourable feeling* (a warm feeling) towards a POLITICIAN you should place him/her somewhere between 0 and 50 degrees; if you have a *don't feel particularly warm or cold* (have no feeling towards the politician at all) then you should place him/her somewhere at 50 degrees; where would you place these Irish politicians?

And where would you place each of the following PARTIES;
and where would you place these CANDIDATES who ran in your constituency in the general election in May?

The second non-ipsative preference question is also asked for politicians (i.e., the national leaders of the major parties) and parties. It follows the formulation and format used in the CSES.⁷ For the political parties this question reads as follows:

⁵ The study was funded under the Programme for Research in Third Level Institutions (PRTLTI). This project was co-directed by Michael Marsh and Richard Sinnott. The post-election survey was conducted by the Economic and Social Research Institute (ESRI) using face-to-face interviews. For details of the methodology see Marsh et al 2008; but see also Laver 2005; Marsh 2004, 2006b, 2007.

⁶ The Dutch national election studies has occasionally included the thermometer scale in addition to the propensity to support question, thus also allowing some comparisons of the characteristics of the respective items. The Irish National Election Studies in 2007 and 2011 also included two non-ipsative preference questions. We refer to some findings from these studies in our concluding section.

⁷ The CSES –Comparative Study of Electoral Systems– is an invaluable and ambitious attempt to increase comparability of elections by insertion of a standardized block of questions in the election studies of all

I'd like to know what you think about each of our political parties. Please rate each party on a scale from 0 to 10, where 0 means you STRONGLY DISLIKE that party and 10 means that you STRONGLY LIKE that party. If you come to a party you haven't heard of or you feel you do not know enough about, just put a tick in the 'Don't know' box on the right.

The third question probing non-ipsative preferences for parties was originally developed by van der Eijk and Niemöller in the early 1980s for the Dutch national election studies (van der Eijk and Niemöller 1984; DPES 1982). It has been used in all Dutch national election studies since 1982, in all European Election Studies since 1989, and in a growing number of election studies in European countries, including Britain, Germany, and Spain. This formulation of this so-called 'propensity to vote' (PTV) question is:

We have a number of political parties in Ireland each of which would like to get your vote. How probable is it that you will ever give your first preference vote to the following parties? Please use the numbers on this scale to indicate your views, where '1' means 'not at all probable' and '10' means 'very probable'.

Each of these questions was asked for the 6 main Irish parties: Fianna Fáil (FF), Fine Gael (FG), Greens (Gr), Labour (Lab), Progressive Democrats (PDs) and Sinn Féin (SF). In addition, the propensity to vote question was also asked for 'an Independent candidate'. The three sets of questions were separated by significant intervals so as to reduce the risk of response set. Nevertheless, it is not unlikely that respondents are increasingly primed into a party evaluation mind frame as the interview progresses, which should make it easier for them to respond (and respond consistently) to later party preference questions. It is therefore relevant to know that the PTVs were located earliest in the main questionnaire, and the thermometer questions later.⁸ The likes/dislikes question came much later, in the drop-off section of the questionnaire.⁹

Institutional context: voting under STV

Ireland uses the single transferable vote system. The ballot contains the names, photographs and party affiliations of the candidates in the constituency. Voting consists of indicating a preference ranking for as many candidates as one wishes, and voters know that the 'transferable' aspect minimises the chances of a vote being 'wasted'. Parties may field more than one candidate in a given constituency, while at the same time not always fielding candidates in all constituencies. In 2002, for example, the PDs did not have candidates in 25 of the 42 constituencies. The particular combination of choice options available to voters may thus differ from constituency to constituency. Just under one third of all voters could choose between candidates of all parties and a further third again were

participating countries. For more details about the project and the data that it has generated, see <http://www.cses.org>.

⁸ The questionnaire was a split one. For a random half of the sample, the intervening questions were open ended ones about what each party stood for; the other half were asked to place each party's position on both a scale of attitudes to the EU and a scale on environmental protection.

⁹ In practice many drop-off sections were completed as part of the main interview, contributing to a success rate of almost 90 percent completion of this section.

faced with all parties except the PD. The remaining constituencies were more varied, with SF fielding candidates in only 34/42 constituencies, Greens in 30/42, Labour in 39/42. Independent candidates competed for votes in all but three constituencies, and not without success.¹⁰

The variety of party availability, and the variety of ways in which voters may use their preferences has to be kept in mind when analyzing and interpreting our survey data. In contrast to election surveys from many other countries we cannot look simply at party choice, but have to distinguish 1st preference votes from 2nd preference votes and so on. The question whether the vote is a sincere expression of preference is of little relevance as the most common reason for casting an insincere preference in other electoral systems – the fear of ‘wasting’ a vote – is irrelevant in STV, as unused votes are transferred to candidates who acquired lower preference votes. Arguably the need for coalitions could undermine sincere preferences. However, it could be argued that the uncertainty about post-coalition options is normally so great as to make it very difficult for any voter to know how to use lower preferences to maximise the chance of her preferred party being in government (see Laver 2000 and Marsh 2010 for discussions).

Finally, it is a somewhat open question whether Irish voters vote for parties (who field candidates), or for candidates (most of whom happen to be associated with a party). It is widely assumed that candidate factors are of great importance, yet exactly how important candidate and party factors are vis-à-vis each other is a matter of debate (see, e.g., Marsh 2007; Marsh et al 2008 Chapter 8). However, to a considerable degree, candidate and party loyalties seem to coexist: if a voter likes a candidate, she likes that party and vice versa. The debate is essentially about how often the party is the chicken and how often it is the egg in this process.

Non-ipsative preferences and choice

As discussed earlier, for non-ipsative preferences to be valid as indicators of electoral utility the highest preference score should be given to the party actually chosen. This provides a straightforward and observable criterion for construct validity. This logic is most compelling for voters who see the election in terms of a choice between parties, rather than as a choice between candidates. In our analyses we distinguish therefore between candidate-oriented voters, party oriented ones, and a group with mixed orientations.¹¹ We expect that the party with the highest preference score will more often diverge from the one actually chosen for candidate-oriented voters than for party-oriented voters.

¹⁰ In 2002 the number of Independent candidates was 95 nationwide, who amassed 9.5% of first preference votes nationwide, and 13 of whom were elected (total size of the Dáil is 166 representatives).

¹¹ This typology was developed by Marsh et al (2007, Chapter 8) and is based on respondents indicating whether party or candidate was the most important factor determining their choice, and whether they would still vote for the same candidate if (s)he were to switch to a different party. Self-attributions of motivations are often a disputable basis for such distinctions, (see, e.g., Kaplan 1964 on the distinction between act meaning and action meaning), but to the extent that voters differ in this respect, it is plausible that these self-attributions will reflect this at least to some extent.

Table 2 reports the concurrence between the highest score on the non-ipsative preference measures and first preference vote on the ballot.¹² In order not to bias our assessment of the thermometer instrument compared to the other two non-ipsative instruments, we used the thermometer scale both in its original form (scores ranging from 0 to 100), and in a recoded 10-category form.

Table 2: Percentages of highest-preference voting for different non-ipsative preference measures and different groups of voters
Data: INES2002; valid N: 1466-1643

	PTV	Thermo- meter	Therm. recoded	Like / dislike
Candidate-oriented voters	82.0	59.6	60.7	73.4
Voters with mixed orientations	87.5	66.3	67.4	86.1
Party-oriented voters	94.5	70.5	70.7	92.1
TOTAL	87.1	64.9	65.7	82.9

Table 2 shows large differences between the three questions. For the entire sample as well as for each of the three (roughly equally sized) subgroups, we find that PTVs perform in this respect much better than the other two instruments, while the like/dislike question clearly outperforms the thermometers. It should be kept in mind that the criterion of highest preference voting becomes more demanding as the number of options increases. In the case of the PTVs this criterion is based on 7 options (6 parties plus Independents), whereas only 6 options underlie this test for the other two instruments. The incidence of highest preference voting is –as expected– highest among party-oriented voters. Yet, the PTV measure, in particular, shows that even candidate-oriented voters behave mostly (in excess of 80%) in a party-driven fashion: voting for the *party* that has their highest preference. The question remains, of course, whether party preferences are driven by candidate preferences, or the other way around, or whether party and candidate preferences affect each other reciprocally. Yet, such questions are outside the remit of this paper.

In the analyses reported in Table 2 we disregarded the differences between constituencies (and thus between respondents) in terms of the slate of parties available. We checked in a number of ways whether taking such differences into account alters the results. We do not report these results as they are exceedingly similar to those reported in Table 2.¹³

The results so far point unequivocally to the PTV as the most valid of the three questions as indicator of electoral utility. This assessment is based, however, on first preference votes only. We will turn to lower preference votes next.

¹² The N's for the different cells vary somewhat as a consequence of missing data.

¹³ Details are reported by Van der Eijk and Marsh (2007) and are available from the authors on request.

Non-ipsative preferences and lower preference votes

Comparing lower preference votes with party preferences only makes sense when the lower preference vote involves a different party than the first preference vote. This restriction leads to the loss of some cases: those respondents who did not give any of their lower preference votes to a different party, or if they did, voted for one of the minor ‘other’ parties for which we have no preference ratings.¹⁴

For simplicity or presentation, we report only findings for respondents who could choose from all major national parties.¹⁵ We assess whether or not the next party chosen is the one that was given the next highest rating. In view of this we have to distinguish a number of possible relationships between preferences and choice, as displayed in Table 3.

Table 3: Relationship between non-ipsative preferences and lower preference vote (percentages)

Data: INES2002

	PTV	Thermometer	Like/dislike
Category 1: 2nd party best	5.5	3.2	6.1
Category 2: 2nd party tied for best with first	28.5	10.4	13.1
Category 3: 2nd party next best after 1st	33.0	28.0	36.9
Category 4: 2nd party better than first but not best	0.00	9.3	6.6
Category 5: 2nd party not best after 1st	33.0	49.1	37.3
Total	100.0	100.0	100.0
Sum of categories 2 and 3	61.5	38.4	50.0
Valid N	397	279	244

Note. The large differences in the N for each column are partly attributable to the fact that the PTVs cover 7 choice options, and the other two measures only 6. They also reflect differences in non-response.

Category 1 in Table 3 consists of respondents whose second preference vote is for the party to which they gave the highest preference score, while the party of their first preference vote was given the second highest preference rating. STV allows voters to vote insincerely – although provides no incentive to do so – by giving their first preference vote to a party that is not the most preferred, and a subsequent preference vote to the party that is most preferred. The data do not allow us to assess whether this motivation underlies the reversal of parties in the voting order compared to the ordering of the preference scores. Therefore we have to consider this category as equivocal in

¹⁴ Only 1 per cent of or sample supported such a party as their second (party) choice and less than 1 per cent did so as a first choice.

¹⁵ Results are very similar when distinguishing between voters on the basis of the parties available to them.

terms of our assessment of the construct validity of the three non-ipsative preference measures.

Categories 2 and 3 unambiguously indicate the extent to which the non-ipsative preference measures are valid indicators of electoral preference. Category 2 describes respondents whose highest preference score was tied for two parties, and whose first and second preference votes were cast for those tied parties. Category 3 represents the straightforward situation where the order of preference votes follows that of preference ratings: the second party chosen is the second highest in terms of preference scores.

Categories 4 and 5 are equivocal, for the same reasons as category 1. Category 4 applies where neither the first preference vote nor the second preference vote was scored highest on the non-ipsative preferences. These respondents may have given a yet lower preference vote to a third party that they scored highest on the preference measures. Category 5, finally, reflects that the second preference vote is not for the party that scored second-highest in preference, while the first preference vote was for the party that did score highest. For neither of these two categories can we assess unambiguously whether they reflect strategic use of the possibilities that STV offers to voters, or that they reflect lack of construct validity.

Only categories 2 and 3 unambiguously reflect patterns of responses that should be observed for the non-ipsative measures to have construct validity. PTVs perform considerably better in this respect than either of the other two measures. The results for second preference votes in conjunction with those for first preference votes (in the previous section) imply that the concept of electoral utility is more validly indicated (by a considerable margin) by PTVs than by likes/dislikes or thermometer scores. The reasons for these differences can be of different kinds, each leading to its own implications for the use of non-ipsative measures in electoral research. In the next section we probe two possibilities: differences in reliability, and differences in latent meaning.

Latent structure analysis of non-ipsative preferences

The differences between the three non-ipsative preferences reported in the previous section may be caused by differences in the phenomena tapped by each of them. But they may also originate from differences in reliability. The latter would imply that the concurrence between preferences and votes is lower because of random error, the former that each of the three questions reflects a (somewhat) different kind of preference. If the three measures differ only in terms of reliability, it should be possible to model them jointly as indicators of a latent variable. Latent variable models are widely applied to, e.g., attitude items. Modelling our current measures has to take a slightly different form, as we are here interested in the homogeneity of instruments that are not single items, but sets of multiple items. We do this by first analysing each of the instruments separately so as to provide a baseline that can then be compared with models in which information from several of these instruments is analysed jointly.

For each of the instruments the responses to the items (i.e., the parties) indicate a similarity between respondent and party in a latent space (the characteristics of which are as yet unknown). The further away a respondents 'ideal party' is from the actual party in question, the lower the preference score. These distances can originate because a party

offers too little, or too much of what is desired (e.g., it may be too left, or too right; too idealistic, or too pragmatic, etc.). In the jargon of measurement models we have ‘non-monotonous single stimulus data’ cf. Coombs (1964), which cannot be examined validly by factor analytic methods (cf. van Schuur 1993; van Schuur and Kiers 1994). Suitable models are available in the tradition of IRT (Item Response Theory); from this family we use a nonparametric unfolding model, MUDFOLD (Van Schuur 19xx). As all unfolding models, MUDFOLD attempts to construct a single representation of persons and parties in a (joint) latent space, so that the distances between persons and stimuli reflect as well as possible (inversely) the empirically observed preferences of each of the persons. Such representations –if they exist– have two important substantive implications. First, they demonstrate that respondents use largely the same mix of criteria when expressing preferences for the parties. And secondly, they demonstrate that respondents’ perceptions of party locations in the latent space are very similar.

How well the best possible MUDFOLD representation (or scale) fits observed data is expressed in the coefficient of homogeneity – H – and this can be compared across instruments and subgroups of respondents. H attains an upper limit of 1 if the constructed scale represents the data perfectly, without any violations. In that case all items reflect the same latent phenomenon. It has the value 0 when the items do not relate at all to a common latent variable, and thus do not relate to each other either. Generally, the lower boundary for acceptable fit (or scalability) is set at $H=0.30$ for survey data. H -values surpassing 0.40 indicate a medium-strong scale, and above 0.50 a strong scale. In addition to the H coefficient for a scale in its entirety, item-specific coefficients – H_i – indicate how well the items fit in a given scale.

Unfolding PTV ratings of parties

Unfolding the 7 PTV items yields a medium strong scale, as reported in Table 4.¹⁶ The table reports the order of the parties in the underlying latent dimension, and the degree of scalability of the set of items as a whole ($H=0.40$) and for each of the parties as part of this scale (H_i).

The order of the parties that represents the observed preferences best is:¹⁷

FF / PDs / FG / Independent / Labour / Greens / SF

*Table 4: Unfolding scale of PTV ratings of Irish political parties and Independents
Data: INES2002*

	H_i (parties)
	H (scale)
FF	0.36
PDs	0.42
FG	0.34

¹⁶To avoid comparability problems owing to different rates of nonresponse, we conducted all unfolding analyses on the group of respondents who gave valid responses to all items of all three non-ipsative party preference questions (N=1801).

¹⁷The same ordering was also found in PTVs in different surveys; for details see Marsh 2006a.

Independent	0.33
Green	0.42
Labour	0.41
SF	0.37
Scale	0.38

The unfolding model implies that voters who express a high preference for a particular party tend to also have relatively high preferences for adjacent parties, and increasingly lower preferences for parties that are further removed. The fit of this unfolding model is quite acceptable. Further inspection indicates that preferences for Independents generate more violations than preferences for the other parties. This is not surprising in view of the heterogeneity of independent candidates. Replicating the analysis for only the 6 national parties –leaving out Independents– yields a scale with the same ordering, and somewhat better fit, as reported in Table 5.

Table 5: Unfolding scale of PTV ratings of Irish political parties, without ‘Independents’
Data: INES2002

	H_i (parties)
	H (scale)
FF	0.39
PDs	0.46
FG	0.39
Green	0.44
Labour	0.45
SF	0.37
Scale	0.42

All parties fit reasonably well in the scale, as indicated by their H_i coefficients. SF is the weakest item. The differences in H_i are not of a magnitude, however, that PTV scores for SF have to be interpreted in a different light than those for other parties.

Unfolding of Thermometer and Like/Dislike ratings of parties

Similar analyses were done for the thermometer and like/dislike ratings. The results, reported in Table 6, are remarkably similar to those reported for PTVs (see Table 5). The ordering of parties on the unfolding scale is the same, and both the thermometer and the like/dislike ratings are scalable at an acceptable level. The scalability of these sets of preference ratings is only marginally weaker than for the PTVs.

Table 6: *Unfolding scale of Thermometer ratings and Like/Dislike ratings*
 Data: INES2002

	Thermometer H_i (parties H (scale))	Like/Dislike H_i (parties H (scale))
FF	0.36	0.33
PDs	0.45	0.42
FG	0.38	0.36
Greens	0.37	0.42
Labour	0.45	0.41
SF	0.37	0.35
Scale	0.40	0.38

Unfolding all non-ipsative party preference ratings jointly

The similarity of the unfolding results for the three different measures of non-ipsative preferences is sufficiently striking to suggest that each of them could express the same kind of preference. A test of this hypothesis consists of jointly unfolding the items of PTV, thermometer and like/dislike ratings. Performing such a joint unfolding does not result in a single scale, however. More detailed analyses also lead to a resounding rejection of the hypothesis that all three instruments reflect the same latent phenomenon.¹⁸

The only conclusion that must be drawn from these failed attempts at finding a shared latent structure across the three measurements is that they do not measure the same thing. Warm and cold feelings as expressed in a thermometer scale do not express the same thing as do PTVs, nor do like/dislike ratings. Evidently, there are different kinds of preferences, and the logical next question is what differentiates them. We address that question by using each of the non-ipsative preference measures as the dependent variable, and estimating the effects a common series of independent variables in the hope that different patterns of relationships will elucidate the difference in the character of the three preference measures.

¹⁸ The most telling results were obtained from the following analyses. We analysed preferences for all 6 parties, using 5 preference scores from one of the three measurement instruments and one from one of the other instruments. There are 36 different ways in which this can be done with the available data. In each of these 36 instances, the entire scale was significantly weaker when compared to the unfolding scale with all 6 items from the same instrument (average loss of H was .08). In 30 out of 36 instances the H_i for one or several items dropped below the lower bound of 0.30, often much below that criterion. In all these cases the 6th item (from a different measurement instrument) was amongst the items dropping out of the scale.

Correlates of non-ipsative preferences

Our analyses so far established first that the PTV ratings are more valid measures of electoral utility than thermometer and likes/dislikes ratings, when using as validating criterion the concurrence of highest preference rating and party actually voted for. Second, we found that this difference cannot be attributed merely to a difference in noise. Had that been the case, preferences as indicated by the different questions should reflect the same underlying latent variable, but that was very clearly not the case. The different questions each reflect therefore a somewhat distinct kind of preference. In order to identify the nature of each of these preferences we assess their relationships with a number of other variables that can be expected to impinge on either affect or electoral utility.

The analysis design that we use is that of three different regression analysis with each of the preference ratings in turn as dependent variable, arranged in a ‘stacked’ format in which the units of analysis are respondent*party combinations, so that the analysis pertains simultaneously to intra- and inter-individual variation in preferences (see van der Eijk et al. 2006). We use here only a limited number of limited variables, including party identification (which we expect to be particularly indicative of affect), party size (indicative of electoral utility), distance between respondent and party in left/right terms (possibly more indicative of utility than of affect), credit or blame for the state of the economy (idem), and whether or not a party had contacted the respondent at his/her house (tentatively more relating to affect). In view of the strong correlations between the three different kinds of preferences we do not expect really strong contrasts between these analyses, but we do expect differences in the mix of factors associated with these preferences. The results are reported in Table 7.

Table 7 Regression coefficients for three non-ipsative preference measures

	PTV	Thermometer	Likes/Dislikes
L/R distance	-0.38	-0.30	-0.39
Party identification	3.31	2.65	2.83
Economy	0.18	0.21	0.24
House call	0.83	0.64	0.61
Party size*10	0.21	0.18	0.08

Note: all of these coefficients are significant at $> .01$ level.

The outcomes of these regressions indicate clearly that each of the preference questions is characterised by a distinct mix of covariates. In other respects they are less clear cut. We do find that PTVs are most strongly related to the non-affective variable party size, but also most strongly to the one that seems to us the most outspoken affective variable: party identification. If at face value thermometer scores (‘feelings’ in terms of ‘warm’ and ‘cold’) and ‘likes/dislikes’ appear the most affect-laden preferences, it is not clear at this stage why they are less strongly related to party identification than PTVs.

These outcomes do not clarify fully (let alone intuitively) what the specific character of the three kinds of preferences is, but whatever it is does generate a much greater concurrence between the highest preference rating and actual vote choice for the PTVs.

Conclusion and implications

We argued that non-ipsative preference ratings are used in the literature mainly as indicators of affect or of electoral utility. Assessing these preferences in terms of their construct validity for these different theoretical constructs in a convincing way requires that we can specify observable implications from these theoretical terms. We were able to do so for electoral utility. For preferences to be valid indicators of the electoral utility of parties, the highest of these preferences should be for the party actually chosen on the ballot. In terms of this criterion the PTVs are clearly more valid than the other preferences. Not only are they more valid, the absolute level to which this criterion is fulfilled is very high: 87% of first preference votes are cast for the party with the highest PTV, and this is 94% for voters who clearly see the election as a contest between parties rather than between candidates. The success rates of the other preference questions are considerably lower, and in the case of thermometer ratings this criterion holds for only 65% of the sample (and 70% for party-driven voters).

For validating the preferences as indicators of affect we could not specify a convincing observational implication, and hence no compelling empirical validating criterion. One could, in principle, think of the strength of party identification as such, were it not that the relevance of this concept outside the USA, and its operationalisation in multi-party systems is still very much contested (cf Crewe, Budge and Farlie 1976). Most compelling evidence could be derived from observations not commonly available, and impossible to obtain in mass survey settings, namely scans of brain activity known to be of an affective character in response to these kind of preference questions.

We could establish that the difference in performance of the three questions on the validating criterion for electoral utility were not generated by merely differences in noise. Would that have been the case, the different preferences should all reflect the same latent variable, and the unfolding analyses clearly refuted that. That leaves the conclusion that the three questions tap related, yet distinct and not interchangeable psychological phenomena.¹⁹ Our attempts to clarify the substantive differences between them by analysis of their covariates was only partially successful: PTVs in this analysis more strongly than other preferences reflect non-affective considerations such as party size,²⁰ yet they also most clearly reflect a seemingly affective factor such as party identification.

The excellent performance of the PTVs in terms of the empirical validating criterion is not unique for Ireland or for Ireland in 2002. In European countries we generally observe a success rate in this criterion in excess of 90% (cf. Van der Eijk et al 1996; Van der Eijk et al. 1999; Van der Brug et al. 2007b). Most of those analyses pertained to the second-order national elections for the European Parliament, but Dutch national election data showed that the measure performs equally well in first-order contexts (Van der Eijk and Niemöller 1984; Tillie 1995).

Only occasionally do election surveys contain more than a single form of non-ipsative measures, but in the few cases where that occurs and analysis could be carried out, similar results were found as reported in this paper. Kroh (2001) compared systematically the performance of thermometer ratings and PTVs in the Netherlands. His conclusions mirror ours: thermometer ratings are much more loosely connected to actual party choice than are propensity ratings, and therefore are much less valid as measures of electoral utility.²¹ It might be surprising that we found such a strong link between PTVs and party choice in the Irish context, as there is so much evidence that Irish voters are less party centred than voters in most other countries. However, these findings demonstrate that even if voters are drawn to a party via a candidate, they still tend to vote for that party which has the highest utility rating, and moreover, that such voters tend to give their second preference vote to party next highly rated on PTVs.

A general implication of our findings deserves specific mention. Using thermometer ratings or like/dislike ratings in analyses aiming to clarify the calculus of actual party choice involves serious threats to the (internal) validity of one's conclusions. The large amount of slippage between these ratings and actual choice means that no convincing predictions can be derived from these ratings about actual choice (as is commonly done, however, in analyses of directional voting, and comparisons of directional and proximity

¹⁹ From the unfolding algorithm a score can be calculated that represents respondents' positions on the underlying continuum. Unsurprisingly, the scores on these latent variables are significantly correlated. The strongest correlations are found between the PTV-based score and each of the other scores (0.66 and 0.62 respectively), whereas the scores on the latent dimensions expressed in thermometer and like/dislike ratings are intercorrelated at 0.57. Although these correlations are of a respectable magnitude, they also indicate that the shared variance of respondents' positions on these latent variables ranges somewhere between 0.32 and 0.43, which leaves ample room for mutually incompatible meaning components. In a later publication we will use these scores in a more elaborate latent-variable model explicating the relationships between these and other central variables in the explanation of electoral choice.

²⁰ Party size is actually a proxy for the perceived capacity of parties to influence the course of public policy, a factor of instrumental importance to voters (cf van der Eijk and Franklin 1996, 2009).

²¹ Using data from the 1994 Dutch Parliamentary Election Study he finds that 72.6% of the respondents vote according to the highest rating on the thermometer measure, while the corresponding percentage for the propensity rating was 94.6.

voting). Similarly, the use of non-ipsative preferences in the construction of theoretically relevant counterfactuals –e.g., how would votes and election outcomes have turned out had inflation been higher, or lower, cf van der Brug et al 2007) leads only to convincing and informative results if those preferences have a high degree of validity as measures of electoral utility.

References

- Budge, Ian, Ivor Crewe and David Farlie 1976. *Party identification and beyond: representations of voting and party competition*. New York: Wiley.
- Coombs, Clyde. 1964. *A Theory of Data*. New York: Wiley.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Brug, Wouter van der, Cees van der Eijk, and Mark N. Franklin. 2007a. *The Economy and the Vote*. Cambridge, MA: Cambridge University Press.
- Brug, Wouter van der, Cees van der Eijk, and Mark Franklin. 2007b. “EU Support and Party Choice,” in *European Elections and Domestic Politics. Lessons from the Past and Scenarios for the Future*, W. van der Brug and C. van der Eijk (eds.). Notre Dame: University of Notre Dame Press, pp.168-188.
- Brug, Wouter van der, Meindert Fennema, and Jean Tillie. 2000. “Anti-immigrant parties in Europe: Ideological or protest vote?” *European Journal of Political Research* 37: 77-102.
- van der Eijk, Cees, Wouter van der Brug, Martin Kroh, and Mark Franklin. 2006. “Rethinking the Dependent Variable in Voting Behavior – on the Measurement and Analysis of Electoral Utilities,” *Electoral Studies* 25: 424-447.
- van der Eijk, Cees, Mark N. Franklin, and Wouter van der Brug. 1999. “Policy Preferences and Party Choice,” in *Political Representation and Legitimacy in the European Union*, Hermann Schmitt, and Jacques Thomassen (eds.). Oxford: Oxford University Press, pp.161-185.
- van der Eijk, Cees, Mark N. Franklin, and Erik V. Oppenhuis. 1996. “The Strategic Context: Party Choice”, in *Choosing Europe? The European Electorate and National Politics in the Face of Union*, Cees van der Eijk, and Mark N. Franklin (eds.). Ann Arbor: The University of Michigan Press. pp.332-365.
- van der Eijk, Cees, and Erik V. Oppenhuis. 1991. “European parties’ performance in electoral competition”, *European Journal of Political Research* 19: 55-80.

- van der Eijk, Cees, and Kees Niemöller. 1984. "Het potentiële electoraat van de Nederlandse politieke partijen," *Beleid en Maatschappij* 11: 192–204.
- Kaplan, Abraham. 1964. *The Conduct of Inquiry. Methodology for Behavioral Science*. Scranton: Chandler.
- Kroh, Martin 2001. *An Empirical Analysis of the properties of the Probability of Future Vote Question (PFV) and the Party Sympathy Scale (PSS)*. Amsterdam: Amsterdam School of Communication Research (internal manuscript)
- Kroh, Martin, Wouter van der Brug, and Cees van der Eijk. 2007. "Prospects for Electoral Change," in *European Elections and Domestic Politics. Lessons from the Past and Scenarios for the Future*, W. van der Brug and C. van der Eijk (eds.). Notre Dame, IN: University of Notre Dame Press, pp. 209-225.
- Laver, Michael. 2000. "STV and the politics of coalition", in *Elections in Australia, Ireland and Malta under the Single Transferable Vote: reflections on an embedded institution*, Shaun Bowler and Bernard Grofman (eds.), Ann Arbor: Michigan University Press, pp. 131-152.
- Laver, Michael. 2004. "Analysing structures of party preference in electronic voting data", *Party Politics* 10::521–41
- Laver, Michael. 2005. "The electorate," in *Politics in the Republic of Ireland*, John Coakley and Michael Gallagher (eds). London: Routledge (4th ed.) ,
- MacDonald, Stuart Elaine, Ola Listhaug, and George Rabinowitz. 1991. "Issues and Party Support in Multiparty Systems," *American Political Science Review* 85: 1107–32.
- Marsh, Michael. 2004. "None of your post-modern stuff around here: grassroots campaigning in the 2002 Irish general elections" in *British Elections and Parties Review*, Roger Scully et al (eds.), Volume 14. Frank Cass: 245-67.
- Marsh, Michael. 2006 a. 'Stability and change in structure of electoral competition 1989-2002', in *Facing change in a new Ireland: An analysis of Irish social and political attitudes*, Diane Payne, John Garry and Niamh Hardiman (eds.) Dublin: Liverpool University Press.
- Marsh, Michael. 2006 b "Party identification in Ireland: An insecure anchor for a floating party system" *Electoral Studies* Volume 25(3) 2006, pp. 489-508
- Marsh, Michael. 2007. "Candidates or parties? Objects of electoral choice in Ireland", *Party Politics* 13 (2) 2007, pp.500-27

- Marsh, Michael, Richard Sinnott, John Garry and Fiachra Kennedy. 2008, *The Irish Voter: The Nature of Electoral Competition in the Republic of Ireland* Manchester: Manchester University Press.
- Powell, G. Bingham, Jr. 2000. *Elections as Instruments of Democracy*. New Haven, CT: Yale University Press.
- Rabinowitz, George, and Stuart Elaine MacDonald. 1989. "A Directional Theory of Issue Voting," *American Political Science Review* 83: 93–121.
- van Schuur, Wijbrandt H. 1993. Nonparametric unidimensional unfolding for multicategory data. *Political Analysis* 4, 41-74.
- van Schuur, Wijbrandt H., Kiers, Henk A. L. 1994. "Why Factor Analysis Often Is the Incorrect Model for Analyzing Bipolar Concepts, and What Model To Use Instead.", *Applied Psychological Measurement*, 18: 97-110
- Tillie, Jean. 1995. *Party Utility and Voting Behavior*. Amsterdam: Het Spinhuis.